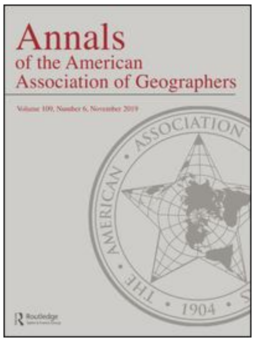

Annals of the American Association of Geographers

ISSN: 2469-4452 (Print) 2469-4460 (Online) Journal homepage: https://www.tandfonline.com/loi/raag21

Toward Smart Foodsheds: Using Stakeholder Engagement to Improve Informatics Frameworks for Regional Food Systems

Allan D. Hollander, Casey Hoy, Patrick R. Huber, Ayaz Hyder, Matthew C. Lange, Angela Latham, James F. Quinn, Courtney M. Riggle & Thomas P. Tomich

To cite this article: Allan D. Hollander, Casey Hoy, Patrick R. Huber, Ayaz Hyder, Matthew C. Lange, Angela Latham, James F. Quinn, Courtney M. Riggle & Thomas P. Tomich (2019): Toward Smart Foodsheds: Using Stakeholder Engagement to Improve Informatics Frameworks for Regional Food Systems, Annals of the American Association of Geographers, DOI: 10.1080/24694452.2019.1662764

To link to this article: https://doi.org/10.1080/24694452.2019.1662764

Published online: 16 Oct 2019.

Toward Smart Foodsheds: Using Stakeholder Engagement to Improve Informatics Frameworks for Regional Food Systems

Allan D. Hollander,* Casey Hoy,† Patrick R. Huber,‡ Ayaz Hyder,§ Matthew C. Lange,¶ Angela Latham,† James F. Quinn,* Courtney M. Riggle,‡ and Thomas P. Tomich#

*Department of Environmental Science & Policy, University of California, Davis
†Initiative for Food and AgriCultural Transformation, The Ohio State University
‡Agricultural Sustainability Institute, University of California, Davis
§Division of Environmental Health Science, College of Public Health and Translational Data Analytics Institute, The Ohio State University
¶International Center for Food Ontology Operability Data and Semantics, Department of Food Science, University of California, Davis
#Agricultural Sustainability Institute and Department of Environmental Science & Policy, University of California, Davis

A foodshed is a concept analogous to a watershed, describing the catchment of the sources of food for a region. As such, it portrays linkages ranging from local communities out to the global food system. Inefficiencies exist at all stages of the food supply chain, resulting in the challenges of inequitable access to healthy and safe food. Many of these inefficiencies are informational; for instance, food being wasted that could be donated to food banks were there communication of the need. These informational inefficiencies can be ameliorated by a stronger semantic characterization of the links between actors and resources in the food system, allowing for the development of smarter software technologies to facilitate interconnections. We discuss an iterative process to improve informatics frameworks for the foodshed by engaging with regional stakeholders to identify important issues and information needs. *Key Words: food systems, ontologies, semantic web, smart foodsheds, stakeholder engagement.*

食物域的概念与流域相近, 即一个地区食物来源的总和, 体现了小到地方社区、大到整个全球食物系统之间千丝万缕的关联。在食物供应链的每个阶段都存在效率低下的问题, 进而引发一个重大挑战：很多人无法公平获得健康安全的食物, 而问题的根源往往都源于信息不畅。例如, 被浪费掉的食物本可以捐赠给有需要的食物银行（贫穷群体或无家可归者可领取捐赠食物的地点）。有一种方法可以改善这种信息效率低下的现状：使用更强的语义描述食品系统中参与者和资源之间的联系, 开发更智能的软件技术促进相互关联性。在本文中, 我们探讨了一个致力于改善食物域信息框架的迭代过程, 通过区域利益相关者的参与确定重要事项和信息需求。*关键词：食物系统、本体论、语义网、智能食物域、利益相关者参与。*

Una cuenca alimentaria es un concepto análogo al de cuenca fluvial, que, para este artículo, describe la zona en donde están las fuentes alimentarias de una región. Como tal, ese tipo de cuenca retrata vínculos que van desde las comunidades locales hasta la totalidad del sistema global alimentario. Las deficiencias se dan en todas las etapas de la cadena de suministro de alimentos, lo cual desemboca en los retos sobre la desigualdad de acceso a alimentos saludables y seguros. Muchas de estas deficiencias son de carácter informativo; por ejemplo, el desperdicio de alimento que podría donarse a bancos alimentarios si hubiese existido comunicación sobre la necesidad. Estas deficiencias informativas pueden reducirse mediante una caracterización semántica más fuerte de los vínculos entre actores y recursos en el sistema alimentario, facilitando el desarrollo de tecnologías de software más inteligentes que faciliten las interconexiones. Nosotros discutimos un proceso iterativo designado para mejorar los marcos informáticos de la cuenca alimentaria involucrando a los interesados con el fin de identificar cuestiones importantes y necesidades de información. *Palabras clave: compromiso de interesados, cuencas alimentarias inteligentes, ontologías, red semántica, sistemas alimentarios.*

Annals of the American Association of Geographers, 0(0) 2019, pp. 1–12 © 2019 by American Association of Geographers
Initial submission, November 2018; revised submissions, April and June 2019; final acceptance, June 2019
Color versions of one or more figures in the article can be found online at www.tandfonline.com/raag
Published by Taylor & Francis, LLC.

The term *foodshed* has emerged as an analogy to a watershed in describing the food system. It was likely first coined in 1929 when Hedden used it to describe the catchment of food production surrounding New York City (Hedden 1929; Peters et al. 2009). More recently the term has entered into the discourse around localizing the food system (Feenstra 1997; Butler 2013). Horst and Gaolach (2015) highlighted how the term is seen in three different ways: first, spatially, imagining a relocalized food system; second, analytically, examining the data on a region's food production and consumption patterns; and, finally, as a starting point for action in organizing local food policy and cultural shifts.

Considerable discussion exists about whether localizing the food system generally contributes to food system sustainability (e.g., Weber and Matthews 2008; Butler 2013). Taking a foodshed perspective, though, leads to interesting new analyses concerning the range of interactions in the food system, as illustrated in a review of foodshed analyses in North America (Horst and Gaolach 2015). Such analytical approaches will lead to better characterization of flows across scales and improve our ability to address some of the inefficiencies in the food system. The foodshed concept provides a good framing because it is both local (a specific region) and global (linked to the global food system). Improved information technologies have a substantial role to play in increasing sustainability in food systems and agriculture (e.g., Gebbers and Adamchuk 2010; El Bilali and Allahyari 2018), so improved characterizations of foodsheds should lead to better utilization of information technology in the food system.

In this article we discuss the prospects for creating "smart foodsheds": using innovations in informatics to enhance the local food system. A smart foodshed would be one that could take advantage of a surfeit of data such as sensor and other Web-enabled data sources throughout the supply chain and analyze food system inefficiencies, having some of the following characteristics:

- *Traceability:* The smart foodshed would collect and synthesize data from sensors in the food supply chain, enabling tracking of food flows and mitigating food waste.
- *Transparency:* The smart foodshed would improve information flows linking producers with distributors and increase consumers' knowledge of food sources and nutrition.
- *Trust:* The smart foodshed would protect the needs for privacy and confidentiality of all actors.

As Maye (2018) elaborated, however, there is a tension between the notion of "smart cities," from which our concept of a smart foodshed derives, and the aims of the urban food movement. The smart city concept pulls together the trends of instrumentation becoming ever more pervasive in cities and the belief that a knowledge economy fosters socioeconomic progress. For Maye, the urban food movement, by contrast, broadly connects issues of food security, social justice, well-being, and environmental impacts, arguing that the work of creating a "smart food city" calls for attention to cultural and social innovations and practices as much as technological solutions (Kirwan et al. 2013). Next we describe how in creating a smart foodshed we aim to merge both technical and social innovations.

Semantic Technologies for Smart Foodsheds

Our approach for developing a smart foodshed begins with the project of developing a common semantic framework for information exchange. This approach builds on a tradition of studying food systems as networked entities, examples being Chiffoleau (2009), who examined social relationships in farmers' markets in Languedoc-Roussillion; Colloredo-Mansfeld et al. (2014), who used sketch mapping to study consumer knowledge of food stores; and, more recently, Trivette (2019), who used social network analysis to examine ties between local food producers and consumers in New England. A smart foodshed is by its nature transdisciplinary, creating a need for a shared semantic characterization across diverse stakeholders. As the notion of a shared semantic web (an extension of the web where the meaning of information is formalized, enabling machine interoperability) has grown from conceptualization (e.g., Berners-Lee, Hendler, and Lassila 2001) toward maturity (e.g., Zeng and Mayr 2018), several themes have emerged. These include formalization of ontologies as shared knowledge schemas, the use of graph data to portray information networks, and prospects for linking open data.

Ontologies

In computer science terms, an *ontology* is a knowledge organization system that formally describes the types of things that exist and the relationships that connect them in a particular domain; for example, foodsheds (Allemang and Hendler 2011). Ontological technologies provide a shared semantic framework for linking disparate pieces of information together and have become important in creating a coherent knowledge network to integrate data, ensure interoperability across systems, and enable inference across the relationships that connect these data.

Ontologies build on previous knowledge organization systems such as controlled vocabularies. In the domain of food and agriculture, much of the initial work here came from the Food and Agricultural Organization (FAO) of the United Nations, beginning with the development of the controlled vocabulary AGROVOC (Caracciolo et al. 2013) in the 1980s. More recently, a number of efforts seek to aggregate these information resources, such as the AgroPortal Map of Standards (see http://vest.agrisemantics.org/), which describes existing standards for the exchange of food and agricultural data, including classification schemes and ontologies.

Representing Information in a Graph Data Structure

Because a major way we now think about relationships between actors is through social networks, it follows that graph data structures composed of nodes and linkages have become important technologies. The semantic web formulation of graph data uses a model called the resource description framework (RDF; Schreiber and Raimond 2014). This model decomposes a graph into primitive data elements called *triples*, with each triple representing a single logical statement combining a subject, predicate, and object, an example being the statement "California [subject] produces [predicate] tomatoes [object]." These triples are linked into an ontological framework using standardized nomenclature for these resources (usually termed URIs, for uniform resource identifier; Schreiber and Raimond 2014). Vocabularies such as AGROVOC have published URIs for terms, enabling shared references to these terms across a global namespace.

Linking Open Data

The principles of linked open data (Berners-Lee 2006; Heath and Bizer 2011) are summarized under Berners-Lee's (2006) five-star open data scheme: (1) Value is added to information by making it available with an open data license; (2) data should be in a machine-readable format; (3) data should be available in a nonproprietary format; (4) data should use open nomenclature to identify things; and (5) data should be linked to other data sets to provide context. Data interlinking flourishes when knowledge organization systems such as vocabularies and ontologies become more interoperable (Zeng and Mayr 2018). In the context of a smart foodshed, data sets will span the range from being fully open (e.g., governmental agricultural production data) to necessarily proprietary (e.g., private financial data). Even in the case of proprietary data, though, structuring the data using consistent globally referenced nomenclature can bring advantages in terms of data set queries.

One major aim of developing knowledge organization systems is providing a shared information space to describe resources. Some vocabularies have been set up by standards organizations, but others have emerged more locally. These ontologies are composable: resources can be described by selecting elements from different ontologies. An individual resource can be catalogued in multiple classes and linked to other resources using properties coming from several ontologies. For instance, one might describe a person's relationships to organizations using properties from the VIVO ontology (intended to enable discovery of persons and resources in academia; Mitchell 2018) and describe a person's contact information using properties from the FOAF schema, an ontology describing persons, their activities, and relationships to other people (Graves, Constabaris, and Brickley 2007).

We would like to ensure that the ontologies we develop to characterize foodsheds mesh with existing ontologies, forming what we term a multiontology framework. In so doing, we aim to fill in gaps in ontology space, because there are domains of interest that are not well covered by ontologies. For instance, conservation planning is a domain with well-developed schemas for management (e.g., the Open Standards for the Practice of Conservation; Schwartz et al. 2012), but the domain has not developed an ontology.

Developing the Smart Foodshed Iteratively

Food systems can be viewed from at least two different perspectives. One focuses on supply chain activities spanning from agricultural input supply and

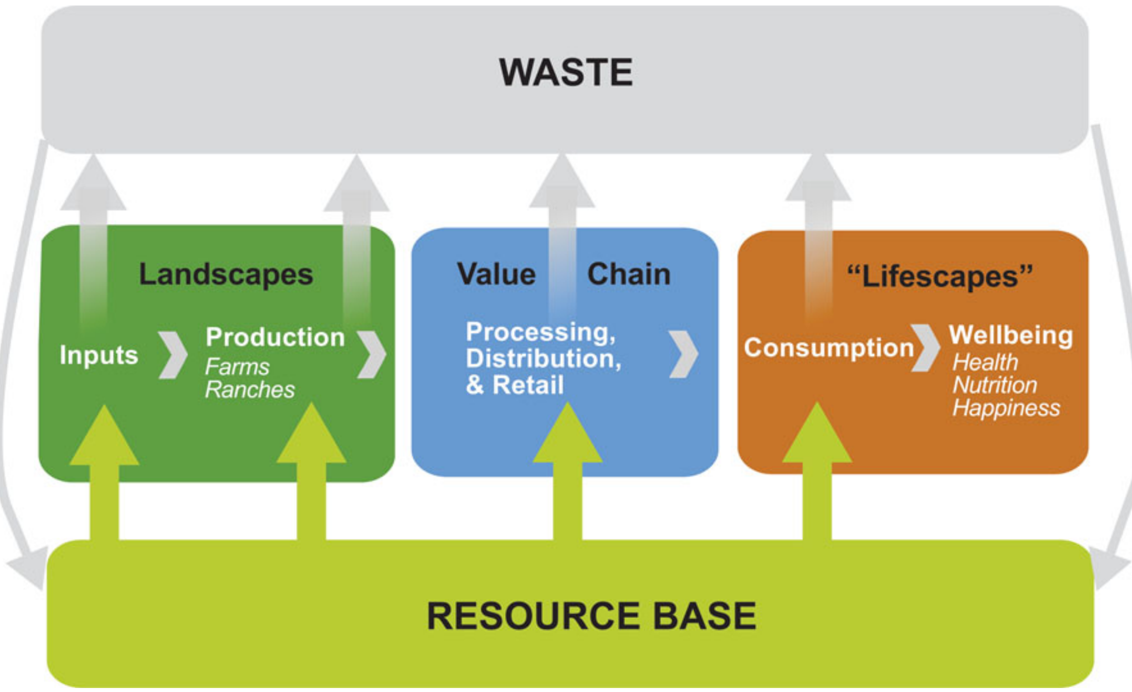

Figure 1. Food systems framing of food supply chains. From Tomich et al. (2018); adapted from a figure by Michele Grant, World Food Systems Center, ETH Zurich. Used with permission.

resource use for production through processing and marketing (Figure 1). On the supply side, these resource-intensive production activities are shaped strongly by their landscape context. Fundamental changes over the past generation (Reardon et al. 2018) require broadening our framing from fields and farms to the entire food system, extending to the health, nutrition, and overall well-being of people. Sociocultural lifescapes on the demand side of the food chain (Figure 1) are also grounded in place, an important component of local food movements described earlier. A second perspective focuses on social relationships within the food system. Even a "simple" typology of these food system stakeholders can be extremely complex (Figure 2), with most individuals holding multiple "stakeholder" identity types. Increasingly, relationships are approached in international practice as inclusive, multistakeholder processes spanning multiple spatial scales (Tomich et al. 2019). If a systems view is inescapable, the challenge becomes how to conceptualize, benchmark, and model (at least qualitatively) the myriad dimensions of economic, environmental, and social issues within these food systems (Huber et al. 2015; Springer et al. 2015). We aspire to create a food systems framework to link data sets ranging across many different data types, measurement units, and spatial scales, enabling development of more sustainable land use practices, improved human nutrition and health outcomes, and more resilient food systems.

We envision an iterative process to improve the informatics frameworks portraying the global food system via constructing knowledge schemas at the local foodshed scale through participatory work. We assert that constructing an overall model of the food system is a very difficult task, likely without a direct path forward to generate a formal ontology. Rather, we believe that better understanding of the global food system will emerge through improved interoperability of data, especially through linked open data protocols. This interoperability will be eased by adopting common vocabularies with referenceable URIs. Initiatives such as the Global Agricultural Concept Scheme (GACS; Baker et al. 2016) are already generating a pool of identifiers for data and metadata creators to draw on.

It is widely recognized that ontologies are efficiently created by iteratively working with subject matter experts (e.g., De Nicola, Missikoff, and Navigli 2009), where a knowledge engineer refines the ontology through engagement with subject-matter experts. We expand this methodology into a participatory design process involving local communities. This notion is that communities can

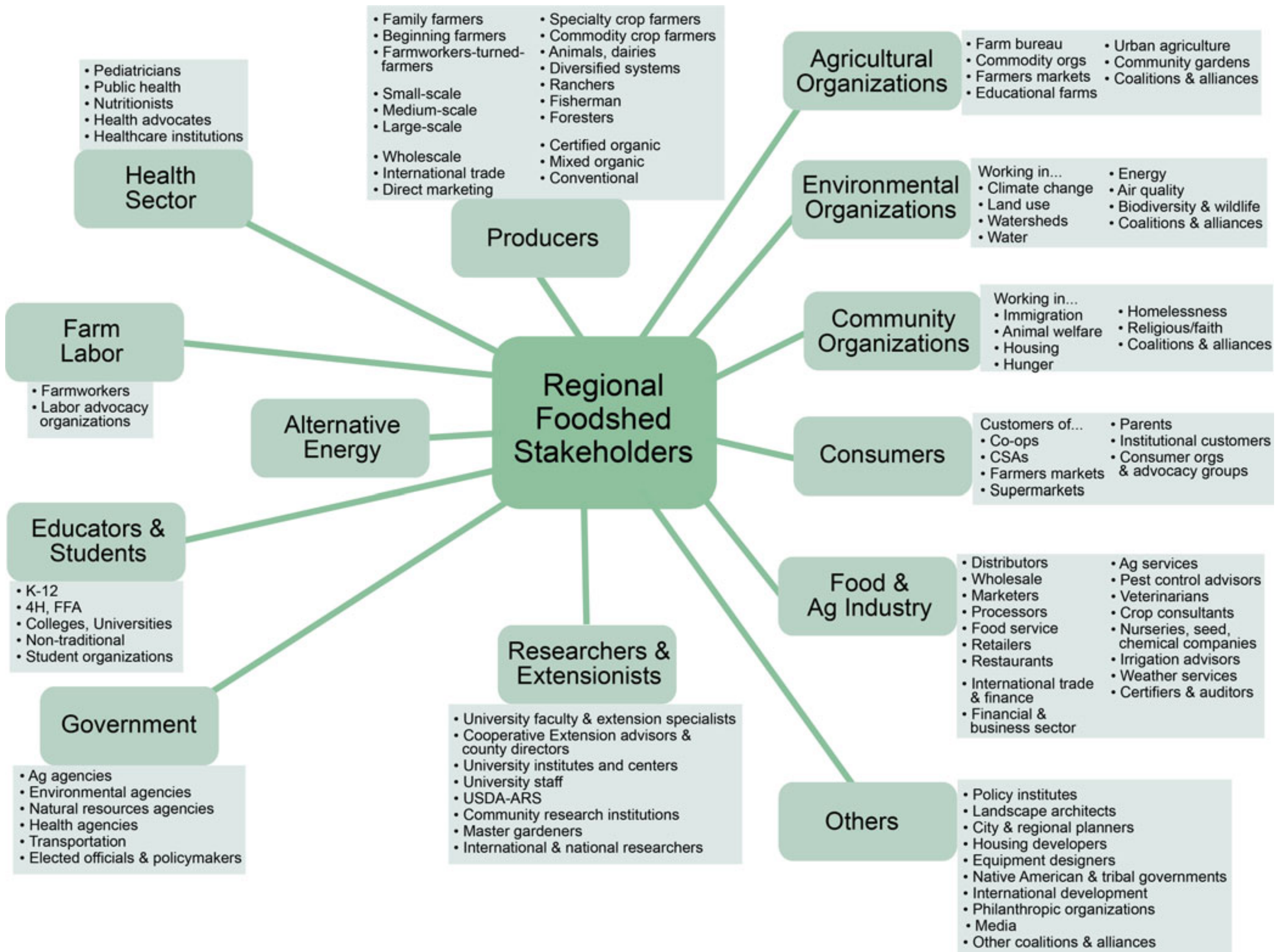

Figure 2. Stakeholders in the regional foodshed. USDA-ARS = U.S. Department of Agriculture, Agricultural Research Service; CSA = community-supported agriculture.

work with knowledge engineers to develop data schemas representing informatics concerns of local interest. An example of this process is as follows. A research group carries out interviews with stakeholders and devises a data template to code the interviews. Fields in this template are drawn, if possible, from properties in existing ontologies, but several new properties are needed to describe the content in the interviews. These properties are added into the data template, and the research group shares the template with other researchers carrying out similar projects in different cities. Later on, a knowledge engineer takes the additional properties defined in the template and extends existing ontologies with these. Data from these research projects can then be transformed into linked open data using common standards and made available as a global resource.

With these newly mobilized data and the additions to knowledge schemas describing the food system, we are able to refine our understanding of the food system. The linked open data graphs we create can be analyzed for connections through social network analysis, suggesting clusters of nodes describing resources. This in turn allows us to create better ontologies in the multiontology framework, thus continuing the iterative loop to improve the informatics framework around the food system, as well as allowing visualizations for new conceptualizations of the food system.

Approaches for Creating a Smart Foodshed

Data Sources for the Smart Foodshed

The smart foodshed can draw on a large variety of data sources, but these differ widely in their

Table 1. Examples of data sources for the smart foodshed

Sector	Data source	Accessibility to public	Degree to which data already exist	Ease of mobilization
Natural systems	Federal agency land, soil, biodiversity data sets (USGS, NASA)	High (public domain)	High	High
	State agency data sets; e.g., for biodiversity	Moderately high, although confidentiality for sensitive species	High	High
Agricultural production	USDA Census of Agriculture	Moderately high, confidentiality maintained at individual scale	High	High
	County crop statistics	High	High	High
	Sensor data tracking food flow through supply chain	Low (proprietary data)	Moderate (technology still developing)	Low
	Private-sector data on precision agriculture	Low (proprietary data)	High	Low
Physical infrastructure	Government agency transportation data	High	High	High
	Private-sector "Internet of Things" data on vehicular and rail transport	Low (proprietary)	Moderate	Low
Financial	Retail sales of goods and services	Low (proprietary)	High	Low
	Banking/loans for food system investment	Moderate (proprietary)	High	Moderate
Social and political	University/other research institution data on social and religious networks	High	Moderate (research generally limited in scope)	Moderate
Human health	Public health records	Moderate (confidentiality issues)	High	Moderate

Notes: USGS = U.S. Geological Survey; NASA = National Aeronautics and Space Administration; USDA = U.S. Department of Agriculture.

characteristics for use. Table 1 shows examples of sources, arranged by sector. The first column describes accessibility to the public. Data produced by the federal government are generally in the public domain, except for cases where confidentiality needs to be protected, an example being the Census of Agriculture, which is collected at the individual farm scale and then aggregated to counties to preserve confidentiality. At the other extreme are the proprietary data collected by many precision agriculture systems. Here we distinguish between proprietary data and confidential data, the former being related to commercial concerns and the latter to privacy issues.

The second column describes the degree to which data for each example already exist. In many cases data are being collected as part of long-standing surveys (e.g., agricultural censuses) and can be readily queried. Similarly, in the private sector, financial records (e.g., loan information) are maintained as a function of business operations and are potentially available for analysis. On the other end of this scale, some data are not presently generated. For instance, the flow of food from production through the supply chain creates data, but these data are currently too fragmented to track individual food items from farm to fork.

The final column summarizes the degree to which each source of data can be mobilized for use. Rankings in these columns are as follows. High are data sets that are available online or through straightforward enquiries. Moderate are data sets that are obtainable through some investment, such as negotiated agreements to maintain confidentiality. Low are data sets that are available only with high investment in negotiations or technical development. A couple of points emerge here. First, the scale of the data set relates to the degree to which it can be mobilized. Large compilations produced by state or federal governments are readily used but are generally only available at a coarse scale (e.g., county-level statistics). Conversely, universities have the capacity to generate fine-scale studies of the networks of actors in the food system but do not typically have the resources to expand these to a wide scope. A second point is that much data in the foodshed will remain private, whether due to confidentiality concerns or due to their proprietary nature. An informatics system that successfully links across a foodshed must address the boundaries between public and private data as part of its very design.

Efforts toward Stakeholder Engagement

A wide range of stakeholder individuals and organizations can be identified in association with a regional foodshed, as illustrated in Figure 2. This network of stakeholders in a foodshed is quite complex, and it is a significant challenge to engage with a representative subset of actors. An initial step here is to solicit information about data availability and needs, particularly focusing on themes of food system coordination, development of food system infrastructure, and elimination of food system waste. These stakeholders can be engaged through focus groups and by developing use cases.

Developing Use Cases for Interaction

The purpose of a use case is to explicitly capture how an end user will use or interact with a product. The concept of use cases comes from software development and product design (e.g., Adolph, Cockburn, and Bramble 2002), and we use it to illustrate how potential stakeholders in the foodshed can use developed tools and applications. The approach for a creating a use case involves the following steps: (1) develop a use case template, usually a set of open-ended questions; (2) conduct interviews of potential stakeholders; that is, the data providers and users in the foodshed; and (3) use the answers in the use case template to design tools for specific end users. From this gathered information we will be able to refine our existing knowledge schemas for the food system.

Creating an Ontology for Stakeholder Relationships

One element of creating a smart foodshed is to be able to describe in a computable manner the networks of relationships among stakeholders. This will enable development of software to connect actors and entities for applications that are suggested by the use case surveys. Our approach here is to develop an ontology for the network of relationships. We have named this ontology PPOD, for persons–projects–organizations–data sets (Hollander 2019). Most of the classes and properties in the PPOD ontology have been extracted from existing ontologies such as VIVO. These four classes have emerged as core groupings in our experience in working with the information-sharing activities of research and

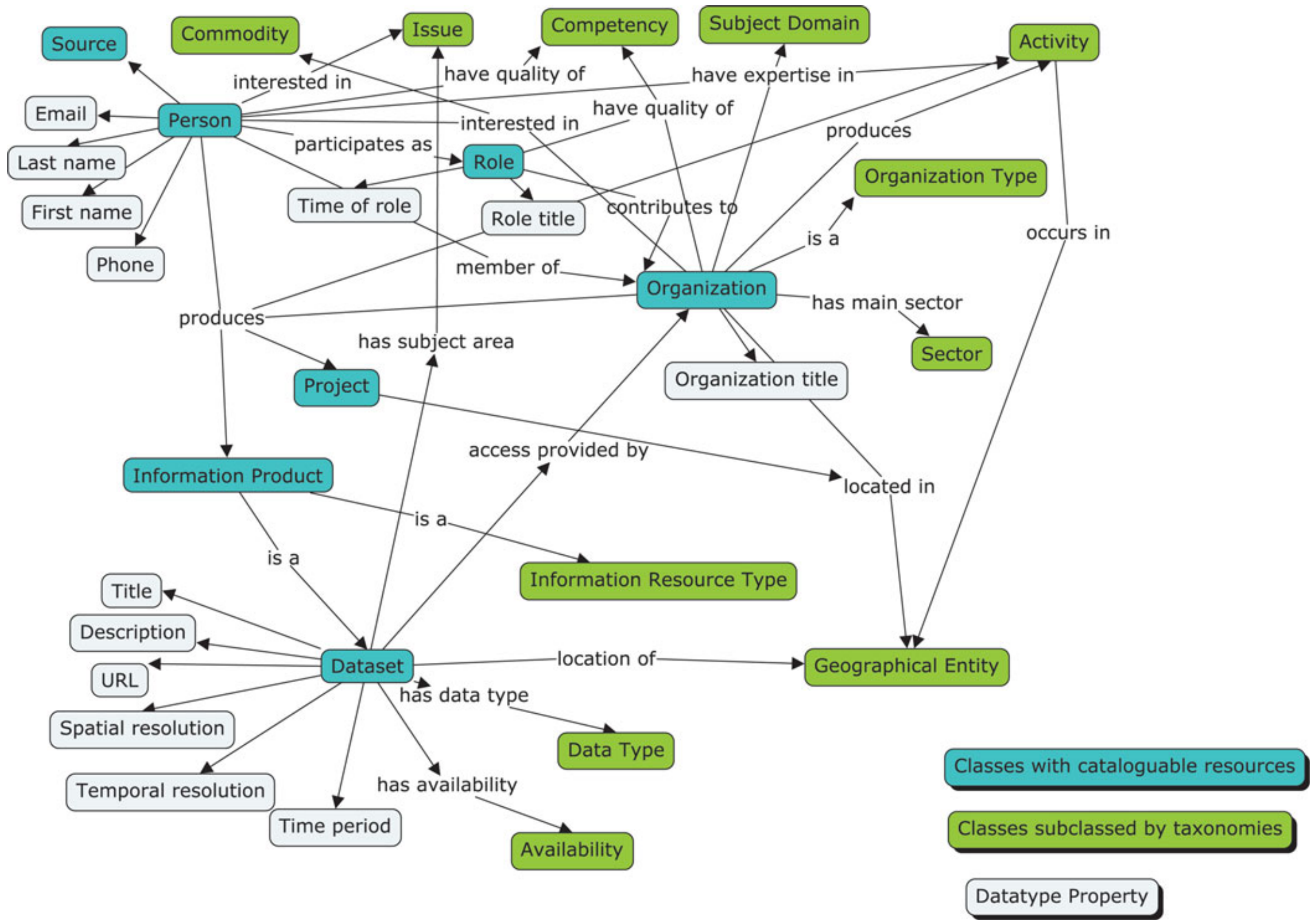

Figure 3. Concept map of the PPOD (persons–projects–organizations–data sets) ontology.

management communities. Figure 3 is a concept map illustrating the main structure of the PPOD ontology. Nodes in aqua are classes that we envision cataloguing, and nodes in green represent taxonomies. The nodes in gray are datatype properties (e.g., strings). Finally, labels annotating the links between nodes represent object properties. For instance, "located in" is an object property that links a project (a catalogable resource) with a geographical entity (a taxonomic resource coming from a gazetteer).

Participatory Design

It is important that technology for the smart foodshed has the support and interest of the stakeholders who would engage with it. Participatory design is a tradition dating back to the 1970s that could supply strategies for developing appropriate technologies (Bannon, Bardzell, and Bødker 2018). Gooch et al. (2018) discussed challenges and opportunities for applying participatory design to the smart city. Some of the difficulties here include that of scale: The original participatory design approach featured a limited set of people (workers in a single workplace): How does one extend the approach to include all the people in a city? The approaches that have been tried to date in the smart city movement have their challenges. For instance, the strategy of releasing open data has the difficulty that often these data lack sufficient context and description. Additionally, the smart city agenda is often disconnected from the concerns of socially disadvantaged communities. Gooch et al. (2018) developed a four-stage model for citizen participation in smart city development. The first is engaging with the community to identify its problems. The second and third steps are facilitating the collection of citizens' ideas and encouraging turning these into projects. The final step is ensuring enough investment in projects such that their success is sustainable.