frontiers
in Sustainability

Securing Data in Life Sciences—A Plant Food (Edamame) Systems Case Study

Susan E. Duncan[1]*, Bo Zhang[2], Wade Thomason[2], Margaret Ellis[3], Na Meng[3], Michael Stamper[4], Renata Carneiro[1] and Tiffany Drape[5]

[1] Department of Food Science and Technology, Virginia Tech, Blacksburg, VA, United States, [2] School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA, United States, [3] Department of Computer Science, Virginia Tech, Blacksburg, VA, United States, [4] University Libraries, Virginia Tech, Blacksburg, VA, United States, [5] Department of Agricultural Leadership and Community Education, Virginia Tech, Blacksburg, VA, United States

Efforts to identify specialty crop genetics and agronomics, such as for edamame (vegetable soybean), that improve crop yields, resilience, and sustainability often fail to account for data on nutritional content, sensory profile, and/or consumer acceptability. Limited exchange of data across agricultural and food sectors challenges the design of specialty crops that meet consumer needs and expectations and the value chain, and can increase cyberbiosecurity risks. Communication and collaboration within the multi-sector system are essential to address cyberbiosecurity issues related to privacy of data producers, ownership of original data, risks of data sharing, security protection for data transfer and storage, and public perceptions of the food supply chain. This paper introduces a new exploration to design domestic (U.S.) edamame, which is based on both our domain knowledge of life science and our information-sharing mechanisms across the agriculture and food sectors. A case study, involving a multidisciplinary team of breeders and non-breeder researchers with expertise in crop/food production, processing, quality, and economics, serves as a model. We introduce the value chain attributed to combining and linking data from different sectors in the research and development phase, and explain why we believe such data-sharing mechanisms can facilitate better analyses that resonate throughout the full system, from seed to consumer. Our perspective illustrates that, by securely sharing and interpreting data across sectors and identifying cyberbiosecurity risks, we can design promising agricultural and food systems to better meet consumers' need and protection of life science data.

Keywords: edamame, cyberbiosecurity, data, value chain, specialty crop, genetic, consumer, agriculture

INTRODUCTION

For life science data to have the most significant impact and realized value, data digitization, and sharing across sectors inform and create the evidence chain allowing for innovation, trust, and application. In the agriculture and food (A+F) systems, the exchange of data information across sectors is limited and can create opportunities for data security challenges (Duncan et al., 2019; George, 2019). These challenges create vulnerabilities that affect public trust and confidence in the food supply and also in loss of opportunity for innovations. Computerized physical

systems and information systems, with internet-connected devices, are now essential parts of data collection, management, and information exchange. The integration of cyber-physical systems and information technology with life sciences leads to data analytics and technology opportunities that can accelerate and expand our knowledge of biological resilience as well as innovations for creating value (Duncan et al., 2019; Mantle et al., 2019; Reed and Dunaway, 2019).

Emerging from this integration is the need for security of life sciences data and the integrated cyber-physical systems in a new discipline: cyberbiosecurity (Murch and DiEuliis, 2019). Cyberbiosecurity is defined as "understanding the vulnerabilities to unwanted surveillance, intrusions, and malicious and harmful activities which can occur within or at the interfaces of commingled life and medical sciences, cyber, cyber-physical, supply chain and infrastructure systems, and developing and instituting measures to prevent, protect against, mitigate, investigate and attribute such threats as it pertains to security, competitiveness, and resilience" (Murch et al., 2018). This adds the role of responsibility and accountability for data and cyber-physical systems protection to life science researchers, small business owners, and major industries within the specific agricultural sector as well as across the system. This is, in fact, a concern of relevance to national security and protection of our bioeconomy (George, 2019; National Academies of Sciences, Engineering, and Medicine, 2020).

The need for cyberbiosecurity strategies in the A+F systems to secure the domestic and global food supply has already been identified (Duncan et al., 2019). In modern science, large biological datasets and big data analytics have an increasing importance in life sciences, including agriculture technology and economy. For instance, public genoma databases are known to benefit research and innovation in many academic institutions, government entities, and companies, which have important economic value (Vinatzer et al., 2019). However, several countries still lack policies, laws, and deterrents to prevent malevolent cyberbio activities (George, 2019). Despite the fact that the annual investment on generating genomic data corresponds to millions of dollars, there is still a lack of discussions and research focused on cyberbiosecurity risks associated with online biological databases, as well as possible alternatives to mitigate those risks and protect this knowledge (Vinatzer et al., 2019). Nevertheless, it is essential that cyber- and biological areas and perspectives are both considered concurrently to achieve sufficient and effective cyberbiosecurity (George, 2019). However, in specialty crop breeding, which often occurs within academic research universities, there is often a lack of integration of disciplines in the early stage decisions. The breeder is focused on creating the pre-breeding materials that have the best resilience to agronomic conditions and achieving the targeted metrics familiar to the crop. The timeframe from original genetic materials through a licensed final variety may be 10 or more years. Often, there is a gap in data at the latter segments of the value chain, such as the producer, processor, retailer, or consumer, where creating value is important in driving the market (Carneiro et al., 2020). Where do the producer, processor, retail, and consumer needs and values enter the breeder's perspective and how do those needs influence the behaviors of the parties related to the data-driven decisions at the early stages of specialty crop breeding and the related value chain (Fishbein and Ajzen, 2009)? If the stakeholders do not view the work as part of a whole system, but individualized fields, lapses in security can occur due to inconsistent communication, handling of sensitive materials, and general lack of education around protocols and habits that will secure the data.

The theory of reasoned action (TRA) proposed by Fishbein and Ajzen (2010) and the theory of planned behavior (TPB) proposed by Ajzen (1985) guided this study by examining the influence of behavior and decision making related to the edamame value chain. Individual attitude is directly related to the perception of closely affiliated peers (Fishbein and Ajzen, 2009). Knowing this condition already exists, researchers who are part of a larger system can be trained to understand and be aware that their work is not happening within an individual system, but as one part of a large system. Changing behavior is influenced by the need and desire to change. If people in the value chain are not aware of the other actors in a system, they will often not change their intention or their behavior related to securing data. The theories suggest that it is an individual's intention to perform a certain behavior that predicts their actual behavior. These intentions can also be determined by the people's attitudes, cultural norms, and beliefs that influence existing behaviors and attitudes and are also critical in determining the cause of a particular behavioral effect in a community (Fishbein and Ajzen, 2009). Without knowledge and training, it is difficult to change or address someone's intention and behavior.

Communication and collaboration across sectors of the A+F system can provide a clearer understanding of the value of the data to be protected. Both communication and collaboration are also important to identify vulnerabilities, map potential risks (for example, associated with data sharing, transfer, and storage), allocate resources, and build an effective risk management plan. This is especially true for specialty crops, where the large agricultural enterprises and support systems in place for commodities such as corn, wheat, and soybeans do not exist. The risks associated with time and investment of breeding specialty crops are magnified if there is no long-term value and market. Moreover, losses of data or intellectual property is exacerbated because there are no other data sources on which to fill the gaps and rapidly overcome the loss, as is available for the commodity feed and food crops.

Within the U.S. agricultural enterprise, specialty crops are a major economic contributor, estimated in 2017 at nearly $65B and representing 17% of total agricultural sales (USDA, 2019). Specialty crops are differentiated from other federally-supported commodity crops and defined by USDA (Johnson, 2017; AMS, n.d.). Included within the USDA definition of specialty crops are fruits and tree nuts, vegetables, culinary herbs and spices, medicinal herbs, and other horticulture products [further defined], and all nursery and greenhouse crops. In addition, further guidance stipulates that specialty crops must be (1) cultivated or managed; and (2) used for human consumption, medicinal purposes, or for creating aesthetic value/gratification. They may consist of both fresh and processed markets.

Edamame (vegetable soybean) is a specialty crop in the United States. The domestic market for edamame is expanding as U.S. consumers incorporate more protein-rich vegetable foods into their diet. However, most edamame in the U.S. is imported from Asian countries (mainly China) and is not commonly available as a fresh product. The limited domestic production of edamame is attributed to several challenges, such as yield, mechanization, adaptation, as well as identifying and meeting the market needs for the U.S. processing, retail food service and grocery sectors for fresh and frozen edamame products. These rely on the U.S. consumer finding the domestic product that meets their needs and desires for product flavor, nutrition, economics (price) as well as sustainability values. In fact, consumer perspective is not typically addressed in the early stages of the specialty crop breeding process (Carneiro et al., 2020), thus increasing the risk of selecting varieties that meet desired agronomic characteristics but do not address consumer demands.

Within the context of an ongoing multidisciplinary research project on edamame development for domestic (U.S.) production, cyberbiosecurity in an academic setting was explored in a moderated focus group discussion (Section Materials and Methods). In Results and Discussion section, this qualitative approach was used to help us understand the perceptions, existing knowledge, feelings, and ideas of a diverse team of academic researchers about the relevance and value of their research data to improve decision-making for creating domestic edamame varieties. Section Increasing Value With Secure Data describes the concept of increasing economic and security of data through the value chain. The importance of cyberbiosecurity in the design of the value chain (Section Cyberbiosecurity Threats and Strategies for Protection) is described and the need for training in cyberbiosecurity for agriculture, food, and life science contributors (Section Training in Cyberbiosecurity) to protect against threats and secure data across the entire edamame system is explained. The study limitations are acknowledged in section Limitations of the Work. We conclude (Section Conclusions) with the recognition that the importance of the food and agriculture system to the bioeconomy begins with understanding the value chain and the importance of preparing the workforce for contributing to data and economic security. We offer these insights as an illustration for the importance of data communication at the A+F systems level, with the potential value of reinforcing consumer confidence as well as the importance of data security.

MATERIALS AND METHODS

The research protocol for this edamame case study was approved by the Virginia Tech Institutional Review Board (IRB 19-692). One focus group meeting was held (Blacksburg, VA, USA) in November 2019; off-campus participants were able to join the meeting through Zoom (Zoom Video Communications, Inc., San Jose, CA, USA). A total of 11 researchers (8 faculty, 1 postdoctoral researcher, and two doctoral students) participated in the group discussion. Most participants were Virginia Tech employees; one participant was a researcher at the University of Arkansas. Participants providing responses were all part of the Specialty Crop Research Initiative (SCRI) project on domestic edamame variety development (USDA-NIFA, Grant No. 2018-51181-28384; Accession No. 1016465). Five observers who were not part of the SCRI edamame project, including co-authors on this paper, occasionally asked clarifying questions. The moderator was a member of the SCRI edamame project.

The SCRI project participants represented several areas of the A+F systems: plant breeding and genetics, smart farm technology, plant production, crop, soil and environmental science, plant pathology, food processing, flavor chemistry, sensory evaluation, consumer studies, agricultural economics. Observers were researchers in computer science education, software engineering, cyberbiosecurity initiatives, and agricultural education.

The meeting was audio recorded and transcribed for analysis. Transcripts were reviewed and coded by two coders. Audio was transcribed verbatim and open coded to find common themes. After open coding, focused coding took place, and themes emerged (Charmaz, 2008; Krueger and Casey, 2009; Yin, 2011).

RESULTS AND DISCUSSION

Evaluation of the group discussion identified three principal themes: (1) issues with data collection and techniques to protect it; (2) intellectual property issues; and (3) capacity building in human knowledge. From this focus group discussion, it was clear that the recognition of the entire process of the system, from genomic design through marketing and consumer decisions were not foremost in the researcher minds. Nevertheless, the development of a broader perspective of the process and the identification of its potential risks are essential to implement cyberbiosecurity practices (Peccoud et al., 2018). Therefore, we identified the need to clarify and further discuss the overarching edamame (specialty crop) system, illustrate where data influences the value chain, and also describe the rationale behind the design and the threat models with potential relevance for guiding selection of security mechanisms to protect the data and the system. Finally, we identified the need for integrated training for developing and advanced scientists to encourage them to change their intentions and behaviors to encourage adoption of rigorous practices to work toward more consistent handling practices of sensitive data. Such training is necessary to assist researchers, supply and equipment manufacturers, growers, food processors, distributors and retailers with the adoption of and implementation of cyberbiosecurity strategies for protection of specialty crops (edamame) under development and production in the country.

Increasing Value With Secure Data

The case study focus group discussion provided information used in conceptualizing the influence of the research activities on the data flow and economics involved in building the domestic edamame market. **Figure 1** presents a general framework for building the domestic edamame market. Within this framework, the relationship of the value chain is illustrated in four phases.
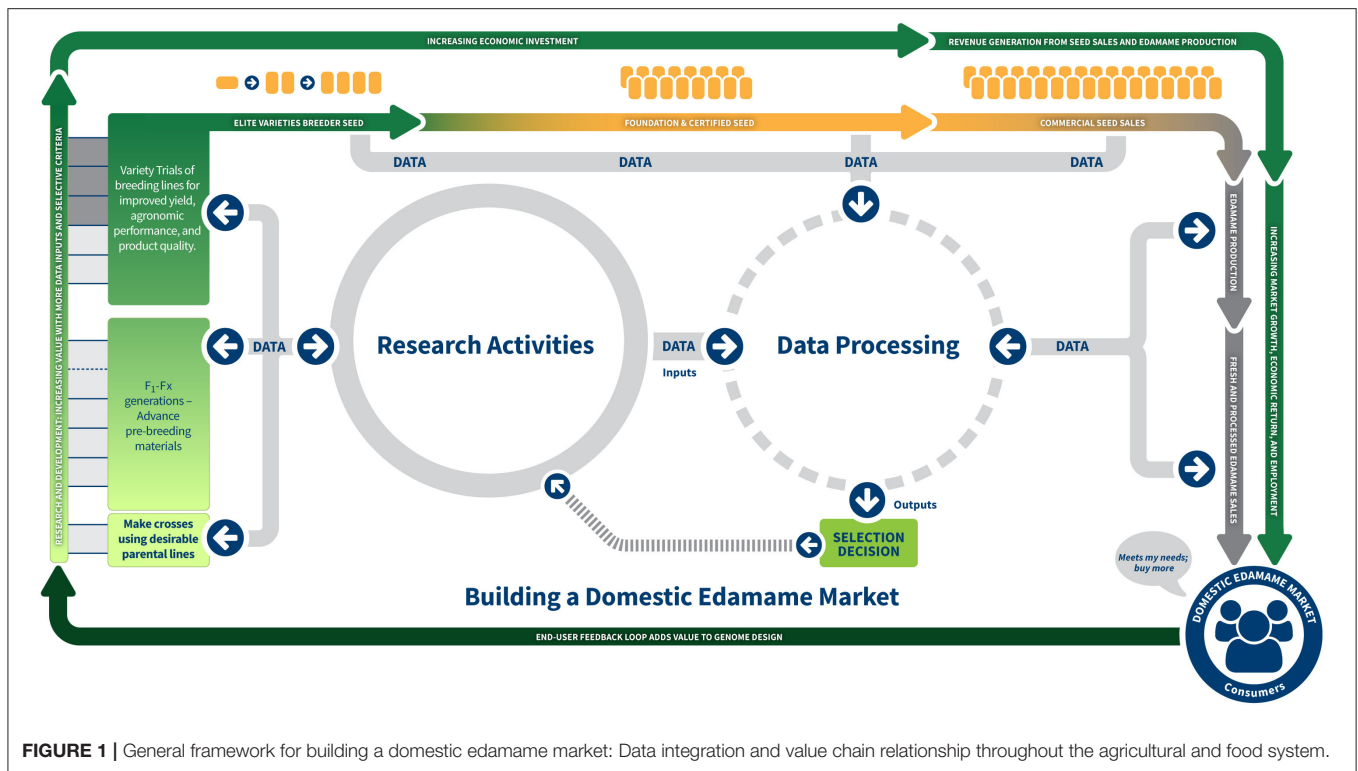
FIGURE 1 | General framework for building a domestic edamame market: Data integration and value chain relationship throughout the agricultural and food system.

Phase I (left side of **Figure 1**) illustrates the increase in value as research and development investments are high with the goal of selecting elite varieties that have targeted yield, agronomic performance, and product quality.

Phase II (top of **Figure 1**) illustrates expanding elite seed availability (gold "seed" images) and initiating seed sales, which also requires economic investment. Through this phase, seed is the primary product and, as seed availability is expanded, commercial seed sales contribute to revenue generation. Approved varieties may be commercially marketed and breeders provide breeder seed to produce foundation and certified seed. These steps are essential for expanding the seed base and validating that seed purity and quality are retained. When sufficient certified seed is available, the specialty crop planting for the purpose of marketing begins. This may take another 1–3 years post-approval of the variety before sufficient seed is available for marketing for crop seed (top of **Figure 1**). At this point, the investment of the intellectual property through the breeding and elite varietal selection process and the establishment of the certified seed for crop production has all been built in anticipation of the potential success of the future market. Data and seeds are the primary products.

Phase III (right side of **Figure 1**) characterizes the economic growth as suggested by market growth, economic return, and employment expansion as edamame production, processing, and fresh and processed edamame sales are increased. Note that consumers are instrumental in the success of this market as they identify edamame as an option in their diet, consider the value to their nutritional and dietary needs and sensory

enjoyment, and repeat purchase, thus increasing demand. The cost of intellectual properties and seed generation is recovered as more seed is sold and more edamame is harvested, processed, sold directly or through retail markets, and consumers respond favorably. Understanding the "end-user" pivot point is essential for providing feedback (Phase IV; bottom of **Figure 1**) as it influences the research and development of improved edamame varieties that address yield, performance, and improved product quality. The feedback brings a better understanding of the need and value for improvements and value-added quality, thus creating a cyclic value chain.

Building the edamame market requires the exchange of information, identified as "data" in this image. Data inputs are not one directional, but influence decisions throughout the framework. **Figures 2A,B** further illustrate the importance and relevance of this data exchange in building the edamame market. **Figure 2A**, which represents the detail of the "Research Activities" circle in **Figure 1**, describes the research disciplinary domains, knowledge contributed, activities and data processing relationships to decisions for selecting desirable parental lines, advancing the pre-breeding materials, and then local and regional/national variety trials for selecting the elite varieties. The center portion of **Figure 2A** illustrates that the transdisciplinary exchange of information helps the breeder make informed and improved decisions at the earliest stages of the breeding process (selecting parental lines and making crosses).

In the early stages of new specialty crop (edamame) development, breeding crosses are made to improve genetics for enhancing yield and composition, pest and disease resilience,

FIGURE 2 | (A) Transdisciplinary research and data integration for the research and development process of edamame breeding; **(B)** Data management and processing influences decisions through the value chain.

and plant architecture, which require the engagement of breeders and several other plant and agronomic scientists. Compositional chemistry, such as quantification of total sugars for evidence of potential sweetness, contribute analytical measures for selection that are supported by the food science team. When advanced breeding lines are selected for field trials, portable computer tablets are used for data entry and smart farm technologies and sensors are used for in-field assessments of plant architecture, canopy closure and health as well as environmental conditions such as weather and soil types. Then, these advanced breeding lines are further selected based on fresh yield, agronomic traits, pest resistance, etc. Quality and sensory characterization (e.g., appearance, flavor) and consumer acceptability for each variety enhances the selection process (Carneiro et al., accepted) and increases the confidence that a variety has market potential. Thus, for this research team, elite varietal selections are those that achieve highest integrated agronomic, quality, and consumer responses across multiple environments.

This process of defining, developing, and selecting the best elite varieties takes at least 10 years, from crossing the initial parental lines to the final elite variety selection. As defined in this case study, elite varieties are submitted for review by the university Plant Germplasm Committee and subsequent approval. This approval serves to provide protection of the intellectual property and allow for royalties to return to the program when elite seed is purchased. A return on this investment requires a strong edamame market and protection of the intellectual property associated with the breeding and genetics of the varieties. The general framework described for building a domestic edamame market (**Figure 1**) might serve as a template for other plant (food and feed) systems.

Figure 2B further illustrates the data management and processing and influence of data on the decisions in each phase of the value chain. Throughout this value chain, multimodal data creates challenges in integration and interpretation. The protection of the genetic improvement of the specialty crop is imperative as loss or corruption of this information stops the value chain cycle. To secure the exchange and storage of data, implementing the SSL/TLS (Josh Fruhlinger, 2018) protocol encrypts the communication channel, and encrypts the file system to secure files in local machines and the cloud. In addition, implementing authentication and authorization mechanisms ensures that (1) only authenticated users can read data, and (2) only users with appropriate privileges can modify data.

The domestic (U.S.) edamame market chain is in its infancy. Most (70%) edamame is imported from China and other Asian countries, though there is an increased production of edamame as producers are seeking alternative crops (Carneiro et al., 2020; Carneiro et al., accepted). In this case study, access to data from seed expansion, commercial seed sales, production, processing, and final product sales was not available as there is limited domestic edamame production and processing. **Figures 1**, **2B** illustrate where these data streams can be useful in influencing the market. As the domestic market expands, integrating the production, processing, and market data with the research and development data further guides the selection process. Building

the domestic edamame market requires that the data that guides the decisions must be available, fact-bound, protected, and shared.

Cyberbiosecurity Threats and Strategies for Protection

According to focus group participants, the type of generated data in the edamame research activities and data processing (**Figures 2A,B**) included genotypic and agronomic data, drone images and digital photos, sugar and amino acid profiles, gas chromatography mass spectral data, sensory (consumer) and economic data. Several team members reported the use of Microsoft Excel spreadsheets for data entry. Regarding data protection alternatives during the data collection step, one participant recommended locking Excel templates for data creation in order to prevent users from modifications or wrong manipulation. More specific software, targeted to the function or discipline, were also reported to be used for data acquisition and/or storage: for example, Compusense® Cloud (Compusense Inc., Guelph, Ontario, Canada) for consumer studies, ChemStation (Agilent Technologies, Inc., Santa Clara, CA, USA) for food analysis, and ARM—Agricultural Research Manager (GDM Solutions, Inc., Brookings, SD, USA) for management of agriculture research experiments (plant breeding, crop production). Researchers also used Labellmg for plant architecture analyses. The need to login with a user identification, password, and/or multiple-factor authentication to access the software was only mentioned for Compusense® Cloud. Some researchers also reported the use of physical laboratory notebooks and hard drives to document and store data. Google Drive (cloud storage) was identified as the main tool used for data storage and backup plan; all team members had access to a shared folder, which was managed by the project principal investigator. Shared folder access was linked to the university email account of each individual. Furthermore, at Virginia Tech, individuals are requested to sign into their Google accounts using a strong password and two-factor authentication.

Awareness of cyberbiosecurity weaknesses and challenges in life sciences, such as pathogen genome databases, is growing (Murch and DiEuliis, 2019; Vinatzer et al., 2019). As online genoma databases can be used in breeding programs to facilitate development of improved food crops such as edamame, several of the cyberbiosecurity weakness and challenges presented in the literature also apply to the SCRI edamame project explored in our case study. For instance, confidentiality, data integrity, data availability, attack on physical hardware, and physical exploitation were identified as the major security threats to genoma databases (Vinatzer et al., 2019). Several of these threats were also identified in our focus group meeting. One participant, for example, expressed concerns regarding data corruption and loss of sensitive information. Overall, participants discussed the uniqueness of their data, which was mostly attributed to the fact their edamame genotypes were developed and selected to be adapted to local growing conditions. One participant emphasized the fact that the group collected economic data in particular years. In this interdisciplinary project, consumer acceptability

and economic impact were both considered, which is not typical (Carneiro et al., 2020). Participants also raised concerns about intellectual property protection and shared the fear of having their data stolen, for example, by companies that would be interested in commercializing edamame seeds. In fact, a year of lost data could be devastating in a process that takes a decade or more of work. The genomic data of all breeding lines should be protected. Someone who can reverse engineer the genome from the late (X+5 years or elite) seeds can capture multiple years-worth of intellectual property development and quickly enter the market with the same varieties, which would also affect our bioeconomy.

Suggestions for mitigating such risks include improvements in data integrity and protection, data availability and longevity, and access control (Vinatzer et al., 2019). The use of cryptographic checksums, secure multiparty computation, or encrypted databases was suggested by Vinatzer et al. (2019) for data integrity checks and protections, for example, to prevent potentially malicious users from modifying data records or injecting useless data to public databases. Regarding data availability and longevity, the authors encouraged researchers to deposit their data in well-established databases maintained by national governments as the lack of funding to securely maintain online databases may result in data loss. Then, as some databases do not require user authentication or a strong password, implementing two/multi-factor authentication or login from third party accounts in addition to requesting a strong password were suggested as possible alternatives to increase security and improve access control (Vinatzer et al., 2019). We believe these are necessary steps to protect research in life science.

Training in Cyberbiosecurity

The main theme that emerged as a result of the edamame case study was "the need to train all parties" who would engage in this kind of work, which may include senior and young researchers (breeders and non-breeders), growers, food processors, supply and equipment manufacturers, and food distributors and retailers. From field trials with open space to labs and other settings, the largest identified gap was in training and education on how to secure sensitive and proprietary data during all phases of edamame growth and development for consumption. Training all parties listed above would lead to knowledge gain and intention to change behavior to be more secure and rigorous when working with proprietary data. Likewise, Peccoud et al. (2018) acknowledged employee training as the first step to implement a culture of cyberbiosecurity awareness and increase security and resilience in life sciences organizations (e.g., biotech companies or academic institutions). According to the authors, the next steps would be the identification of risks and vulnerabilities though a systematic analysis of the workflow, and the development of new policies to prevent and detect possible incidents. The workflow and value chain of the edamame research was presented in the previous sections of this work, as well as its potential vulnerabilities. Nevertheless, building a cyberbiosecurity culture is still needed.

Moreover, Vinatzer et al. (2019) suggested in their work that the little attention that is currently dedicated to cyberbiosecurity

risks to online biological databases may be associated with the fact cyberattacks to such databases would directly affect fewer people (mostly researchers) than attacks to large enterprises. Applying this perception to the SCRI edamame project discussed in our focus group meeting, we suggest that the large edamame dataset, generated by several departments of multiple public academic institutions, is possibly more vulnerable to cyberattacks than crop data generated by large seed companies, for example. During the group meeting, participants discussed how their large dataset could be attractive to several agribusinesses. Another vulnerability of the academic environment is the fact that the project principal investigator often is the project manager. In most cases, academic researchers in life sciences do not have formal training in project and risk management, which would be beneficial for identification and mitigation of potential cyberbiosecurity threats, especially in large and extended duration research projects. In contrast, larger corporations commonly invest in management training and sometimes even have dedicated project manager(s) in their leadership teams. Often these corporations also comply with international standards and pursue certifications that request a broad risk management plan, such as the ISO 9001: 2015 certification, which defines the international requirements for a quality management system. However, there is a consensus among public and private sectors regarding the current need to increase cyberbiosecurity training (George, 2019). While the cybersecurity workforce has numerous training programs and resources (Newhouse et al., 2017), the life science workforce needs resources to build cyberbiosecurity capacity (knowledge, skills, and abilities) and a common, consistent lexicon to describe and address the unique challenges within A+F. Specifically, A+F workforce needs to be able to understand sufficiently the software, hardware, data exchange system as well as the life science system in order to provide a basic risk assessment and identify potential threats. There is also a need for a cybersecurity workforce that effectively communicates with small and private agribusinesses to further advance cyberbiosecurity. Based on the edamame group discussion, it was evident that increased awareness and training about data protection and cyberbiosecurity would encourage consistent behaviors related to securing data and create benefit in the process of developing new specialty crop varieties. Future training efforts may include online or in person workshops, seminars, and roundtable discussions, which could be organized by either public or private institutions, or both working in collaboration. Over the long term, the development of a cyberbiosecurity curriculum for formal undergraduate and graduate programs may also be considered (Richardson et al., 2019). By extension, such efforts would create a workforce more readily able to secure data, prepared for protecting the economic value of the food and agricultural system, through undergraduate and graduate student training in cyberbiosecurity.

Limitations of the Work

A known challenge of focus groups is that facilitator and group members can bias responses (Jones et al., 2019). In our work, we acknowledge that the moderator was a member of the SCRI edamame research team as were most participants, which may

be a potential source of bias. A single focus group meeting was performed with a large group (16 people total, as the 5 observers were also engaged in the discussion), participants were not randomly chosen, and a few SCRI team members ($n = 3$) contributed only a few comments. Also, this work represents only the perspective of one academic research group. Although the outcomes provided useful insights, they are not generalizable and do not represent a larger population (e.g., researchers in life sciences).

CONCLUSIONS

The importance of data in the research and development and throughout the entire value chain of specialty crops, such as edamame, is illustrated in this case study. Secure data increases value of A+F and life sciences research and also contributes to protecting the bioeconomy (Duncan et al., 2019). Nevertheless, research conducted in academic settings may be more vulnerable to cyberattack than research conducted by large corporations, thus education on cyberbiosecurity is needed to change this scenario. The edamame case study demonstrated that critical analysis of the workflow and value chain is important to identify possible cyberbiosecurity threats and create strategies for protection. Further work is recommended to have a bigger picture of how representants of other fields in life sciences (e.g., animal sciences, biochemistry, human nutrition) perceive cyberbiosecurity. This bigger picture is important for the development of training materials, which may range from small seminars to the development of new undergraduate and graduate programs.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Virginia Tech Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Ajzen, I. (1985). "From intentions to actions: a theory of planned behavior," in *Action Control.* Berlin, Heidelberg: Springer, 11–39.

AMS (n.d.). *USDA Definition of Specialty Crop.* Available online at: https://www.ams.usda.gov/sites/default/files/media/USDASpecialtyCropDefinition.pdf (Retrieved July 24, 2020).

Carneiro, R. C. V., Duncan, S. E., O'Keefe, S. F., Yin, Y., Neill, C. L., and Zhang, B. (2020). Sensory and consumer studies in plant breeding: a guidance for edamame development in the U.S. *Front. Sustain. Food Syst.* 4:124. doi: 10.3389/fsufs.2020.00124

Charmaz, K. (2008). "Reconstructing grounded theory," in Ch. 27. *The SAGE Handbook of Social Research Methods*, eds P. Alasuutari, L. Bickman, and J. Brannen (London: Guilford Publishers), 397–412. doi: 10.4135/9781446212165

Duncan, S. E., Reinhard, R., Williams, R. C., Ramsey, F., Thomason, W., Lee, K., et al. (2019). Cyberbiosecurity: a new perspective on protecting U.S. food and agricultural system. *Front. Bioeng. Biotechnol.* 7, 1–7. doi: 10.3389/fbioe.2019.00063

Fishbein, M., and Ajzen, I. (2009). *Predicting and Changing Behavior: The Reasoned Action* Approach. New York, NY: Psychology press.

Fishbein, M., and Ajzen, I. (2010). *Predicting and Changing Behavior: The Reasoned Action Approach.* New York, NY: Psychology Press.

George, A. M. (2019). The national security implications of cyberbiosecurity. *Front. Bioeng. Biotechnol.* 7, 1–4. doi: 10.3389/fbioe.2019.00051

Johnson, R. (2017). Defining "specialty crops": a fact sheet. *Congr. Res. Serv.* R44719. Available online at: https://crsreports.congress.gov/product/pdf/R/R44719/5

Jones, K. R., Gwynn, E. P., and Teeter, A. M. (2019). Quantitative or qualitative: selecting the right methodological approach for credible evidence. *J. Hum. Sci. Ext.* 7, 61–87. Available online at: https://www.jhseonline.com/issue/view/88

Josh Fruhlinger, J. (2018). What is SSL, TLS? And how this encryption protocol works? Retrieved August 22, 2020 from https://www.csoonline.com/article/3246212/what-is-ssl-tls-and-how-this-encryption-protocol-works.html Dec. 4, 2018.

Krueger, R. A., and Casey, M. A. (2009). *Focus Groups: A Practical Guide for Applied Research.* Fourth ed. Los Angeles: SAGE Publications, Inc

Mantle, J. L., Rammohan, J., Romantseva, E. F., Welch, J. T., Kauffman, L. R., McCarthy, J., et al. (2019). Cyberbiosecurity for biopharmaceutical products. *Front. Bioeng. Biotechnol.* 7:116. doi: 10.3389/fbioe.2019.00116

Murch, R., and DiEuliis, E., editors. (2019). *Mapping the Cyberbiosecurity Enterprise.* Lausanne: Frontiers media. doi: 10.3389/978-2-88963-213-8

Murch, R. S., So, W. K., Bucholz, W. G., Raman, S., and Peccoud, J. (2018). Cyberbiosecurity: an emerging new discipline to help safeguard the bioeconomy. *Front. Bioeng. Biotechnol.* 6:39. doi: 10.3389/fbio.2018.00039

National Academies of Sciences, Engineering, and Medicine (2020). *Safeguarding the Bioeconomy*. Washington, DC: The National Academies Press. doi: 10.17226/25525

Newhouse, W., Keith, S., Scribner, B., and Witte, G. (2017). *National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework*. National Institute of Standards and Technology Special Publication 800-1818, 144. doi: 10.6028/NIST.SP.800-181

Peccoud, J., Gallegos, J. E., Murch, R., Buchholz, W. G., and Raman, S. (2018). Cyberbiosecurity: from naive trust to risk awareness. *Trends Biotechnol*. 36, 4–7. doi: 10.1016/j.tibtech.2017.10.012

Reed, J. C., and Dunaway, N. (2019). Cyberbiosecurity implications for the laboratory of the future. *Front. Bioeng. Biotechnol*. 7:182. doi: 10.3389/fbioe.2019.00182

Richardson, L. C., Lewis, S. M., and Burnette, R. N. (2019). Building capacity for cyberbiosecurity training. *Front. Bioeng. Biotechnol*. 7:112. doi: 10.3389/fbioe.2019.00112

USDA (2019). *2017 Census of Agriculture*. United States Summary and State Data. Vol. 1. Geographic Area Series. Part 51. AC-17-A-51.

Vinatzer, B. A., Heath, L. S., Almohri, H. M. J., Stulberg, M. J., Lowe, C., and Li, S. (2019). Cyberbiosecurity challenges of pathogen genome databases. *Front. Bioeng. Biotechnol*. 7, 1–11. doi: 10.3389/fbioe.2019.00106

Yin, R. K. (2011). *Applications of Case Study Research*. 3rd ed. Washington, DC: Sage Publishing, Inc. 264pp.