sustainability

MDPI

Article

Workflows for Knowledge Co-Production—Meat and Dairy Processing in Ohio and Northern California

Allan D. Hollander [1,*], Casey W. Hoy [2], Kevin S. Armstrong [3], Michael R. Dimock [4], Patrick R. Huber [1], Douglas Jackson-Smith [3], James F. Quinn [5], Courtney M. Riggle [6] and Thomas P. Tomich [1]

1 Food Systems Lab, Department of Environmental Science and Policy, University of California, Davis, CA 95616, USA
2 Department of Entomology, Agroecosystems Management Program, Ohio State University, Wooster, OH 44691, USA
3 Agroecosystems Management Program, Ohio State University, Wooster, OH 44691, USA
4 Roots of Change, Public Health Institute, Oakland, CA 94607, USA
5 Department of Environmental Science and Policy, University of California, Davis, CA 95616, USA
6 International Center for Food Ontology Operability Data and Semantics (IC-FOODS), Davis, CA 95616, USA
* Correspondence: adhollander@ucdavis.edu

Abstract: Solving the wicked problems of food system sustainability requires a process of knowledge co-production among diverse actors in society. We illustrate a generalized workflow for knowledge co-production in food systems with a pair of case studies from the response of the meat and dairy production sectors in the wake of the COVID-19 pandemic. The first case study serves as an example of a scientific workflow and uses a GIS method (location allocation) to examine the supply chain linkages between meat and dairy producers and processors in Ohio. This analysis found that meat producers and processors are less clustered and more evenly distributed across the state than dairy producers and processors, with some dairy processors potentially needing to rely on supply from producers up to 252 km away. The second case study in California adds an example of a stakeholder workflow in parallel to a scientific workflow and describes the outcome of a series of interviews with small and mid-scale meat producers and processors concerning their challenges and opportunities, with the concentration of processors arising as the top challenge faced by producers. We present a pair of workflow diagrams for the two case studies that illustrate where the processes of knowledge co-production are situated. Examining these workflow processes highlights the importance of data privacy, data governance, and boundary spanners that connect stakeholders.

Keywords: knowledge co-production; assessment workflow; food systems; meat production; dairy production; stakeholder engagement; geographic information systems; boundary spanning

check for updates

Citation: Hollander, A.D.; Hoy, C.W.; Armstrong, K.S.; Dimock, M.R.; Huber, P.R.; Jackson-Smith, D.; Quinn, J.F.; Riggle, C.M.; Tomich, T.P. Workflows for Knowledge Co-Production—Meat and Dairy Processing in Ohio and Northern California. *Sustainability* **2023**, *15*, 9991. https://doi.org/10.3390/su15139991

Academic Editor: Jen-Yi Huang

Received: 1 May 2023
Revised: 10 June 2023
Accepted: 14 June 2023
Published: 23 June 2023

1. Introduction

How does one mobilize science to engage with difficult policy problems, such as the "wicked" problems of sustainability? The extensive literature on the relations of science to policy (e.g., [1,2]) suggests that solving sustainability issues necessitates a scientific enterprise that engages actively with society at large (e.g., [3]). There have been a wide variety of proposed approaches towards research that engages with society. Knapp et al. [4] identify nine approaches to such research that included: participatory action research, indigenous/local knowledge, knowledge with action, knowledge integration, transdisciplinarity, evidence-based practice, translational science, citizen science, and collaborative adaptive management. These approaches differ in their scale of application (local to global), the degree of power sharing between researchers and the community, the relative orientation between research and action, and the level of inclusivity. For instance, translational science, originating in the problem of how to translate basic research in biomedical science to clinical practice, is characterized by relatively low levels of power sharing and inclusivity. Knowledge with action approaches (e.g., [5,6]) emphasize processes of knowledge co-production through social learning, stakeholder collaboration, and establishing norms around knowledge governance. Because the food system is driven by and impacts a broad swath of social actors, solving sustainability problems in our domain can benefit from more inclusive knowledge co-production processes, as the range of actors in the food system ultimately includes all of society.

To highlight what this can look like we describe two case studies of workflows for food system knowledge co-production about the meat and dairy supply chains in Ohio and Northern California. Both case studies are located in regions of high value agricultural production and processing, and with large and economically varied communities. The term *workflow* refers to the series of steps involved in a knowledge production process, and has wide usage in informatics, especially describing information flows in business processes (e.g., [7,8]) and developing systems to automate data transformations in scientific research (e.g., [9,10]). These case studies are examples of an *assessment workflow* that combines both information processing aspects as in scientific workflows and methodologies for stakeholder engagement [11].

This pair of case studies has grown out of an NSF-funded Research Coordination Network on smart regional foodsheds. The two institutions anchoring this network (The Ohio State University and the University of California, Davis) are both land grant universities situated within rich but very different agricultural regions. The collaborators in this research network share expertise in food policy, experience with building information infrastructure, and connections to the broader community beyond the university. This research network was launched in part as a capacity building process across two states that could serve as a useful model for a much larger network encompassing similar organizations elsewhere. With a common background in information infrastructure, the group realized it would be valuable to start to develop coherent approaches to data identification, sharing, and research analyses for negotiation support and policy work in sustainability initiatives.

The work described here is part of a broader project [11] to develop the field of food system informatics. A number of themes described in [11] are especially pertinent. First is the recognition of "knowledge controversies", which stem from deficiencies in access to information and not just value conflicts. From that recognition, it follows that improved frameworks are needed for curating information. Examples of such frameworks include the development of ontologies to better represent materials involved in the food value chain (e.g., [12]). Second is the realization that management in the real world involves the interaction of many actors with multiple complex and often discordant goals [13]. If information systems are to be effective, they need to be developed through engagement of a diverse set of stakeholders. Also, because of this complexity, [11] find that understanding the food system has to come bottom-up through examination of use cases rather than using a top-down approach.

These case studies are used to illustrate the differences between a *scientific workflow* and a *stakeholder workflow*. The Ohio case is an example of the former, the Northern California case is primarily an example of the latter, though it includes some elements of a scientific workflow. In a scientific workflow, one starts with available data resources and derives insights into processes and issues through data analysis. In a stakeholder workflow, one seeks input from stakeholders to drive an interactive loop of knowledge acquisition and collaboration building. The motivation for beginning these studies was the COVID-19 crisis in early 2020, which heavily affected meat and dairy supply chains [14]. We were interested in ways in which new data and knowledge could be gathered and used to address bottlenecks, inefficiencies, and user-driven questions or requests to make these systems more resilient. The aim of comparing these two cases is to draw out the common themes of workflows used in sustainability efforts, the goal being to generalize major features of these workflows.

2. Materials and Methods

2.1. Ohio Case Study

The Ohio meat and dairy analysis originated as an effort to reduce the vulnerabilities in the food supply chain in Ohio that were revealed by the COVID-19 outbreak. In particular, disruptions in processing capacity, often due to illness or prevention of illness in the processing industry labor force, resulted in sudden and sometimes economically catastrophic loss of market access for producers. Informal discussions with producers, processors, and colleagues led researchers to ask how supply chains might be redirected if key processors or producers in the system were temporarily or permanently lost? In particular, we wanted to analyze whether and to what extent reconfigured supply chains at local and regional scales could maintain the flow of farm products from farm to consumers in Ohio. Meat and dairy supply chains were chosen for study because they are among the more highly regulated food industries, making study data about them relatively available, and because some of the most severe disruptions from COVID-19 were observed in meat and dairy processing. In March of 2020, available datasets were gathered describing food producers and processors to analyze the potential for a more resilient regional food supply chain in the face of current and future disruptions. Some of the questions of interest were:

- What areas of Ohio had the most and least capacity in meat and dairy production and processing?
- If one or more specific processing facilities were to shut down unexpectedly, how could supply chains be reconfigured quickly across multiple scales to take advantage of the remaining producing and processing capacities, while minimizing the distance that products would need to travel?

Our team identified several datasets from commercial data vendors and state and federal regulatory agencies that provided information about the size and location of meat/dairy producers and processors in Ohio. However, no single dataset contained all the needed information. The proprietary commercial datasets provided information on company size and sales, but could not be shared publicly, and ground truthing revealed numerous errors, usually miscategorized entities or missing information. State and federal regulatory datasets had more current and accurate business listings, but lacked information on sales, production, or processing volumes. By combining these sources of data, we were able to generate reasonably reliable maps of the meat and dairy supply chain actors in Ohio.

Unsurprisingly, while we were able to assemble data on the size, type, and location of most farmers and processors in Ohio, we were not able to find any systematic or public sources of information that listed the specific business relationships between individual producers and processors. As such, we could not create a map of the *actual* supply chain linkages (e.g., which farms sold to which processor). In the analysis below, we focused on identifying areas within Ohio where production and processing volume could be better connected by minimizing the distance between the two within Ohio. Although we know that Ohio meat and dairy farmers and processors often supply and trade with similar operations in other states and countries, we did not have the resources to acquire similar datasets for all surrounding states, so we did not attempt to extend the analysis to regional, national, or international supply chain partners.

Methodology

The Location-Allocation analysis tool in ArcGIS is designed to identify arrangements of two-point datasets that minimize travel distances from sources to destinations. In this case, we wanted to connect processors that provide a service and producers that utilize the services offered by the processor. For the dairy supply chain, the processor dataset included all dairy processing facilities in the state, and the producer dataset was composed of all commercial dairy farms. For the meat supply chain analysis, the processor dataset consisted of all meat slaughter, butchering and packing facilities and the production dataset included a list of farms that raised animals for meat.

Commercial and government regulatory datasets were both used as sources in developing the datasets for this analysis. First, we identified relevant records by North American Industry Classification System (NAICS) code in the Dun & Bradstreet Mergent Intellect (MI) database available through the Ohio State University (OSU) libraries. NAICS codes provide a standardized system for classifying businesses for economic censuses and statistical analysis [15]. The MI database was readily available to OSU personnel for research purposes, and included information that could be used to infer both business type and production or processing volume (e.g., employee counts, sales volume, and square footage). We found that the MI dataset included numerous errors, including incorrect business type (NAICS) codes and information on businesses that had closed or been sold. To improve the accuracy of these data, and to gather additional attributes of meat and dairy processing firms, data about licensed meat and dairy processors from the state and federal government agencies that regulate these industries in Ohio were obtained as described below.

Dairy Processors: The initial dairy processors dataset included all companies listed in the MI database with the NAICS code 3115 (Dairy Product Manufacturing). This was combined with a list of licensed dairy processors obtained through a public records request to the Ohio Department of Agriculture. The resultant dataset included 216 Ohio dairy processors.

Dairy Producers: The dairy producers dataset consists of all dairy farms included in the MI database with the NAICS code 112120 (Dairy Cattle and Milk Production). This list included 2012 Ohio dairy producers.

Meat Processors: The meat processors dataset combined information from the MI database (NAICS code 3116), a dataset of federally inspected meat processors from the USDA Food Safety and Inspection Service (FSIS), and a list of state-inspected meat processers obtained from the Ohio Department of Agriculture Division of Meat Inspection. The combined dataset included 606 meat processors in Ohio.

Meat Producers: No single authoritative public dataset is maintained on the size or location of all beef, sheep, hog and poultry farms in Ohio. To create a proxy list, the producers linked to NAICS codes between 112111 and 112420 (a range including meat, dairy, and poultry production) in the MI database were combined with farms on a list of Concentrated Animal Feeding Facilities (CAFF) that are permitted by the Ohio Department of Agriculture and Ohio EPA NPDES permit system (as compiled by the Ohio Environmental Council). The combined dataset included 4786 meat producers.

These lists were then combined to create a master list of processors and producers. A few firms in each category that were clearly miscategorized in the MI dataset based on background internet research were dropped from the analysis. Generally speaking, we only conducted the location analysis on facilities that we were confident fit the category (e.g., a meat processing firm) and where we had information about the volume of sales. When sales information was not available for a firm that appeared on the government regulatory databases, the median value from the rest of the observations was used to estimate sales volumes.

Dataset Pre-Processing Step: The first step in pre-processing was to fill in missing data, and verify the records were unique (e.g., not in both producer and processor lists). If the records were duplicated, inspection of available records was used to estimate the best categorization. The datasets were then stripped down to only essential fields and any records that did not have a sales value indicating the average amount of money they made over the last 5 years were assigned the median value for all the companies in the database that had values. A check was run to make sure that there were no records that appeared in both the processing and production datasets. Table 1 shows the numbers of firms included in the source files and the final combined files used for this analysis.

Table 1. Sources of data on meat and dairy producers and processors in Ohio that were combined for the full data set that was analyzed.

Source	Dairy Processors	Producers	Meat Processors	Producers
Mergent	251	2161	546	4784
ODA	152	1678	303	
USDA FSIS			193	
OEC/ODA CAFF				241
OEPA NPDES				31
Combined Full Data	216	2012	606	4786

Location Allocation Process: Once the datasets were compiled, a Python script was used to run the Location-Allocation tool in ArcGIS Pro. The tool was run using a transportation network dataset based on OpenStreetMap roads to identify the nearest producers who collectively would be able to provide the volume required to meet the requirements of each processor. A series of queries were created so that the contents of the datasets could be filtered to include or exclude different types of operations over the course of several runs of the Location Allocation tool to see the difference that certain inclusion criteria would have on the output of the tool. The results of Location Allocation tools were exported to File Geodatabase Feature Classes.

Maps were then drawn to identify the optimal arrangement of supply chain linkages between producers and processors that minimized the distance needed to supply Ohio processors with meat or milk from Ohio producers. These maps document a hypothetical and highly contrived set of supply chains only, but one that could serve as a starting point in engaging processor and producer stakeholders in dialogue on how to increase resilience in their supply chains.

2.2. Northern California Case Study

The Northern California case study is an example of a stakeholder engagement workflow methodology and comes from work carried out between Fall 2020 and Summer 2021 on challenges and opportunities facing small and mid-sized meat producers in the region [16]. The impetus for this study was the disruption facing the meat industry from the COVID-19 pandemic in 2020, but this disruption highlighted many long-standing challenges for the industry.

This case study began with a generic use case template developed to assess information and data needs from stakeholders in local or regional food systems. This use case template was created as part of an NSF-funded Research Coordination Network (RCN) on information infrastructure for smart regional foodsheds [17]. This template has four major sections: framing, data needs overview, detailed data needs, and data collection and implementation. It includes such questions as what is the user's mission, what challenges are they trying to address, and how can data help with the identified challenges. This template was modified to apply more specifically to the meat use case, focusing on elaborating the framing section and having less detail for the data needs section. This template asked interviewees to identify themselves, describe and rank the most important issues affecting animal production in California, name possible opportunities for addressing these challenges, describe important individuals, organizations, and coalitions in the system, and outline important information sources and needs (see Table 2).

Table 2. Questions for interviews of stakeholders involved in meat production and/or processing.

Category	Question
Mission:	(A) Please describe who you are and how you work with animal production, meat processing, distribution and/or retail. (B) What is your organization's primary purpose? (C) Where is your organization/operation located? (with zip code(s))
Challenges:	(A) What are the biggest issues impacting animal production and meat processing in California? (or more broadly?) (B) How are you/your members and your community impacted? (Economic, physical/mental health, animal health, farmland, environment) (C) How has COVID impacted the industry? You? (D) How has climate change (including fire, drought) impacted the industry? You? (E) How would you rank these issues/challenges by importance? (F) On what scale are these challenges best addressed (e.g., local, state or federal)?
Opportunities:	(A) Which of these challenges do you view as actionable? What opportunities/constraints do you see? (B) What are desired outcomes? (C) If the desired change(s) happen—who could be potential winners and losers?/How?
Stakeholders and other important participants in the system:	(A) Who can best address or act on these issues/challenges? Existing institutions or agencies? Individuals? Key collaborations or coalitions? (B) How might these issues/challenges be addressed? Acted on?) and why would those actions take place? (C) How long do you think it would take to see impacts or improvements (Ex: 1 year, 10 years, a generation) (D) What do you think the "quick wins" could be? (E) What are the biggest impediments to actions discussed above? (Ex: industry pressure, stakeholder lack of capacity, lack of information, funding, etc. (F) What/who are the thought leaders and critical voices to include in this conversation(s)? Who may potentially block progress?
Information:	(A) What kinds of information are needed to support evidence-based solutions and/or facilitate your community's action on any front with respect to these issues? (B) What information do you rely on? (How do you access it? Are particular platforms particularly important to you?) (C) What information is missing or hard to access? (D) What data sources might you have that are relevant and could potentially be shared?/Are there impediments to sharing this information? (E) Would you/your organization make better use of raw data, or more developed communication products? (F) What is the best way to communicate this data for your use? (Ex: Maps? Spreadsheets? Images & Graphics? Podcasts? Videos?)

After consulting with a stakeholder group convened by two non-profit organizations, the California Certified Organic Farmers (CCOF) and Roots of Change, 27 individual interviews were conducted with people representing a wide spectrum of activities and points of view within the meat supply chain, each interview lasting approximately one hour. These interview materials were organized using qualitative analysis tagging software (Taguette [18]) and network visualization software (Gephi [19]).

We read through the notes for each interview and developed a set of tags to describe the interview, categorized them into the following list: Challenges (58 tags), Opportunities (64), Commodity (8), Information (24), Roles (12), Stakeholders (27), and Zip Code (28). We linked the sets of tags in the Challenges and Opportunities categories to one or more sustainability issues from a formal list developed by the team in earlier research [20]. We then organized the Challenges and Opportunities tags by the primary issue, and ranked the issues by importance.

Based on the initial 27 interviews, over 50 more informants were interviewed to explore specific questions in greater depth. Separately, all 46 USDA-inspected slaughter facilities in California were called, and responses from 23 were obtained.

In a process parallel to the stakeholder interviews, available data resources on meat production and processing in Northern California were gathered. These included data from the 2017 Census of Agriculture on livestock production in Northern California at a county scale. The data also included information on meat processors from both a proprietary business database (the ReferenceUSA database, openly available via the Sacramento Public Library) and a list of accredited processors made available to us by the California Department of Food and Agriculture.

3. Results

3.1. Ohio Case Study

Figures 1 and 2 show the shortest possible supply chain linkages between meat and dairy producers and processors in Ohio. These dairy and meat supply chain maps highlight the location of both farm producers and processing facilities across Ohio. They represent a first step toward analyzing the spatial relationships that could be relevant to local and regional supply chains in these two important sectors.

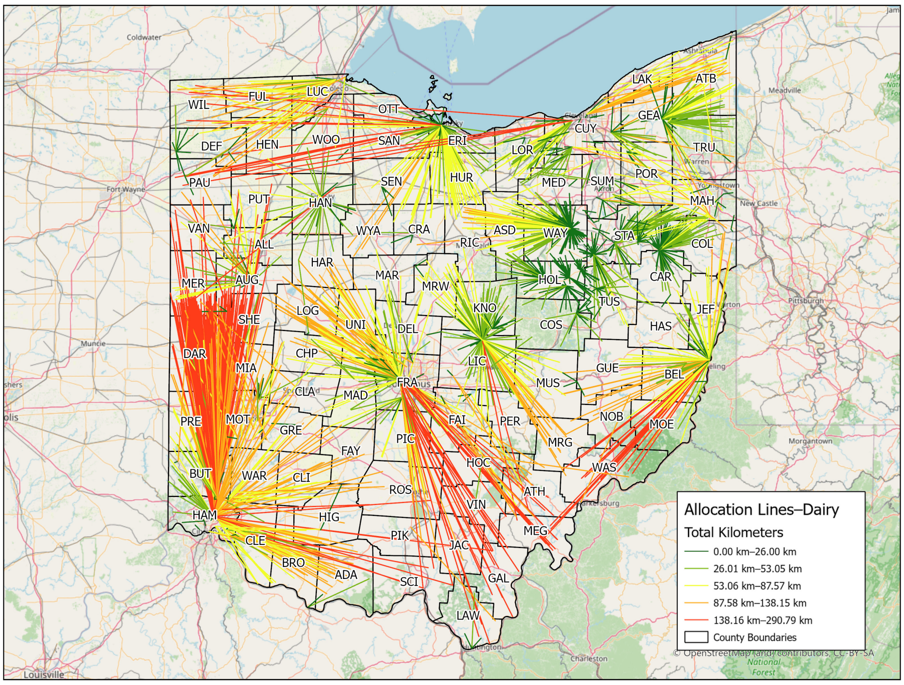

Figure 1. Dairy location allocation network analysis.

Clustering is evident in the distribution of both dairy farms and processors. For example, dairy farms are clustered in Northeastern and West Central Ohio, but can be found more sparsely distributed throughout the state. Dairy processing also is clustered, but not in a way that is spatially well-correlated with production. In some cases, Holmes County (Northeastern Ohio) for example, the production capacity is sufficient near clusters of processors to supply all of the milk that the processors can use from an average radius of 15 km and a maximum distance of 45 km. In other cases, for example Hamilton County (includes Cincinnati in Southwestern Ohio), processors would need to rely on supply from producers at an average distance of 136 km and as far as 252 km. It may be faster and easier to reconfigure the shorter and lower volume supply chains than the longer and higher volume ones after a disruption.

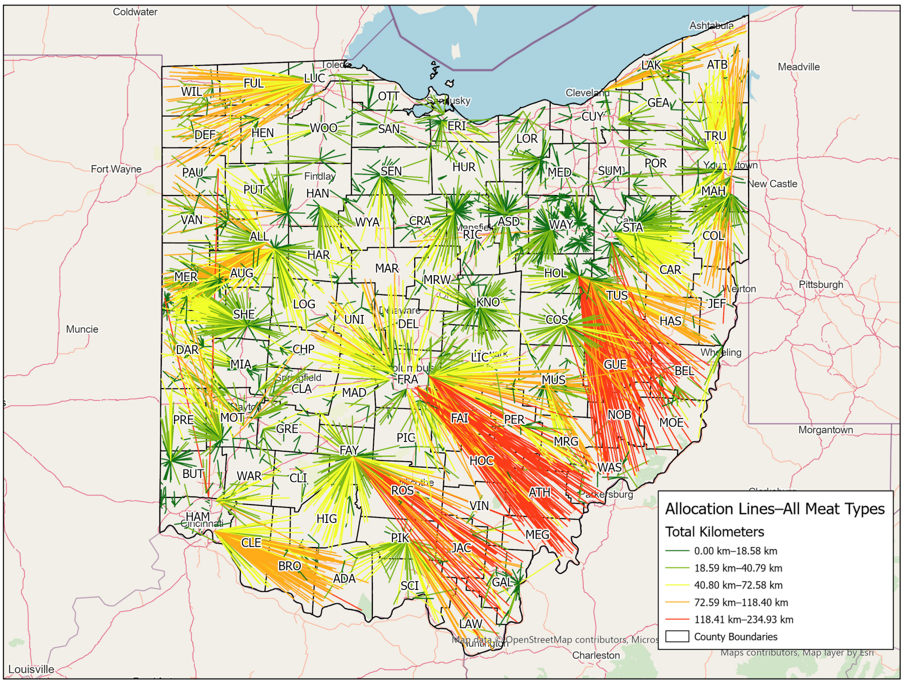

Figure 2. Meat location allocation network analysis.

Compared with dairy, meat producers and processors are less clustered and more evenly distributed across the state. In general, the meat production capacity could be connected with processing capacity at shorter distances than were observed in dairy, although it does vary, and greater distances would be required particularly for production in the Appalachian areas in the southeastern third of the state. For example, in Ashland County (Northcentral Ohio), the production capacity is sufficient near clusters of processors to supply all of the meat animals that the processors can use from an average radius of 17 km and a maximum distance of 42 km, whereas the processors in the Appalachian counties as a whole (Southeastern Ohio) would need to source from an average radius of 57 km and a maximum distance of 192 km. The Appalachian counties have some of the highest concentration of pasture ground that is suitable for grazing livestock, and a large number of producers, but relatively few processors in the area.

The dairy map illustrates how Ohio's urban areas (Cincinnati in the Southeast, Columbus in central Ohio, and Cleveland in the northeast) each have clusters of milk processors. The infrastructure required for milk handling and proximity to large consumer markets makes it more feasible for larger processors to operate in or near urban areas. Consolidation in the industry has also reduced the number or rural small dairy processors in recent years. As urban growth and development pushes livestock operations farther from the urban core, dairy processors are increasingly distant from local producers, and the distances required to transport milk and livestock in Ohio are increasing. By contrast, meat processors are more numerous (Table 1), tend to be smaller scale, and tend to be present across most areas where livestock are raised. Meat processing can be accomplished with less investment in infrastructure, making it possible for smaller processors to coexist with large meat processing facilities.

Dissemination

Graphic illustrations of the spatial arrangement of Ohio's meat and dairy supply chain actors highlighting the minimum distance analysis between Ohio production and processing firms were used to facilitate discussions among Ohio scientists and stakeholders about challenges and opportunities for a more resilient food supply chain for these commodities. The maps have also been shared with food system scholars and informatics scientists participating in the NSF-funded Intelligent Cyber Infrastructure with Computational Learning in the Environment (ICICLE) Artificial Intelligence (AI) Institute as an example of the kind of use cases that could be addressed with additional work in ontology, modeling and application of AI. The Ohio datasets and maps are currently hosted as content on ArcGIS Online, but because of privacy considerations associated with the use of the Mergent Intellect database, access is limited to employees of The Ohio State University.

3.2. Northern California Case Study

The Northern California case study revealed a long list of challenges and opportunities [16]. Based on the responses of the stakeholders in our structured interviews, we developed ranked lists to organize the challenges and opportunities. In order of importance, the top challenges included concentration of processors, drought and wildfire, high cost of doing business, negative public perceptions, NIMBYism, regulations governing water quality and land use, and public health risks. The ranked list of opportunities included creating more processing capacity, demonstrating value as a pathway to profitability, attracting capital, workforce development, rewards for ecosystem services, more innovative governance of food safety and waste management, and collaboration to address sociocultural factors.

A number of the key opportunities identified by stakeholders were interdependent, combining policy changes, new investments, and encouraging more entrepreneurship and innovation. There were variations in how informants viewed opportunities and even some contradictions among interviewees, but taken as a whole, we noted some promising pathways. A policy change was a commonly mentioned opportunity. The results were used by [16] to identify specific policy recommendations for meat processing reform in California to level the playing field for small- and mid-scale, high-value meat producers and cultivate diverse, innovative business models to seize emerging opportunities and achieve commercial viability in California's complex economy. These include specific steps to:

- Increase access to meat processing facilities by removing regulatory barriers, increasing inspection options, and more funding to upgrade, expand, and build new facilities.
- Expand market access for small ranchers by prioritizing public procurement of local, high-value meat and addressing market consolidation through improvements to the federal Grain Inspection, Packers, and Stockyards Act (GIPSA).
- Promote rancher-led solutions by creating a Small Meat Processing Innovation Advisory Panel comprised of producers, processors, financial experts, and academics to expand local meat processing capacity and expand options.

4. Discussion

We developed workflow diagrams to illustrate the steps taken in each of our two case studies. Figure 3 illustrates the steps in the Ohio case study. The upper portion shows the context for the study, concern among stakeholders and scientists about the impacts of COVID-19 on the regional meat and dairy supply chain. In the upper right in the diagram are the state and federal data sources used as input into the analysis, and in the upper left is the input from the business database. The geoprocessing steps are in the middle portion of the diagram, culminating in a series of map outputs noted at the bottom of the figure. The maps do not depict actual supply chains, which reflect more complex considerations that go well beyond distance between producers and processors (e.g., actual production and processing capacities, scheduling, product quality, production methods, transportation logistics, and policy at various scales). Instead, the generated display a picture of theoretically possible connections between Ohio producers and processors that minimizes travel distance and strengthens local and regional connections in the supply chain. These maps were designed to stimulate discussions about the potential for a more diverse set of local and regional supply chains that could provide increased flexibility during future crises. These research products were presented at a number of different venues with academic and non-academic partners as a starting point for stimulating transdisciplinary creativity for experimentation to improve the resilience of food supply chains.

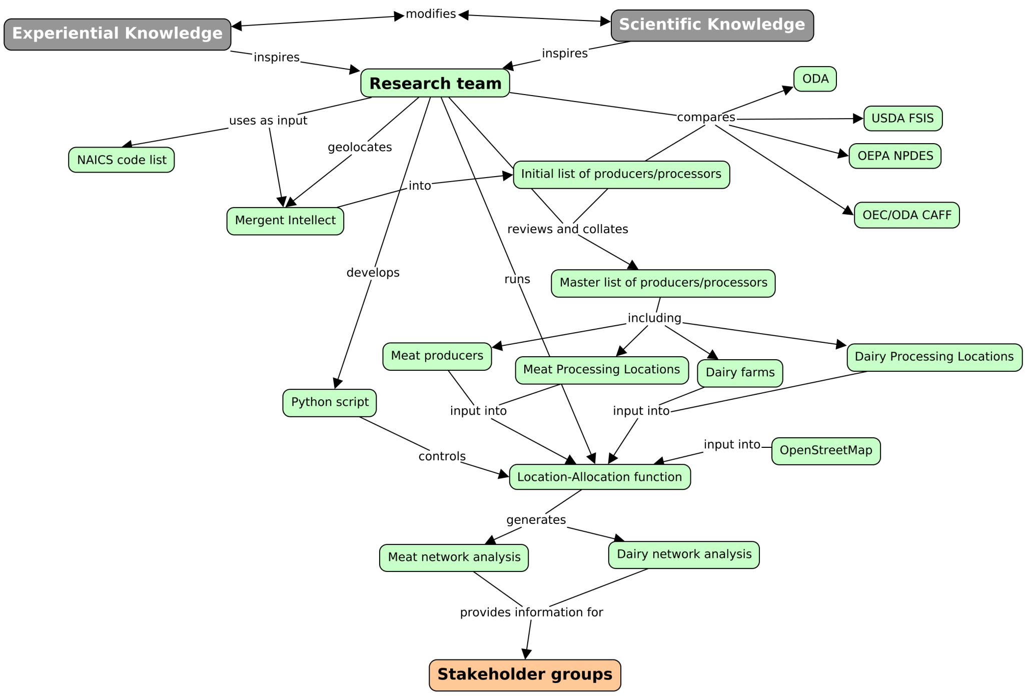

Figure 3. Workflow diagram for Ohio case study.

Figure 4 illustrates the workflow diagram for the Northern California case study. The figure is divided into several major portions. At the top of the diagram is the social and epistemological context of the work. On the right is the scientific workflow portion of the case study, highlighting the use of structured information resources to organize the research analysis cycle. These include working with a use case template, the use of ontologies to describe environmental issues, and utilizing tagging software to assist in qualitative analysis. The left portion of the diagram illustrates the stakeholder workflow. Major actors here include the research team, the funders, the boundary spanner [21], and the lead and key stakeholders. We highlight the role of boundary spanners, who are individuals or organizations that facilitate the exchange between the production and use of knowledge to support evidence-based decision-making [22]. In this portion of the diagram, a collaboration between the boundary spanner and the research team catalyzed the convening of key stakeholders and the formulation of template questions for stakeholder interviews. The research team drew from resources in the scientific workflow to analyze the content of the interviews. In consultation with the coalition of key stakeholders, the research team and the boundary spanner wrote a series of drafts of a white paper for the study. The output of the case study provided input into policy implementation workflows subsequently organized and launched by others.

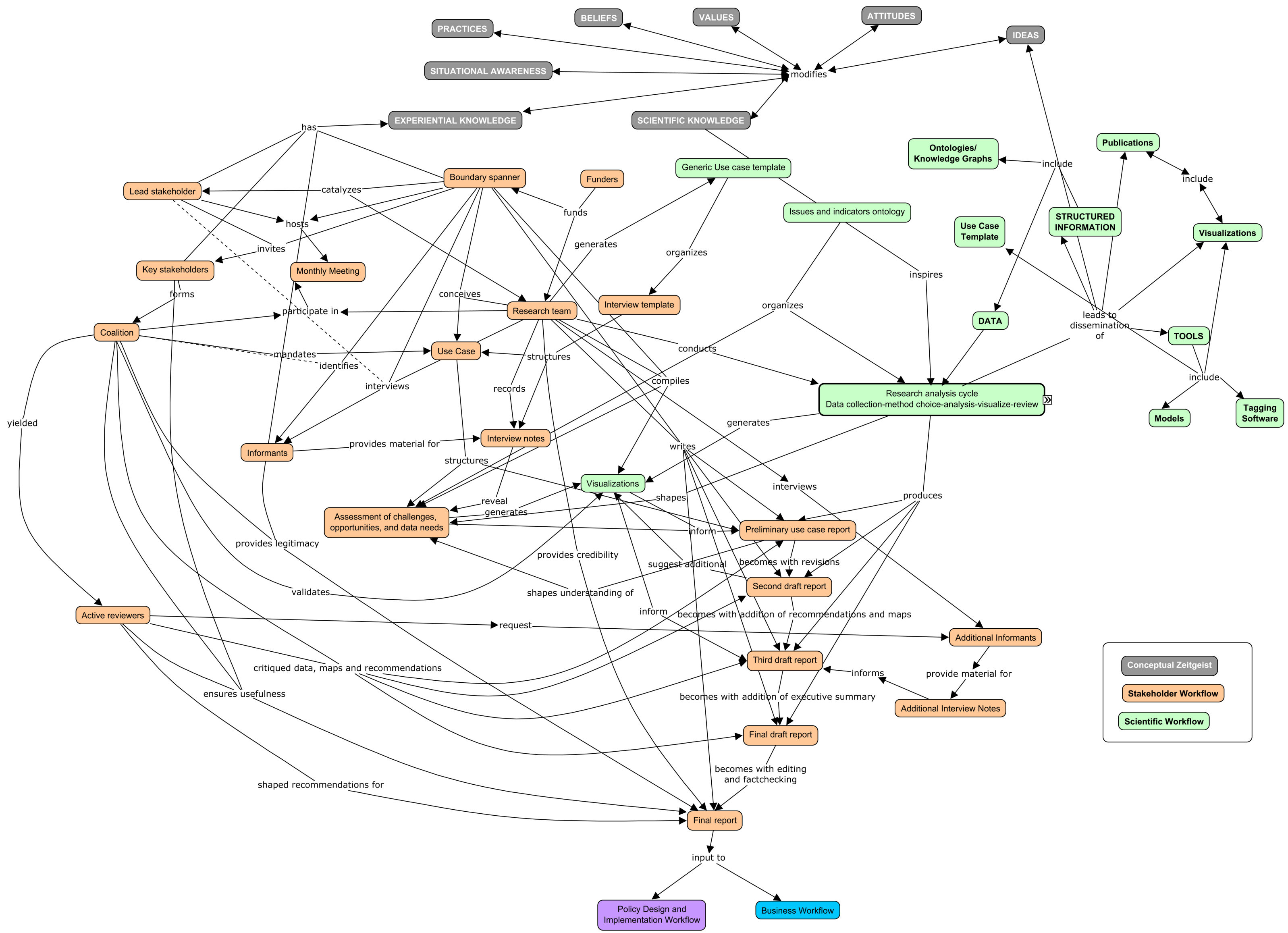

Figure 4. Workflow diagram for Northern California case study.

These two workflow diagrams are generalized into Figure 5, which is a simplification of Figure 3 in [11] and illustrates the main components in workflows for knowledge production. At the top is the overall context, on the left the stakeholder component, on the right the scientific component, and the outputs leading to policy and business processes.

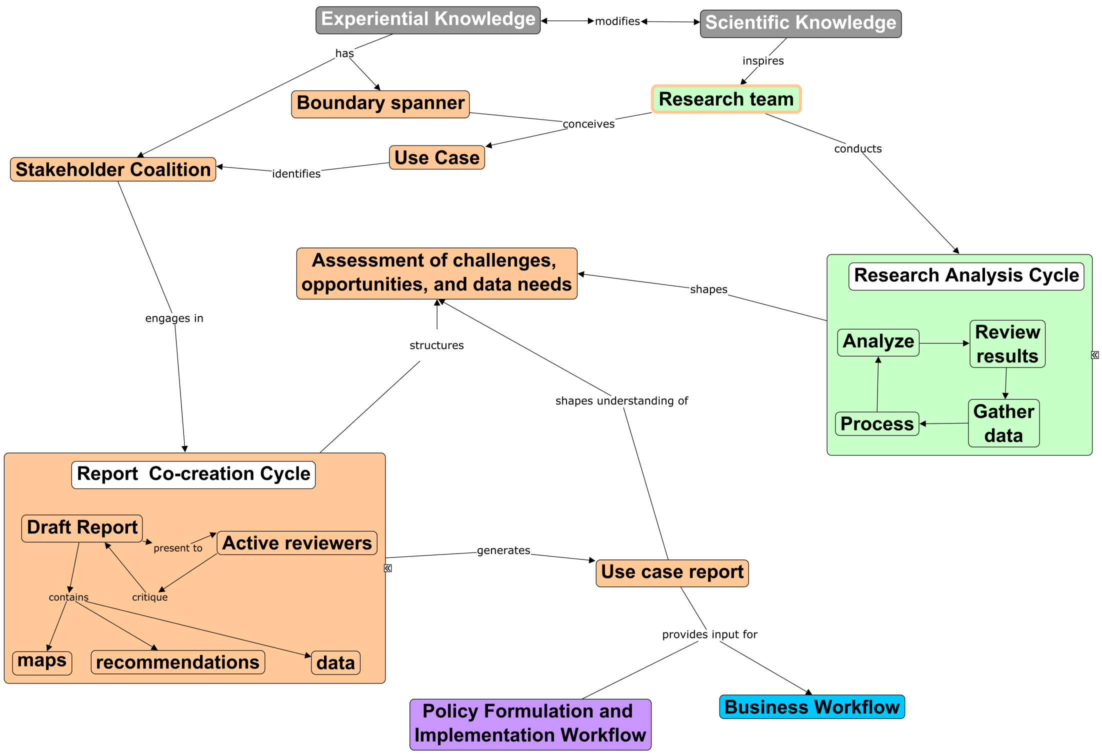

Figure 5. Generalized assessment workflow.

Bringing stakeholder workflow and scientific workflow components together to form a full assessment workflow is an ideal situation, with synergy emerging from the interactions between both sides. These interactions may be mediated by the boundary spanner, and facilitated by sharing visualizations produced by the scientific analysts. But often this synergistic ideal cannot be reached. A group may lack the analytical expertise to pursue a scientific workflow. Conversely, there is a social resource cost to convening a stakeholder workflow, and a group may not be able to engage the key people in a stakeholder network. However, in pursuing only a scientific workflow, a group will be limited by the history and constraints of what data have already been collected, such as US Census of Agriculture data, available only in aggregated form at the county scale, or state regulatory data that do not include complete information about the scale or throughput of processor firms.

In this paper, two case studies are used to highlight the value of visualizing the workflows associated with any knowledge generation project (because workflows happen whether or not they are made explicit). Creating a visual workflow helps identify the relationships between the sociotechnical context in both research and stakeholder engagement are embedded. They also can point to opportunities to develop more participatory processes for knowledge co-production, leading to more information being available to a wider range of stakeholders.

That said, maintaining privacy around data was a concern for many of our stakeholder partners. Privacy issues already restrict the types of data that are publicly available on meat and dairy supply chains. In both case studies, data limitations arose because information is not being collected or released at the farm or firm scale for privacy reasons—for instance, in Northern California, we would like to know more about the distribution of livestock production and transport at approximately a kilometer scale.

There are many concerns about the rise of big data in agriculture, for instance producers being concerned that sharing data will lead to price discrimination (e.g., [23]). There is often a tension between information privacy and the public good.

There are also important equity implications in how data are gathered, questions are framed, and knowledge production is organized. These include procedural dimensions (inclusion of all stakeholders in decision-making), distributional dimensions (equitable distribution of costs, benefits, rights, responsibilities, and risk), and recognitional dimensions (respect for knowledge systems and values of all stakeholders) [24].

Going forward, we see a need for more explicit attention to the details of data governance. These details might include the provision of private benefits from aggregating data for the participating stakeholders, shared leadership in setting standards between the public sector and industry, and cooperative ownership of organizations that collect and manage data [25]. It is also important to pay attention to sociocultural issues in the design and implementation of new data and information systems. There are many cases in agricultural development where inadequate understanding of society and culture led to only the limited uptake of a particular innovation (e.g., [26]) or unintended scale biases and other distributional consequences from emerging technologies [27].

Conversely, participatory engagement by stakeholders can be a way forward in bridging sociocultural gaps associated with data collection, analysis, and innovation. This may include the explicit use of boundary spanners to communicate needs and wishes between scientists and stakeholders [21,28,29]. In developing new technologies, it is important to bring stakeholders into the process early on, otherwise tools and research products tends to fit the interests of the developers rather than stakeholders or users [30]. There are many different approaches to knowledge co-production, and social innovation can occur through overlaps between stakeholder networks, user networks, developer networks, and core communities [31]. It is important to identify emergent patterns so as to bring them forward for effective use by neighboring communities [32]. Recognizing workflow patterns is an important step in this process.

5. Conclusions

We believe we have developed a coherent framework illuminating a key aspect of sustainability assessments, the relationship between stakeholder workflows and scientific workflows, as illustrated in Figure 5. These two types of workflows may be differentiated by the importance of a report creation cycle facilitated by a boundary spanner and a research analysis cycle for the latter. In an ideal situation, a sustainability assessment would involve both sides of this framework, but a group may emphasize one or the other side depending upon its expertise and access to social networks. We do not provide a thorough treatment of the factors affecting which information is hard to access, though privacy is clearly an important concern in both workflows. We also do not examine the details internal to policy formulation and implementation workflows and business workflows, which are activities downstream from the assessment, nor do we look at how the intellectual context (experiential knowledge and scientific knowledge on the figure) affects the directions the assessment takes.

The research network in which we have created this framework has already broadened from its inception as a collaboration between two academic groups. Developing informatics elements in smart foodsheds through building and visualizing knowledge graphs is now a thrust of the aforementioned NSF-funded ICICLE AI institute [33], and collaborators from an additional land grant university (the University of Wisconsin-Madison) are part of that institute's smart foodsheds working group. The next steps for building the framework include creation of ontologies to formalize the relationships it illustrates, thereby facilitating the construction of knowledge graphs describing an increasingly diverse portion of the food system.

This work is a contribution to a body of literature emphasizing the importance of participatory engagement in creating new socio-technical systems for community development (e.g., [30–32]). As information systems are developed to address the needs of the food system community [11], they will be greatly improved by finding the synergies between scientific and stakeholder communities.

Author Contributions: Conceptualization, A.D.H., C.W.H., D.J.-S. and T.P.T.; methodology, C.W.H., T.P.T., P.R.H., K.S.A., A.D.H., C.M.R. and M.R.D.; software, A.D.H. and K.S.A.; validation, M.R.D., C.M.R., K.S.A. and C.W.H.; formal analysis, C.W.H., T.P.T., P.R.H., K.S.A., A.D.H., C.M.R. and M.R.D.; investigation, M.R.D., A.D.H., C.M.R., P.R.H. and K.S.A.; resources, C.M.R., T.P.T., M.R.D., K.S.A., C.W.H. and D.J.-S.; data curation, A.D.H. and K.S.A.; writing—original draft preparation, A.D.H. and K.S.A.; writing—review and editing, T.P.T., C.W.H., D.J.-S., P.R.H., C.M.R. and J.F.Q.; visualization, A.D.H. and K.S.A.; supervision, T.P.T., C.W.H. and D.J.-S.; project administration, C.M.R., T.P.T., C.W.H. and D.J.-S.; funding acquisition, T.P.T., C.W.H. and M.R.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the US National Science Foundation, NSF Award 1737573 as part of a Research Coordination Network project entitled "Developing Informational Infrastructure for Building Smart Regional Foodsheds". The Northern California case study also received additional funding from the TomKat Foundation and USDA Hatch Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Usage of the Ohio case study maps and datasets is restricted to employees of The Ohio State University because of privacy concerns with the Mergent Intellect database. Data from the Northern California case study are freely available upon request.

Acknowledgments: We would like to thank Matthew Lange for his guidance in producing the workflow diagrams.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
FSIS	USDA Food Safety and Inspection Service
GIS	Geographic Information System
ICICLE	Intelligent Cyberinfrastructure with Computational Learning in the Environment
MI	Mergent Intellect
NAICS	North American Industry Classification System
OSU	The Ohio State University
RCN	Research Coordination Network

References

1. Pielke, R.A., Jr. *The Honest Broker: Making Sense of Science in Policy and Politics*; University Press: Cambridge, UK, 2007.
2. Gibbons, M.; Limoges, C.; Nowotny, H.; Schwartzman, S.; Scott, P.; Trow, M. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*; SAGE Publications: London, UK, 1994.
3. Birke, F.M.; Knierim, A. ICT for Agriculture Extension: Actor Network Theory for Understanding the Establishment of Agricultural Knowledge Centers in South Wollo, Ethiopia. *Inf. Technol. Dev.* **2020**, *26*, 591–606. [CrossRef]
4. Knapp, C.N.; Reid, R.S.; Fernández-Giménez, M.E.; Klein, J.A.; Galvin, K.A. Placing Transdisciplinarity in Context: A Review of Approaches to Connect Scholars, Society and Action. *Sustainability* **2019**, *11*, 4899. [CrossRef]
5. Cornell, S.; Berkhout, F.; Tuinstra, W.; Tàbara, J.D.; Jäger, J.; Chabay, I.; de Wit, B.; Langlais, R.; Mills, D.; Moll, P.; et al. Opening up Knowledge Systems for Better Responses to Global Environmental Change. *Environ. Sci. Policy* **2013**, *28*, 60–70. [CrossRef]
6. Clark, W.C.; van Kerkhoff, L.; Lebel, L.; Gallopin, G.C. Crafting Usable Knowledge for Sustainable Development. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 4570–4578. [CrossRef]
7. Georgakopoulos, D.; Hornick, M.; Sheth, A. An Overview of Workflow Management: From Process Modeling to Workflow Automation Infrastructure. *Distrib. Parallel Databases* **1995**, *3*, 119–153. [CrossRef]
8. Van der Aalst, W.M. Business Process Management: A Comprehensive Survey. *Int. Sch. Res. Not.* **2013**, *2013*, 507984. [CrossRef]
9. Sethi, R.J.; Gil, Y. Scientific Workflows in Data Analysis: Bridging Expertise across Multiple Domains. *Future Gener. Comput. Syst.* **2017**, *75*, 256–270. [CrossRef]

10. Ludäscher, B.; Altintas, I.; Berkley, C.; Higgins, D.; Jaeger, E.; Jones, M.; Lee, E.A.; Tao, J.; Zhao, Y. Scientific Workflow Management and the Kepler System. *Concurr. Comput. Pract. Exp.* **2006**, *18*, 1039–1065. [CrossRef]
11. Tomich, T.P.; Hoy, C.; Dimock, M.R.; Hollander, A.D.; Huber, P.R.; Hyder, A.; Lange, M.C.; Riggle, C.M.; Roberts, M.T.; Quinn, J.F. Why Do We Need Food Systems Informatics? Introduction to This Special Collection on Smart and Connected Regional Food Systems. *Sustainability* **2023**, *15*, 6556. [CrossRef]
12. Medici, M.; Dooley, D.; Canavari, M. PestOn: An Ontology to Make Pesticides Information Easily Accessible and Interoperable. *Sustainability* **2022**, *14*, 6673. [CrossRef]
13. Anderies, J.M.; Mathias, J.D.; Janssen, M.A. Knowledge Infrastructure and Safe Operating Spaces in Social–Ecological Systems. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 5277–5284. [CrossRef]
14. Marchant-Forde, J.N.; Boyle, L.A. COVID-19 Effects on Livestock Production: A One Welfare Issue. *Front. Vet. Sci.* **2020**, *7*, 585787. [CrossRef]
15. North American Industry Classification System (NAICS) U.S. Census Bureau. 2023. Available online: https://www.census.gov/naics/ (accessed on 30 April 2023).
16. Dimock, M.R.; Riggle, C.; Hollander, A.; Huber, P.; Tomich, T. A New Era for Meat Processing in California? Challenges and Opportunities to Enhance Resilience. 2021. Available online: https://escholarship.org/uc/item/4r723374 (accessed on 30 April 2023).
17. Hollander, A.D.; Hoy, C.; Huber, P.R.; Hyder, A.; Lange, M.C.; Latham, A.; Quinn, J.F.; Riggle, C.M.; Tomich, T.P. Toward Smart Foodsheds: Using Stakeholder Engagement to Improve Informatics Frameworks for Regional Food Systems. *Ann. Am. Assoc. Geogr.* **2020**, *110*, 535–546. [CrossRef]
18. Rampin, R. Taguette. 2020. Available online: https://www.taguette.org/ (accessed on 30 April 2023).
19. Bastian, M. Gephi. 2021. Available online: https://gephi.org/ (accessed on 30 April 2023).
20. Springer, N.P.; Garbach, K.; Guillozet, K.; Haden, V.R.; Hollander, A.D.; Huber, P.R.; Ingersoll, C.; Langner, M.; Lipari, G.; Mohammadi, Y.; et al. Sustainable Sourcing of Global Agricultural Raw Materials: Assessing Gaps in Key Impact and Vulnerability Issues and Indicators. *PLoS ONE* **2015**, *10*, e0128752. [CrossRef] [PubMed]
21. Goodrich, K.A.; Sjostrom, K.D.; Vaughan, C.; Nichols, L.; Bednarek, A.; Lemos, M.C. Who Are Boundary Spanners and How Can We Support Them in Making Knowledge More Actionable in Sustainability Fields? *Curr. Opin. Environ. Sustain.* **2020**, *42*, 45–51. [CrossRef]
22. Bednarek, A.T.; Wyborn, C.; Cvitanovic, C.; Meyer, R.; Colvin, R.M.; Addison, P.F.E.; Close, S.L.; Curran, K.; Farooque, M.; Goldman, E.; et al. Boundary Spanning at the Science–Policy Interface: The Practitioners' Perspectives. *Sustain. Sci.* **2018**, *13*, 1175–1183. [CrossRef] [PubMed]
23. Sykuta, M.E. Big Data in Agriculture: Property Rights, Privacy and Competition in Ag Data Services. *Int. Food Agribus. Manag. Rev.* **2016**, *19*, 57–74.
24. Clements, H.S.; Selinske, M.J.; Archibald, C.L.; Cooke, B.; Fitzsimons, J.A.; Groce, J.E.; Torabi, N.; Hardy, M.J. Fairness and Transparency Are Required for the Inclusion of Privately Protected Areas in Publicly Accessible Conservation Databases. *Land* **2018**, *7*, 96. [CrossRef]
25. Hutchins, J.; Hueth, B. 100 Years of Data Sovereignty: Cooperative Data Governance and Innovation in US Dairy. *Appl. Econ. Perspect. Policy* **2023**, *in press*. [CrossRef]
26. Díaz Andrade, A.; Urquhart, C. The Affordances of Actor Network Theory in ICT for Development Research. *Inf. Technol. People* **2010**, *23*, 352–374. [CrossRef]
27. Klerkx, L.; Jakku, E.; Labarthe, P. A Review of Social Science on Digital Agriculture, Smart Farming and Agriculture 4.0: New Contributions and a Future Research Agenda. *NJAS-Wagening. J. Life Sci.* **2019**, *90–91*, 100315. [CrossRef]
28. Posner, S.M.; Cvitanovic, C. Evaluating the Impacts of Boundary-Spanning Activities at the Interface of Environmental Science and Policy: A Review of Progress and Future Research Needs. *Environ. Sci. Policy* **2019**, *92*, 141–151. [CrossRef]
29. Sarkki, S.; Heikkinen, H.I.; Komu, T.; Partanen, M.; Vanhanen, K.; Lépy, É. How Boundary Objects Help to Perform Roles of Science Arbiter, Honest Broker, and Issue Advocate. *Sci. Public Policy* **2020**, *47*, 161–171. [CrossRef]
30. Hewitt, R.J.; Macleod, C.J.A. What Do Users Really Need? Participatory Development of Decision Support Tools for Environmental Management Based on Outcomes. *Environments* **2017**, *4*, 88. [CrossRef]
31. De Moor, A. Creativity Meets Rationale: Collaboration Patterns for Social Innovation. In *Creativity and Rationale: Enhancing Human Experience by Design*; Carroll, J.M., Ed.; Human–Computer Interaction Series; Springer: London, UK, 2013; pp. 377–404. [CrossRef]
32. Schuler, D.; de Moor, A.; Bryant, G. New Community Research and Action Networks Addressing Wicked Problems Using Patterns and Pattern Languages. In Proceedings of 7th International Conference on ICT for Sustainability, Bristol, UK, 21–26 June 2020; Association for Computing Machinery: New York, NY, USA, 2020; ICT4S2020; pp. 330–337. [CrossRef]
33. Tu, Y.; Wang, X.; Qiu, R.; Shen, H.W.; Miller, M.; Rao, J.; Gao, S.; Huber, P.R.; Hollander, A.D.; Lange, M. An Interactive Knowledge and Learning Environment in Smart Foodsheds. *IEEE Comput. Graph. Appl.* **2023**, *43*, 36–47. [CrossRef] [PubMed]