

Minimizing Threats Through Responsible Voluntary Confidential Food Safety Data Sharing: Insights from Food Industry Leaders

Linda Kalunga

lk549@cornell.edu

Cornell University <https://orcid.org/0009-0008-4259-6953>

Katherine Koebel

Cornell University

Carrie Alexander

University of California - Davis

Martin Wiedmann

Cornell University <https://orcid.org/0000-0002-4168-5662>

Aaron Smith

University of California - Berkeley

Aaron Adalja

Cornell University

Renata Ivanek

Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University <https://orcid.org/0000-0001-6348-4709>

Article

Keywords: food sector, privacy, trust, data governance, decision making

Posted Date: July 10th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-6960610/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Minimizing Threats Through Responsible Voluntary Confidential Food Safety Data Sharing: Insights from Food Industry Leaders

Linda Kalunga^{1*}, Katherine Koebel¹, Carrie Alexander², Martin Wiedmann³, Aaron Smith⁴, Aaron Adalja⁵,
Renata Ivanek¹

¹Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, New York, U.S.A.

²Department of Science and Technology Studies and AI Institute for Next Generation Food Systems (AIFS), University of California – Davis, Davis, California, U.S.A.

³Department of Food Science, Cornell University, Ithaca, New York, U.S.A.

⁴Department of Agriculture and Resource Economics, University of California – Berkeley, Berkeley, California, U.S.A.

⁵Cornell SC Johnson College of Business, Cornell University, Ithaca, New York, U.S.A.

*Corresponding author: lk549@cornell.edu

Abstract

Voluntary sharing of confidential food safety data among companies can improve food safety and enhance efficiency and decision-making, especially for rare events. However, despite its potential benefits, inter-organizational data sharing in the food industry has lagged, with limited research exploring the reasons for this hesitation. To address this knowledge gap, we conducted in-depth semi-structured interviews with 27 food industry leaders. We found that companies face trade-offs when deciding to share data, weighing risks against benefits. The level of trust that the data sharer has in the stakeholders (e.g., industry peers, regulatory bodies, customers) and technology strongly influences voluntary data sharing decisions. Data governance emerged as a solution to concerns stemming from trust, such as loss of control once data is shared. Our findings provide timely evidence to inform and facilitate data-sharing initiatives and practices, enabling collective analysis of localized incident data to reduce future food safety incidents.

Key words: food sector, privacy, trust, data governance, decision making

1. Introduction

Despite significant advancements in food safety systems, foodborne illnesses continue to pose a substantial public health challenge in the United States (U.S.), with an estimated 48 million illnesses, resulting in 128,000 hospitalizations and 3,000 deaths annually.¹ Numerous foodborne outbreak investigations have traced contamination to the production and processing stages of the supply chain. For example, a 2024 multistate *Listeria monocytogenes* outbreak linked to ready-to-eat meat products was linked to 61 illnesses, 60 hospitalizations, and 10 deaths, traced to a single processing plant.² Similarly, a 2024 *Salmonella* Braenderup outbreak with 551 reported illnesses and 155 reported hospitalizations was linked to cucumber contamination during production.³ While such incidents are very rare, their continued occurrence across the industry points to systemic vulnerabilities that could be addressed through inter-organizational data sharing.

Data sharing represents a potentially transformative solution for food safety practices within the industry, as companies would gain access to safety data not just from their own operations but from the industry as a whole. This expanded access would enable companies to identify patterns of otherwise rare food safety failures, support advanced predictive modeling of potential risks, drive adoption of best practices, and foster continuous improvement throughout the industry. Ultimately, such collaborative efforts could improve production efficiency and reduce foodborne illness outbreaks traced to processing facilities. Industry initiatives such as Western Growers GreenLink project demonstrate how aggregation of member company food safety data can be used to identify common risk factors and collaboratively address industry-wide food safety challenges.⁴ However, despite the recognized potential, comprehensive voluntary food safety data sharing within the food industry remains limited. This implementation gap suggests

the existence of barriers, such as concerns about confidentiality⁵ and lack of standardized data formats⁶ that have been identified in the food industry but remain underexplored in the context of voluntary food safety data sharing. To address these knowledge gaps, we used semi-structured interviews with industry leaders, a method suited for exploring the nuanced experiences, concerns, and perceptions while maintaining consistency across participants.⁷ Specifically, our objectives were to investigate (i) the perceived benefits motivating companies to engage in voluntary, confidential food safety data sharing within the food industry, ii) the barriers that impede such sharing, and (iii) the conditions and strategies that can enable broader adoption of such practices. By understanding these factors, our research offers insights into how voluntary food safety data sharing can be effectively implemented.

2. Methods

2.1. Sampling and recruitment

To achieve the objectives of our study, we used a qualitative research design in the form of semi-structured interviews to allow for flexibility in exploring emerging themes in depth. We recruited participants through purposive sampling, leveraging the professional networks of MW and RI, with additional participants identified through snowball sampling. Participant eligibility included individuals in leadership positions (e.g., president, director or manager) within the U.S. food industry or those representing companies supporting the food industry in a food safety capacity (e.g., laboratory services). Additionally, we only considered individuals who had at least 2 years' experience working in the food industry, fluency in English, and individuals over the age of 18. These criteria ensured participants possessed the necessary expertise and organizational perspective to provide meaningful insights. We emailed an informed consent document, which outlined study procedures, confidentiality measures, and potential risks and benefits to those who

expressed interest in the study. The study was approved by the Cornell University Institutional Review Board for Human Participants (IRB protocol #IRB0147486). To protect participant confidentiality, identifying information was removed from transcripts and quotes, including participant ID numbers. As a result, quotes are presented without attribution to specific participants.

2.2. Data Collection

We designed an interview guide that focused on two key areas: (i) how food safety data is collected and shared within the industry and (ii) instances where food safety data is collected but not shared, along with the reasons behind these decisions. The guide was piloted with three food safety professionals (two representing food industry companies and one with relevant experience in the industry and government sector) and conducted by LK and RI to ensure clarity and relevance before full implementation. Interviews were conducted between June 2023 to January 2024 via teleconference, allowing for flexibility while maintaining consistency in data collection. Each interview began with obtainment of consent for audiovisual recording in addition to an initial inquiry into the participants' employer, role, and tenure contextualization. All interviews were recorded and later transcribed for analysis. The interview guide, which outlines the full set of questions and prompts used in the study, is available in Supplemental Information S1.

2.3. Data Analysis

Interview transcripts were transcribed using an automated transcription service provided by NVIVO Transcription (Lumivero, Denver, CO). Following this, each transcript underwent a manual validation process by LK to ensure that transcripts matched the recordings. This involved listening to the audio recordings while reading the transcripts, allowing for the correction of any discrepancies between the spoken word and the transcribed text.

We used a reflexive thematic analysis approach, which included six phases: i) familiarization with data, ii) generating initial codes, iii) generating themes, iv) reviewing potential themes, v) defining and naming themes, and vi) producing the report.⁸ However, these phases did not occur linearly due to the iterative nature of the analysis. During each phase, LK, CA, KK, and RI, individually familiarized themselves with the data, and collaboratively reflected on and discussed emerging codes and themes, revisiting the data and themes throughout all six phases of thematic analysis. During phase ii, we used a combination of inductive and semantic coding approaches. Inductive coding allowed codes to emerge directly from the data without relying on a pre-existing framework, while semantic coding focused on the explicit content of participants' responses rather than interpreting underlying meanings.⁸

We assessed saturation using a binary-coded matrix of our codes from NVIVO. We examined code emergence by (i) plotting a sequential saturation curve, (ii) comparing the actual interview sequence with a randomly ordered set, and (iii) testing the stability of saturation points by calculating the cumulative percentage of unique codes identified across interviews by running 10,000 bootstrap iterations.⁹ We used the entire dataset as the base size, with the total number of unique codes across all interviews acting as the denominator in saturation calculations. This approach assumes the full dataset represents the universe of possible codes and each interview contributes equally to code discovery. We defined a run length of one interview, meaning each interview was assessed individually for new codes. Interviews were analyzed sequentially in the order conducted, comparing codes in each interview to a cumulative set of previously discovered codes. New codes were added to the cumulative set, and we calculated the cumulative percentage of codes discovered after each interview (S_k) using Equation 1⁹:

$$S_k = \frac{C_k}{C_N} \quad \text{Eq. (1)}$$

where C_k is the cumulative number of unique codes identified up to the k^{th} interview, and C_N is the total number of unique codes in the dataset.⁹ We defined saturation using a 0% new information threshold, consistent with existing studies^{10,11}, meaning saturation occurred when subsequent interviews introduced no new codes (i.e., $S_k = 100\%$). To assess the potential impact of interview order on the saturation points, we conducted a bootstrap analysis by generating 10,000 samples through random sampling with replacement. In each iteration, we recorded the number of interviews required to reach 80% and 90% saturation of all code coverage.^{9,12} We derived median saturation points and 5th and 95th percentiles to characterize the variability in the number of interviews needed to reach each threshold. All analyses and visualizations were conducted using Jupyter notebook via Google Colaboratory (Colab). The data and code is available on the GitHub repository:

https://github.com/IvanekLab/Saturation_Analysis_Interview_Study.

One of the preliminary themes that emerged was on barriers to data sharing. We characterized the identified barriers into three groups, expanding on the categories defined by Panhuis et al. (2014): technical (issues relating to data formats, data interoperability, and standardization), economic (possible economic impact as a result of data sharing), and social (relating to the perceptions, attitudes, and trust of stakeholders towards data sharing).¹³ While Panhuis et al. (2014) originally characterized barriers in the context of public health, we are repurposing them to analyze the challenges related to food safety data sharing in the food industry.

When talking about barriers to data sharing, several participants commented on the need for data governance. We mapped the barriers raised by participants against the data governance domains¹⁴ that would address them, adapting a previously introduced framework.¹⁵

3. Results

We conducted 27 interviews out of the 87 invited food industry leaders, yielding a response rate of 31%. The interview lengths ranged from 19 to 52 minutes. The characteristics of the interview participants are presented in Table 1, which provides an overview of represented food industry sectors, years of experience, and participant positions.

Table 1: Summary of participant characteristics

Characteristics		n
Food Industry Represented	Dairy	6
	Food Manufacturing	8
	Meat ^a	4
	Produce	7
	Laboratory Services	3
Years of Experience		
	2-5	8
	6-10	13
	11-15	4
	≥ 16	2
Participant Role/Position		
	C-suite and/or President	3
	VP of Food Safety & Quality	12
	Director of Food Safety	4
	Manager (plant, food safety, quality)	8
Total Interviews ^b		27

^a Includes beef and poultry

A total of 65 codes were identified across 27 interviews. In the code saturation analysis, no new codes were identified after interview #12 (Figure S1). The comparison of code emergence in the actual vs randomized interview order suggested that the saturation point observed in Figure S1 is relatively robust, though modestly influenced by interview order (Figure S2). The bootstrap analysis showed that a median of 8 interviews reached 80% code saturation

(with the 5th and 95th percentiles of 5 and 12 interviews), and a median of 13 interviews reached 90% saturation (5th and 95th percentiles: 8 and 23) (Figure 1). The binary-coded matrix of all codes and interviews is provided in Supplemental Information S3.

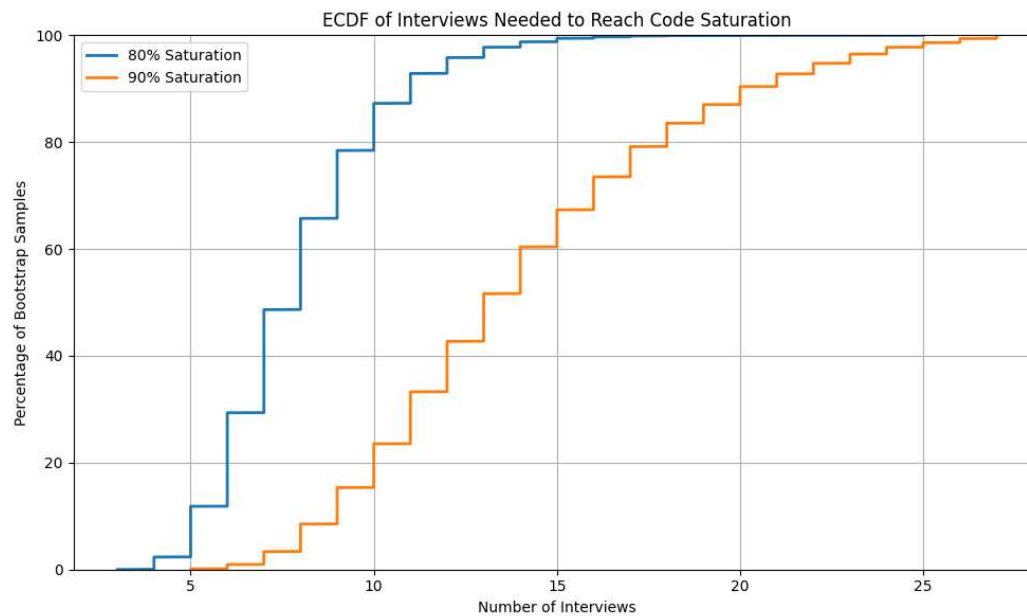


Figure 1: Cumulative percentage of simulations reaching 80% and 90% code saturation by number of interviews. Each curve is based on 10,000 bootstrap iterations, where interviews were randomly sampled with replacement. The x-axis shows the number of interviews, and the y-axis shows the percentage of iterations in which 80% and 90% of all codes were identified by that count. For example, the 90% curve reaches 100% at the 25th interview, meaning that in every simulation, 90% of the total codes had been identified by that point.

Table 2 presents a summary of the key themes we identified through thematic analysis of interview transcripts, highlighting benefits, barriers, and solution aspects of voluntary confidential food safety data sharing. In the following sections, we explore these themes in greater detail, illustrating them with direct quotes from participants. Those additional quotes, organized by theme and theme aspect, are also provided in Supplemental Information S4.

Table 2: Overview of themes identified in semi-structured interviews about voluntary inter-organizational food safety data sharing with 27 food industry leaders

Theme	Category
1) Data sharing can enhance food safety practices and decision-making due to access to more data points	Benefits
2) Technical barriers, such as limited digitalization, standardization of data collection and management practices, limit the feasibility of data sharing	Barriers
3) Trust, or lack thereof, is a key factor influencing data sharing decisions, often underlying specific concerns such as confidentiality, liability risks, and data misuse	Barriers
4) The need exists for sustained collaboration between academia, government, and industry to identify trade-offs in data sharing, establish data governance, and enable improved food safety through information pooling.	Solutions

Theme 1: Data sharing can enhance food safety practices and decision-making due to access to more data points

Participants recognized the potential benefits of food safety data sharing within the food industry to enhance food safety. One of the main recognized benefits was access to more data. Specifically, some participants expressed that the current industry challenge is understanding what causes foodborne outbreaks. This challenge exists because the food safety failures at individual companies are so infrequent and highly variable, making it difficult to identify patterns or determine the root cause of these failures. One participant mentioned that:

"The reason we feel that this data sharing can help in those regards is because the failures as it relates to the data that we're talking about here are so infrequent an individual company might not be big enough to have enough data points to build and understand trends. And so, what the industry talks about is if we could share our data collectively in a manner that would give us more data points to look at, then we might be able to build a trend or understand what's going on from an industry perspective because all of those data points then can give you more things to look at versus in our particular case."

Similarly, participants empathized that the food safety challenges they face are not unique to individual companies, but that industry competitors often encounter similar issues. As one participant explained:

"And since some of the struggles we're facing, we're probably seeing the same thing with our competitors ... To share those data, to understand the struggles that everybody is facing can help to improve and to drive best practice in the same sectors."

In addition to the driving best practices within the industry, interview participants also highlighted the value of data sharing in enhancing predictive modeling and enabling earlier risk identification, prevention, and mitigation due to access to a collective industry food safety data.

"I think the power of interrogating data really comes from as rich and diverse a dataset as possible, and bringing together different parts of the industry and different perspectives, as broad a supply network as possible because the more we do that, the earlier we are likely to notice any early signals and also to be able to move towards prediction and what those key triggers might be and, how do we use big data and modeling to actually create predictive models that enable us to get ahead of some of these events or notice them very much earlier than typically we do now. And I think there's also an opportunity, if we can notice things earlier, we can, as an industry act more quickly."

Some participants viewed data sharing as a way to level the playing field between large and small companies within the food industry. They explained that small companies could benefit from the data and experience of the wider industry, allowing them to improve their food safety practices without having to invest heavily in research to determine their best practices. As one participant explained:

"I think it will be very beneficial, especially for small companies, because the big companies, they can spend all the money they want in researching or finding whether it works, but the small companies don't have those resources and sometimes they struggle to find what is best because they don't have that experience."

One participant representing the produce industry highlighted that food safety risks extend beyond the individual companies and have industry-wide consequences. They noted that in the event of an outbreak, consumers tend to avoid the entire product commodity rather than

distinguishing between brands. Given these conditions, participants highlighted the necessity of data sharing to help their specific industry collectively and continuously improve food safety practices.

"What's unique about the produce industry compared to other industries is that we are much less focused on produce brands ... When an outbreak in that commodity happens, people avoid all of that product. They don't just avoid tomatoes from one company or avocados from one company. They just avoid all tomatoes and in that sense, there is very much the idea that when a food safety incident occurs, it's in the entire industry's best interest to understand what happened and to be able to kind of collectively raise that bar of food safety. That's really the biggest advantage and benefit to sharing data or sharing outcomes is to be able to, like I said, just drive continuous improvement across an entire commodity."

Theme 2: Technical barriers limit the feasibility of food safety data sharing within the industry

Participants identified the technical barriers that hinder the adoption of food safety data sharing initiatives, including the varying levels of digitalization across the industry and the lack of standardization regarding data documentation across the industry. Participants shared differing perspectives on the state of digitalization within the food industry. While some acknowledged ongoing efforts to transition from paper-based to digitalized systems, there was broad agreement that the industry remains on a spectrum, with some companies still relying heavily on paper.

"I guess I'll speak again to my company, we are further behind than I would prefer. We have active projects working to digitalize things. However, we still rely a lot on manual and paper documentation or systems where there's several layers where you have a piece of paper, but it's in a system and then perhaps there's a software that can help actually do something with the data. I would say we're still I would call us, you know, evolving in that space of digitalization."

While several participants discussed industry reliance on paper for data collection, one participant specifically noted factors contributing to this reliance and the varying pace of digitalization, attributing these differences to company size and profitability. Larger companies

were described as having greater capacity to invest in digital infrastructure, while smaller companies often faced constraints that slowed their transition away from paper-based systems, stating:

"The big trend would be larger companies that can buy software programs generally will have digital, the midsize will overwhelmingly have digital, the smaller ones will use more like an Excel sheet or a Google sheet, and then the very smallest, they may use those sheets, but they also may just write things down. It really is tied to the size of the company and how much profit, I guess, they make."

Most participants acknowledged a general trend toward digitalization in the food industry. However, one participant pointed out that despite the potential progress, there is little motivation for digitalization in an environment that resists data sharing. They explained that this resistance is what keeps the industry relying on outdated practices, saying:

"I'll answer it this way. I think that it can be digitalized, but I think in general, industry does not digitalize, and I think we're still working in our old ways of working, and we're still managing via Excel spreadsheets, I hate to say that, but we still are. And the lab reports, we're still managing that way. That doesn't mean we have to. But right now, there's such a resistance to openly sharing data that there's no excitement or line of sight to digitalizing."

While technological advancements are often seen as a key component of improving data sharing practices, one participant shifted the focus to the challenges of standardization. They emphasized that the real difficulty lies not in adopting new technologies, but in aligning industry data practices:

"I don't think the promise is so much around the technology advances as it is around data practices. We've, in our organization, had many debates and conversations that the technology part is easy, it's the getting the data into the state it needs to be, is the heavy lifting and what takes resources and funding to do so. I think it's really more about setting the standards and the ability to meet those standards, especially in places where you have legacy systems in place."

Participants highlighted the need for data standardization, such as standardized protocols for data collection, nomenclature, and formatting, which is required for meaningful inference

from shared data. They described how inconsistencies in data structures across companies could make the interpretation of shared data difficult. One participant noted:

"A good example that I commonly use in the conversation is that the way we capture what an ingredient or what a product or material is called within our organization is not the same way that someone outside of our organization would capture that same ingredient, product or material. How do you know that those two things are the same thing? Or how does someone on the other side who's receiving that data in the exchange, how do they interpret that information? Or at what level you're providing that data?"

Beyond differences between companies, one participant also highlighted inconsistencies in data formatting practices within their own organization, noting variations in data collection practices.

"I think we even have a challenge within our business around how we standardize data, let alone between different businesses, units, decimal places. I think it's a minefield. And I think if we are going to be sharing data, we got to figure out a clearer nomenclature around some of these things that does probably involve some element of standardization and a few key things. Or are we going to always be cleaning up data?"

Participants also expressed concerns about potential discrepancies in food safety programs such as environmental monitoring programs (EMPs). Specifically, the participants worried that positive test results could be misjudged, as it would be unclear whether they reflected actual safety issues or simply more rigorous testing. As one participant noted:

"Generally, people that do test and do get positives, do the right actions, but sharing that information becomes problematic because if there's, one entity who never finds a positive, does that mean that they're good or that they're not looking in the right manner? And one who finds a bunch of positives, are they a bad entity or are they doing the right thing and they're testing in the right manner and they're actually then thus keeping out the right amount of product from commerce."

Theme 3: Trust, or lack thereof, is a key factor influencing data sharing decisions, often underlying specific concerns such as confidentiality, liability risks, and data misuse

Trust emerged as an overarching theme influencing data sharing decisions within the food industry. Multiple participants identified trust as the primary factor in determining whether companies are willing to share data or not. In this context, participants described trust-related

concerns, particularly regarding the risks and vulnerabilities they perceived. Some framed it as uncertainty about how their data would be used, or if the privacy and confidentiality safeguards in place would be sufficient to ensure anonymity. One participant noted explicitly that the biggest hurdle to data sharing is trust:

" But it's going to take another issue which is not scientific, which is not data, which is trust. "

Another participant highlighted the primary challenge in data sharing as the human element rather than technological limitations. They noted that concerns about being identified contribute to data sharing reluctance, with decisions often shaped by emotional responses rather than purely rational considerations. Additionally, the same participant pointed out an underlying conflict between the fear of being identified and the goal of protecting consumer health:

"And that's what I've seen doing this and trying to apply it is again the trust that you will not be identified. That, I think, is the biggest hurdle to get over. We could do it technologically now, but the problem is the human reaction, and the trust has to happen. We always have to remember there's a human factor to making these decisions. And it becomes emotional at times, not necessarily rational either, or logical, but the human component to making decisions about this has to be considered, because the technology exists to do everything that you want to do right now, and to have it blinded or everything else. But I think we have to think about what's the human reaction that keeps us from taking that step, especially when we know that, like I said, most people want to protect public health. "

As trust emerges as a key factor in data sharing practices, participants discussed a variety of concerns that influence their decisions. We found that these concerns, which are often intertwined, reflect both economic and social barriers to food safety data sharing within the industry. For example, participants expressed a reluctance to share data due to concerns about how the data might be used once they engage in data sharing. Many feared that once data was shared, they would no longer have oversight, making them vulnerable to unintended consequences of engaging in data sharing. One participant noted:

"Once I give that data away, unless I'm absolutely confident that it's protected, it can be used as a weapon against me. So as long as I don't give it away, they may have a gun, but I haven't put the bullets in it, too."

Some participants expressed concerns that data could easily be taken out of context, noting that even if corrective actions were properly implemented, those reviewing the data might only focus on whether a company had positive test results, leading to negative economic consequences. As one participant explained:

"Well, the first thing that comes to mind of why we wouldn't share is just so that I guess if any data was out of the specification or the normal range, what kind of impact would that have on our sales, on our reputation, on the ability to grow? And often times any data or information can be misinterpreted or taken out of context where even though a corrective action was in place and worked correctly, if all the information isn't shared, then the appearance of what the data may be representing could be wrong and detrimental to a company."

Others mentioned their reluctance to share data despite having good practices in place, due to fear of data being misinterpreted or being unfairly implicated. One participant highlighted that, in conversations within the food industry, this fear is common, with many not wanting to share data to avoid potential implications, even when their practices are sound:

"...when you speak to people [in the] industry they say, well, I'd rather not share in case I'm implicated, you know, despite the fact that they have good processes and everything in place and they have their corrective actives, and corrections and they have a good food safety management system in place, you could get a misinterpretations of the data"

A prime example of this fear was highlighted in relation to regulatory bodies: that they might somehow access and identify the data sharer and use it to impose increased oversight or regulatory action. They worried that even well-intended transparency could lead to unintended consequences, such as increased scrutiny or regulatory action based on misinterpretations of their data. Some participants explained that efforts to voluntarily share information were overshadowed by how government entities perceived their data, reinforcing their hesitation to share. One participant said:

"I've had a couple of experiences where when you offer up those voluntary information and next thing you know, sure enough, you get a phone call inquiring about it. It's not a positive phone call when inquiring about what's going on. And you know the old adage, we're here to help, but that's not always the case because there's a financial penalty if items are found. So, a company is opening themselves up for scrutiny. And while that might not be the cause of an issue while they're looking they find something else they don't like. And I've seen that happen several times and I think there should be some guardrails or guidelines around that as well. And again, it gets back to trust."

In addition to regulatory fears, participants also expressed apprehension about how customers might react if they had access to certain data. The primary concern was that, without proper context, customers could misinterpret the information, leading to loss of patronage or unnecessary reputational harm.

"Other aspects of the industry, especially the purchasing area of an industry, if they learn about a risky area or specific subsection of a region, they can quickly say, 'I don't want anything from that area and that becomes problematic, right? So of course, people want to protect their own data because they would rather fix something versus being eliminated from a marketplace, let's say. So that becomes an area of concern. Why would I share my data if you're just going to use it against me?'"

Concerns about data sharing extend beyond regulatory action against companies and scrutiny or loss of customers, with participants also expressing fears about how shared data could be leveraged within the industry itself. Specifically, there was a recurring concern about whether companies could safeguard their competitive advantage while engaging in data sharing initiatives. Some participants highlighted the risk of inadvertently harming other businesses or being harmed themselves due to the way data is used or interpreted.

"For that kind of data sharing to happen, the first one is to have legal involved, unfortunately or fortunately, because there's probably aspects that I'm not aware of. What are the implications of company X sharing their food safety data with company Y? Well, this could happen, this could happen; we could unintentionally try to sabotage each other. I know that's an awful thing to say, but it's very competitive. So, if Company X says, oh, you got Listeria hits, okay. So, how do we protect each other from using that information against each other to have a competitive edge in a very competitive landscape?"

A related concern centered on questions of ownership, such as who owns the rights to the insights derived from shared data? One participant raised the issue of how companies could use shared data to develop proprietary information/intellectual property, potentially giving advantage over others who contributed to the dataset.

"I think, concerns around how the insights might be used to benefit some of the players more than others. You know, if you submit your data and then somebody takes the data and creates a proprietary tool that they own, that you haven't got a stake in, I think that would be, yeah. So, I think there might be a lack of trust about what happens to the insights, who owns the insights? I think the legal framework could be very tricky. And it has to be really clear that either everyone has or no one has the right to create intellectual property from the data."

Additionally, some participants highlighted a persistent mindset in the industry that views food safety as a competitive advantage rather than a shared responsibility.

"I think another challenge is potentially not everyone's on the same page with food safety not being a competitive advantage."

While concerns about data sharing were often framed as industry-wide hesitations, some participants noted that the resistance frequently stemmed from legal teams and executive leadership rather than from individuals in food safety or quality roles, as shown below:

"I see the benefit to data sharing. But what underpins it all is an understanding of, I think, trust between the organizations... There are companies that do share with each other, but it is, not as open because I think there's a fear of, if I give all my information, it's usually not the food safety individuals or quality. It usually comes from a legal or a higher leadership standpoint that says there's proprietary information here. They are technically competitors; we shouldn't share more than we need to with industry."

Many participants expressed concerns about the ability of existing technology to ensure the security of shared data, reflecting a lack of trust in the current systems. While technology is often seen as a key enabler for data sharing, participants doubted whether it could adequately protect their confidential food safety data. One participant raised a trust-related concern about whether regulators or legal entities could access the shared data during investigations or lawsuits,

knowing that the data clearing house holds multiple industry data points, even if their company was not directly at fault.

"If there were technological lockdowns that gave sufficient confidence that those concerns would not happen, we would be more inclined to participate in a data sharing event. But in reality, if our company is involved in a lawsuit, that information is going to be discoverable. But I can control how that information is released to plaintiff's counsel. I can control what is released. But so having that control is important to us. It's our data. We own it."

Additionally, when asked about knowledge of existing privacy-preserving technologies for data sharing, most participants explained not being aware of any such tools:

No, I don't. I don't have that knowledge. That's too much technology for me.

Another participant highlighted the need for robust privacy protections to encourage data sharing, emphasizing the concern that without strong security measures, companies would be hesitant to participate in such data sharing initiatives.

"So, if you, if we want to create a bit of a level playing field and encourage industry partners to engage, we will have to provide privacy protection for individual companies. And build trust, I think, in what the intent is and what it isn't is equally as important, if not more important, around how the data is going to be used and how we can anonymize them and protect people's privacy."

Theme 4: The need exists for sustained collaboration between academia, government, and industry to identify trade-offs in data sharing, establish data governance, and enable improved food safety through information pooling.

Participants proposed several strategies to overcome barriers to data sharing, focusing on the importance of third-party governance, problem-specific objectives, and fostering collaboration. A common solution proposed was the involvement of third parties, such as academic institutions, to facilitate data sharing as well as spearheading data governance with related privacy-preserving technologies. One participant explained that:

"There's probably something we can do to build some confidence in like, I think there's some governance that could be applied to data sharing that would generate some comfort on those that are a little bit paranoid about, like, what they're sharing or where it's going. I think there's some governance we could apply to this idea that could be generated by, like, a third independent party, a neutral party that really does not win or lose or gain in any of this. They're just governing for us."

Additionally, some participants highlighted the importance of having a safe space, in this context, third-party legal experts, with whom companies could consult on food safety and regulatory-specific questions. As one participant said:

"But we've also developed relationships with third-party legal folks that have formerly worked for government. So, that's a nice partnership for us to kind of talk through these exact things about, like, how much can we share? You know, what would you recommend for us to share so that we can, again, better ourselves and better our suppliers?"

Another key solution discussed was the need for problem-specific data sharing objectives. Several participants stressed that defining clear and focused goals would make the data sharing process more practical and transparent. As one participant emphasized, data sharing should not be done just for the sake of sharing but rather to address specific problems within the right context:

"And yeah, so, and we shouldn't be sharing data just to share data. If we're sharing data to solve a problem, that is the right context. But then again, it also needs to be protected and because sharing data, sharing again in today's environment, just sharing to share is, there's being transparent and then there's being like naked walking down the street."

Similarly, participants expressed the need for clear incentives to participate in data sharing initiatives. While they acknowledged the benefits to public health, they noted a lack of incentives discourages them from participating. As one participant explained:

"If somebody can articulate the benefit to the person who's putting the data in, not a nebulous public health, 'it's good for public health,' okay, but public health? I'm feeding people, that's public health, I'm doing lots of things that are public health. The downside of this public health contribution is, I could get sued for billions of dollars and put myself out of business. I'm not going to do it. There's no incentive right now..."

The final solution proposed by participants centered around the belief that effective data sharing within the food industry is possible, but significant efforts are required to develop guidelines for these initiatives in a way that fosters trust. Participants emphasized that careful planning and clear frameworks are essential for building confidence in data sharing. One participant stated:

"I fundamentally believe that that [data sharing] can happen. You just need to put some more structure around it such that people have more trust in it."

In addition to building trust, participants noted the importance of engaging with entities that the industry often views as adversarial, such as regulatory bodies that may gain access to the confidential data shared within the industry.

"And maybe having more collaborative conversations with FDA on how would they use that data and what regulatory risk might exist. I know some of those conversations are being had with FDA within certain divisions, certain centers of FDA. I think more effort is needed to understand what are the expectations of the federal agencies."

Lastly, participants emphasized that a critical part of making data sharing feasible involves ensuring robust confidentiality and security measures. They acknowledged the potential risks of data misuse and the fear that once shared, information could be accessed by unauthorized parties, including regulatory bodies or the customers, despite intentions for it to remain within the industry. One participant noted:

"I think it would take significant work to ensure the confidentiality of the data and the security of the data so it couldn't be misappropriated. I think there would also be a need to work with regulatory authorities to align how the data would be used and what the expectations of the authorities would be relative to the data that they receive. I think these would be the primary concerns that would need to be addressed in order for data to be shared more freely. Particularly because data without context is not necessarily so useful from, particularly from, an enforcement perspective."

These insights reflect a belief that data sharing in the food industry is possible if the right frameworks are in place to address data governance, security, and trust. To identify barriers that

can be addressed by data governance, we mapped the risks associated with data sharing mentioned by participants against domains of data governance: data principles, quality, metadata, access, and lifecycle (Table 3). This mapping shows that participants' concerns span multiple governance domains. These findings suggest that many perceived risks are linked to governance structures.

Table 3. Mapping risks associated with data sharing raised by participants against the data governance domains

Data Governance Domain	Definition	Risks Identified by Participants											
		Ownership	Loss of control	Misinterpretation	Misuse	Bad publicity	Poor data quality	Taken out of context	Unfairly implicated	Privacy	Security (breaches)	Punitive action	Loss of sales
Data principles	Establish the link with business operations, assign ownership, treat data as an enterprise-wide asset, guide policies, promote data sharing, and define roles	*	*						*				
Data quality	Ensures data meets specific usage requirements, including accuracy, timeliness, completeness, and credibility, for its use to be effective			*	*	*	*						
Metadata	Metadata relates to maintaining concise and consistent data description to aid in the interpretation of data			*		*	*	*	*			*	*
Data access	Involves assigning value to different data and ensuring confidentiality, integrity, and availability through risk analysis and compliance with legal standards	*	*		*	*			*	*	*	*	*
Data lifecycle	Understanding how data is used, usage duration, retention, and storage				*					*	*	*	

4. Discussion

The attitudes, perceptions, and practices of food industry leaders regarding the voluntary sharing of confidential food safety data reflect a complex interplay of personal beliefs, industry norms, legal landscape, and lived experiences. Understanding these dynamics is crucial for enabling data sharing within the industry. Our main findings reveal that: i) companies face trade-offs when deciding to share data, weighing risks against benefits, ii) the level of trust that the data sharer has in the stakeholders (e.g., among the industry peers, regulatory bodies, and customers) and technology strongly influences voluntary data sharing decisions, and iii) participants highlighted data governance as a solution to concerns that stem from trust such as loss of control once data is shared. These findings are elaborated in the following paragraphs.

The Trade-offs and Complexities of Data Sharing Decisions

Our interviews with food industry leaders highlight their perceptions of the efforts required, the potential benefits, and the risks involved in engaging in food safety data sharing. We begin by examining the efforts and investments needed for companies to engage in data sharing as identified by participants. The first is the need to digitize existing records of food safety data for sharing purposes and to transition to a digital approach for data collection, which participants noted is still largely paper-based. An interview study with food industry stakeholders conducted by Alexander et al. (2023) supports this statement and shows that some companies within the food industry delay digitizing their data due to limited time and financial resources to convert paper to digital records.¹⁶ Although Alexander's study was about barriers in adopting new artificial intelligence (AI) technologies, we believe that similar constraints may apply to data

sharing efforts. Therefore, efforts to offset the financial constraints experienced by companies could help accelerate the digitalization process within the food industry.

Additionally, participants emphasized the need for data standardization (defined as the establishment of standards across the data value chain to ensure consistency and interoperability)¹⁷ for effective data sharing. Qian et al. (2025) demonstrate the utility of data standardization by developing a public repository that uses three food safety datasets from different sources.¹⁸ Authors had to align and preprocess these data formats to create benchmark datasets, facilitating the development of machine learning models for food safety. Mahmood (2025) describes how the lack of data standardization and the confinement of biomedical data behind institutional firewalls hinder progress in biomedical machine learning, leading to variability in how models are evaluated.¹⁹

In foodborne pathogen surveillance, swab site metadata is often recorded as unstructured free-text, leading to inconsistent descriptions and poor data interoperability. To address this, researchers developed a structured, machine-readable schema that standardizes swab site metadata.²⁰ This schema has been incorporated into the One Health Enteric Package at NCBI BioSample, a public metadata standard supporting pathogen surveillance across human, animal, and environmental health.²¹ Similarly, FoodOn, an open source ontology, provides a standardized vocabulary for describing food products, their sources, processing methods and packaging to enhance data sharing and traceability.²² While such efforts demonstrate progress, standardization challenges remain prevalent at both the intra- and intercompany levels as stated by our interview participants. Without broader development and adoption of standardized practices, inconsistent data collection and nomenclature will continue to hinder data sharing efforts in the food industry and compromise the reliability of collaborative analyses. Implementing standardized data

practices across the food industry is therefore essential to ensure the reliability and interoperability of shared data. A recent review highlighted that the successful application of Artificial Intelligence (AI) in food safety relies heavily on the availability of high-quality, standardized data,²³ emphasizing the need to address these foundational challenges to unlock AI's full potential in transforming food safety.

Efforts aside, several participants noted that the potential benefits of data sharing include improved public health outcomes due to access to more data points. They explained that this access could enhance predictive modeling capabilities, enabling early identification and mitigation of rare food safety incidents that could lead to foodborne outbreaks. Similarly, voluntary data sharing in the swine industry, led by academic institutions, enables predictive models for forecasting disease outbreaks and informing prevention strategies, highlighting the value of data sharing in protecting animal health.²⁴ Additionally, our participants noted that small businesses could benefit from data sharing insights as they often lack the resources or historical data that large companies have to guide decision-making processes. Illustrating a shared challenge with small businesses having limited data access, a participatory action research study on small business marketing reports that access to retailer loyalty card data helped bridge the information gap between small and large businesses.²⁵ Similarly, sharing food safety data could empower small companies to make more informed and improved risk mitigation efforts. An implicit benefit of data sharing is the creation of larger, more comprehensive datasets that support the application of AI and machine learning to advance food safety goals, provided aggregated data are accurate, unbiased, and representative of the intended use cases.²⁶ This underscores the importance of understanding what the aggregated data represent before drawing inferences to realize the benefits of data sharing fully.

While data sharing has many recognized benefits, participants highlighted that it has inherent risks that could expose companies to negative consequences if sensitive data were leaked or used for an unintended purpose. Specifically, participants were concerned about the possibility that leaked data could be misinterpreted by customers, leading to reputational harm. Bernal (2024) discusses how companies face reputational risks even when they share data for socially beneficial purposes, a practice referred to as data philanthropy.^{27,28} Bernal argues that companies may still suffer unintended consequences if the public misinterprets the shared data. Although Bernal's paper focuses on business-to-government (B2G) data sharing, similar concerns apply in the industry-to-industry data sharing context.

Another concern expressed by participants was the potential for retrospective or prospective liability. For example, participants talked about the risks associated with whole genome sequencing (WGS) data, which could implicate companies in past events or outbreaks yet to be detected, exposing them to legal liability and increased regulatory scrutiny. Although empirical research on this remains limited, a trade journal article by Amezcua et al. (2025) suggests the food industry is hesitant to adopt WGS due to legal and regulatory uncertainties, particularly concerns about regulatory data sharing obligations.²⁹ These concerns underscore that without clear legal or regulatory assurances, companies may prioritize complete avoidance of risks associated with data sharing.

Lastly, participants were concerned about losing their competitive advantage if they engaged in data sharing. Similarly, a survey and interview of fish farmers in Norway reveals that they were hesitant to share data because they feared that sharing would result in a loss of competitive advantage, especially if there was no sharing reciprocity.³⁰ Overall, our findings highlight a broader tension between data sharing to improve public health and business survival concerns.

The efforts, investments, benefits, and risks presented above highlight a misalignment between public health priorities, individual company interests, and broader industry needs. The benefits of improved data sharing would be shared by all stakeholders, including consumers, companies, and the food industry as a whole, but the associated risks and investments are often expected to be borne primarily by individual companies. As one participant said, “... *I'm feeding people, that's public health, I'm doing lots of things that are public health. The downside of this public health contribution is I could get sued for billions of dollars and put myself out of business...*” This statement underscores the extent to which participants perceive data sharing as a potentially existential risk, capturing what they mean when they speak of economic harms.

An interview study of experts from industries, such as agriculture and automotive, reports that organizations tend to focus more on the risks rather than the potential benefits of data sharing.³¹ The authors highlighted that this attitude is due to uncertainties about the added value, both monetary and non-monetary (e.g., customer rewards and strategic market position), from data sharing.³¹ Using the privacy calculus theory (PCT) framework, a study with experts from European data sharing initiatives demonstrates how data sharing decisions are influenced by perceived benefits (e.g., access to more data) and barriers (e.g., lack of trust between actors), resulting in tensions between the perceived gains and risks.³² Our findings suggest the PCT framework could provide a useful lens for understanding how food companies navigate the complex decisions of data sharing in a food safety context.

Because food industry participants view the potential risks as existential, the risks will tend to outweigh any perceived benefits in the data sharing decision calculus. Therefore, future research should focus on mitigating the downside risk faced by individual companies. Such mitigation requires broadening the risk burden. One possible scenario is that the costs of, for

example, a lawsuit resulting from a harmful pathogen being discovered in shared data could be covered by the fund to which all participants contribute. An alternative would be that the government guarantees to cover legal costs or damages using public funds (i.e., taxpayer money). As with any insurance setting, the risk-sharing mechanism would need to address moral hazard, i.e., the fact that firms insured against the worst downside risks of data sharing would have less incentive to follow best practices. The potential joint liability could be reduced with appropriate legal and contractual terms.

In general, future research should aim to develop a clear and compelling value proposition for data sharing that demonstrates how the benefits can offset the perceived and actual risks at both the company level and across the broader industry. Lefebvre et al. (2025) emphasize the importance of creating a joint value proposition for companies involved in horizontal data sharing and provide examples of existing initiatives, such as Skywise in the aviation sector, whose >300 members come together to gain insights on plane component improvement and maintenance anticipation.³³ These efforts, when accompanied by trust-building initiatives, can facilitate food industry-wide voluntary data sharing to support better informed decision making.

The Role of Trust in Voluntary Food Safety Data Sharing within the Food Industry

Participant concerns about data sharing, rooted in a trust deficit, can be categorized into two groups: lack of trust among stakeholders (regulatory, customers, industry peers) and lack of trust in technology. While participants expressed that the government can play a role in facilitating data sharing, many participants (i) expressed concerns about how leaked data could be used against them by regulatory bodies and (ii) shared negative past experiences about how shared data was used against them. As one participant noted, *“I’ve had a couple of experiences where when you offer up those voluntary information and next thing you know, sure enough, you get a*

phone call inquiring about it. It's not a positive phone call when inquiring about what's going on...” The participant further noted that this negative experience occurred despite assurances from the government that the information would be used to improve food safety. Such experiences highlight breaches of trust and discourage future engagement in voluntary data sharing initiatives.

An interview study with regulators and food industry members about their roles in food safety in three countries (UK, Australia, and New Zealand) has explored the complexity in the relationship and communication between food regulators and the food industry.³⁴ The study showed that, on one hand, regulators perceive that the industry’s perception of the regulatory role is to cause problems for the industry, while, on the other hand, industry representatives report that the lack of knowledge or care on the part of regulators led to experiences where data was improperly handled and incorrect information was released to the media.³⁴ These experiences, along with the differing goals between regulators and industry, underscore the challenges in fostering a cooperative and trustworthy environment for voluntary data sharing.

While our participants highlighted that food safety is not a competitive advantage, the unregulated data sharing space provides room for potential unintended consequences if industry peers access data. One participant said, “...*We could unintentionally try to sabotage each other. I know that's an awful thing to say, but it's very competitive...*” highlighting fears that competitors might lose revenue or business as a side effect of data sharing. It is often stated that food safety should not be considered a competitive advantage because it is a non-negotiable condition, invisible to consumers, and yet expected. Also, food safety incidents often create shockwaves reverberating through the entire food industry sectors, meaning any one weak link weakens the whole industry. However, sharing food safety data may result in a loss of business advantage

relative to competitors³⁵ because shared data may allow inferences about proprietary production processes or performance metrics. This highlights the importance of carefully regulated accessibility³⁶ of voluntarily shared food safety data. Otherwise, even well-intentioned transparency could compromise competitive positioning by exposing sensitive operational details.

Our participants also expressed concerns about the potential for the loss of customer trust. Specifically, participants were concerned that if food safety data were leaked or made public without proper context, customers might misinterpret the data, particularly given the technical nature of such data. While participants did not explicitly question customers' ability to interpret such data, their concerns suggest an underlying fear that technical food safety data could be misinterpreted when viewed outside its intended context. Such misinterpretation could lead to significant economic consequences, particularly if the information were perceived as evidence of negligence or risk, even when that wasn't the case. Using the protection motivation theory, which explains consumer behavioral changes in response to health risks such as food safety risk, Liao et al. (2020) show the impact of food recalls events on consumer trust and company reputation, and the role that trust plays in food safety and how violations of this trust can lead to significant reputational damage and loss in sales.³⁷ Similarly, misinterpreted leaked data can lead to a loss in sales as customers may choose to take their business elsewhere due to perceived risks. Both our findings and those of Liao et al. (2020) highlight the need for robust data privacy and security measures to prevent unauthorized access and release of shared data.

Trust in technology also emerged as a critical factor limiting data sharing. Participants questioned whether the existing data systems are secure, reliable, and capable of safeguarding their shared data at all costs. One participant said that they do not believe that the current

technological lockdowns provide enough confidence, and the consequences of poor data security protection would become a reality. Most study participants were unaware of technological and AI advancements in data security, such as differential privacy, federated learning, or blockchain. For example, Gai et al. (2023) emphasizes the importance of robust security measures for protecting shared data.³⁸ They discuss implementing zero-trust principles and blockchain technology to enhance data security and address concerns about the reliability of existing data systems.³⁸ Overall, our findings highlight the need to increase awareness among the food industry stakeholders about emerging data sharing technologies and AI that ensure data security and privacy. Additionally, there is a need for technological infrastructure that restricts data access to those directly involved in a sharing agreement to further facilitate trust among stakeholders.

Data Governance as a Solution to Data Sharing Concerns

Participants expressed that barriers to voluntary data sharing can be addressed through the development of a robust data governance framework. Participants said “... *I think there's some governance that could be applied to data sharing that would generate some comfort on those that are a little bit paranoid about what they're sharing or where it's going...*”. Cue et al. (2021) similarly highlight the need for data governance to foster trust and propose the Farmers Bill of Rights to address data sharing concerns in the dairy industry.³⁹ The bill stipulates that farmers retain ownership of their data and mandates transparent communication of intended use by data users (industry).³⁹ This aligns with the broader conversation on data governance, which the Data Governance Institute (DGI) defines as “...a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.”⁴⁰ This definition underscores the importance of having

clear and structured frameworks that ensure that all stakeholders understand the data sharing rules of engagement. Khatri and Brown (2010) outlined five interrelated data governance domains: data principles, quality, metadata, access, and lifecycle, each addressing key aspects of data management.¹⁴ Our data reported here suggest that those domains would address a number of risks associated with data sharing identified by participants (Table 3).

If data governance has the potential to address the risks associated with data sharing, the question is why such frameworks have not yet been implemented. We drew insights from a systematic literature review, which identified several barriers to the development of effective data governance.⁴¹ Barriers relevant to the industry-to-industry food safety data sharing initiatives include a lack of knowledge about how to implement a governance framework, organizational barriers such as a lack of priority compared to other projects, and a lack of financial resources.⁴¹

To gain an understanding of how a data governance framework specific to industry-owned food safety data could be developed, we can analyze existing frameworks. One of them is the Health Insurance Portability and Accountability Act (HIPAA), which offers several parallels to sharing food safety data (such as environmental monitoring test results). Like health data, food safety data contains sensitive company safety status, has significant regulatory implications, and balances competitive interests with public health benefits. HIPAA's governance approach ensures that health data is stored, processed, and shared in compliance with country laws, and according to clearly defined data use agreements and limitations, standards for data de-identification, and authorization requirements that enable provider/patient control. An interview study highlights the importance of comprehensive policies, procedures, and compliance with HIPAA regulations in

fostering trust.⁴² Comprehensive rules of engagement, protections, and policies are similarly necessary for a food safety data governance framework.

The U.S. federal government played a central role in the development of HIPAA, and similar government involvement is necessary to offset the cost of the food safety data governance framework. However, unlike HIPAA, the participants in our study emphasized the need for oversight by “*a neutral party that really does not win or lose or gain*” from shared food safety data, and they suggested that academic institutions play that role. Our interviews also revealed that participants are concerned about losing control over their data once shared. This suggests that maintaining control would be an important motivator for data sharing and increasing participation in data sharing initiatives. As one participant stated, “... *But I can control how that information is released to plaintiff's counsel. I can control what is released. But so having that control is important to us. It's our data. We own it...*” Based on our findings, we recommend piloting a data governance framework for voluntary food safety data sharing, developed in collaboration with regulatory bodies and academic institutions, to evaluate its practical feasibility and benefits, its effectiveness in addressing concerns around trust and data misinterpretation, and its potential to support secure data sharing.

Limitations

While we took steps to minimize the potential introduction of biases, this study has some limitations. The sample was drawn from specific professional networks, which may have introduced selection bias or resulted in saturation faster due to shared experiences or perspectives within this group, overall limiting the generalizability of our findings. However, considering that the sample included participants from a wide range of sectors in the food industry, we believe the

findings are representative of the target population. The semi-structured interview format allowed for in-depth exploration of participant responses, but this also meant that prompts were sometimes given in different orders, and time constraints prevented the exploration of all scripted questions. With interviews, there is a potential for social desirability bias, as participants may have provided responses they perceived as more acceptable or favorable, especially given their leadership roles within the industry. A graduate student, rather than an established food safety authority, conducted all interviews in an attempt to minimize this potential source of bias. While we emphasized confidentiality, this bias may still have influenced the data. There is also the potential for confirmation bias in the data analysis process, which could result in biased interpretation of responses and an emphasis on evidence that confirms our preexisting beliefs over contradictory evidence. To mitigate this, a four-person team conducted the analysis, including the data familiarization, ensuring thorough discussion and validation of emerging themes against interview transcripts. Finally, because the interviews were conducted between June 2023 and January 2024, the findings reflect the participants' knowledge and experiences, legal landscape, and the industry norms prior to or during that specific timeframe. Despite these limitations, our study provides valuable novel insight into food safety data sharing in the food industry and highlights key considerations for future research and policy.

To conclude, our main findings reveal an interesting paradox: the food industry's enthusiasm for collaborative safety improvements coexists with systemic barriers that prevent the realization of voluntary sharing of confidential food safety data. By identifying incentives and obstacles that drive industry decision-making, our results provide an evidence-based roadmap for unlocking the potential of collective food safety decision-making. This work

establishes the foundation for voluntary data sharing initiatives that could transform the industry's approach to preventing foodborne illness through the collaborative analysis of aggregated food safety data.

Acknowledgements

This study was funded by the Agriculture and Food Research Initiative from the USDA National Institute of Food and Agriculture (NIFA), Competitive Grant no. 2020-67021-32855/project accession no. 1024262 (RI, MW, AS, AA) and the Cornell Institute for Digital Agriculture (CIDA) (RI, MW, LK). The funders had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript. The authors thank the interviewed study participants for their time and participation.

Supporting Information

- S1. Semi-structured interview guide
- S2. Code saturation analysis [Figures S1 and S2]
- S3. Binary-coded matrix of all codes and interviews
- S4. List of interview quotes organized by theme and theme aspect

References

1. CDC. About Food Safety. Food Safety. May 17, 2024. Accessed May 13, 2025. <https://www.cdc.gov/food-safety/about/index.html>
2. Review of the Boar's Head *Listeria monocytogenes* Outbreak - January 2025 | Food Safety and Inspection Service. Accessed June 20, 2025. <http://www.fsis.usda.gov/news-events/publications/review-boars-head-listeria-monocytogenes-outbreak-january-2025>
3. CDC. Salmonella Outbreak Linked to Cucumbers - June 2024. *Salmonella Infection* (Salmonellosis). March 4, 2025. Accessed May 13, 2025. <https://www.cdc.gov/salmonella/outbreaks/africana-06-24/index.html>
4. GreenLink – Utilize your data. Transform your industry. Accessed May 14, 2025. <https://www.greenlinkdata.com/>
5. Kaur J, Hazrati Fard SM, Amiri-Zarandi M, Dara R. Protecting farmers' data privacy and confidentiality: Recommendations and considerations. *Front Sustain Food Syst.* 2022;6. doi:10.3389/fsufs.2022.903230
6. Gottschald M. Advancing food safety through digital traceability, interoperability, harmonized data and collaborative partnerships. *J Consum Prot Food Saf.* 2024;19(3):257-258. doi:10.1007/s00003-024-01522-8
7. Ruslin, Machuri, Rasak, Alhabsyi, Syam. Semi-structured Interview: A Methodological Reflection on the Development of a Qualitative Research Instrument in Educational Studies. *IOSR-JRME.* 2022;12(1). <https://www.iosrjournals.org/iosr-jrme/papers/Vol-12%20Issue-1/Ser-5/E1201052229.pdf>
8. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Qual Quant.* 2022;56(3):1391-1412. doi:10.1007/s11135-021-01182-y
9. Namey E, Guest G, McKenna K, Chen M. Evaluating Bang for the Buck: A Cost-Effectiveness Comparison Between Individual Interviews and Focus Groups Based on Thematic Saturation Levels. *American Journal of Evaluation.* 2016;37(3):425-440. doi:10.1177/1098214016630406
10. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. *PLoS One.* 2020;15(5):e0232076. doi:10.1371/journal.pone.0232076
11. Sharma SK, Mudgal SK, Gaur R, Chaturvedi J, Rulaniya S, Sharma P. Navigating Sample Size Estimation for Qualitative Research. *Journal of Medical Evidence.* 2024;5(2):133. doi:10.4103/JME.JME_59_24
12. Hagaman AK, Wutich A. How Many Interviews Are Enough to Identify Metathemes in Multisited and Cross-cultural Research? Another Perspective on Guest, Bunce, and Johnson's (2006) Landmark Study. *Field Methods.* 2017;29(1):23-41. doi:10.1177/1525822X16640447

13. van Panhuis WG, Paul P, Emerson C, et al. A systematic review of barriers to data sharing in public health. *BMC Public Health*. 2014;14(1):1144. doi:10.1186/1471-2458-14-1144
14. Khatri V, Brown CV. Designing data governance. *Commun ACM*. 2010;53(1):148-152. doi:10.1145/1629175.1629210
15. Basharat A, Bewong M, Krivokapic-skoko B, et al. The role of data governance in addressing potential risks associated with agricultural data exchanges: A systematic literature review. Published online June 12, 2024. doi:10.2139/ssrn.4863615
16. Alexander CS, Smith A, Ivanek R. Safer not to know? Shaping liability law and policy to incentivize adoption of predictive AI technologies in the food system. *Front Artif Intell*. 2023;6. doi:10.3389/frai.2023.1298604
17. Gal M, Rubinfeld DL. Data Standardization. Published online June 1, 2019. doi:10.2139/ssrn.3326377
18. Qian C, Yang H, Acharya J, Liao J, Ivanek R, Wiedmann M. Initializing a Public Repository for Hosting Benchmark Datasets to Facilitate Machine Learning Model Development in Food Safety. *Journal of Food Protection*. 2025;88(3):100463. doi:10.1016/j.jfp.2025.100463
19. Mahmood F. A benchmarking crisis in biomedical machine learning. *Nat Med*. 2025;31(4):1060-1060. doi:10.1038/s41591-025-03637-3
20. Feng J, Daeschel D, Dooley D, et al. A Schema for Digitized Surface Swab Site Metadata in Open-Source DNA Sequence Databases. *mSystems*. 8(2):e01284-22. doi:10.1128/msystems.01284-22
21. CFSAN-Biostatistics/One_Health_Enterice_Package. Published online December 9, 2024. Accessed June 6, 2025. https://github.com/CFSAN-Biostatistics/One_Health_Enterice_Package
22. Dooley DM, Griffiths EJ, Gosal GS, et al. FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. *npj Sci Food*. 2018;2(1):23. doi:10.1038/s41538-018-0032-6
23. Dhal SB, Kar D. Leveraging artificial intelligence and advanced food processing techniques for enhanced food safety, quality, and security: a comprehensive review. *Discov Appl Sci*. 2025;7(1):75. doi:10.1007/s42452-025-06472-w
24. Perez AM, Linhares DCL, Arruda AG, et al. Individual or Common Good? Voluntary Data Sharing to Inform Disease Surveillance Systems in Food Animals. *Front Vet Sci*. 2019;6. doi:10.3389/fvets.2019.00194
25. Donnelly DC, Simmons DG, Armstrong DG, Fearne PA. THE ROLE OF ACTION RESEARCH IN THE STUDY OF SMALL BUSINESS MARKETING AND RETAILER

LOYALTY CARD DATA. Published online 2012.
<https://mural.maynoothuniversity.ie/id/eprint/4091/>

26. Dimitrakopoulou ME, Garre A. AI's Intelligence for Improving Food Safety: Only as Strong as the Data that Feeds It. *Curr Food Sci Technol Rep.* 2025;3(1):15. doi:10.1007/s43555-025-00060-0
27. Bernal J. Private sector trust in data sharing: enablers in the European Union. *Data & Policy.* 2024;6:e30. doi:10.1017/dap.2024.20
28. Lev Aretz Y. Data Philanthropy. *SSRN Journal.* Published online 2019. doi:10.2139/ssrn.3320798
29. Amezquita et al. The Benefits and Barriers of Whole-Genome Sequencing for Pathogen Source Tracking: A Food Industry Perspective | Food Safety. Accessed April 23, 2025. <https://www.food-safety.com/articles/6696-the-benefits-and-barriers-of-whole-genome-sequencing-for-pathogen-source-tracking-a-food-industry-perspective>
30. Sønvisen SA, Lilleng G, Syversen T, Vatn DMK. Data sharing in the fisheries: Exploring the willingness to share data in the Norwegian fishing fleet. *Marine Policy.* 2025;175:106620. doi:10.1016/j.marpol.2025.106620
31. Fassnacht M, Benz C, Heinz D, Leimstoll J, Satzger G. Barriers to Data Sharing among Private Sector Organizations. In: ; 2023. doi:10.24251/HICSS.2023.453
32. Jussen I, Schweihoff J, Möller F. Tensions in Inter-Organizational Data Sharing: Findings from Literature and Practice. In: *2023 IEEE 25th Conference on Business Informatics (CBI).* ; 2023:1-10. doi:10.1109/CBI58679.2023.10187530
33. Lefebvre H, Krasikov P, Legner C, Flourac G. Data management as a joint value proposition—A design theory for horizontal data sharing communities. *Electron Markets.* 2025;35(1):21. doi:10.1007/s12525-025-00755-1
34. Meyer SB, Wilson AM, Calnan M, et al. In the interest of food safety: a qualitative study investigating communication and trust between food regulators and food industry in the UK, Australia and New Zealand. *BMC Public Health.* 2017;17(1):189. doi:10.1186/s12889-017-4118-x
35. Qian C, Liu ,Yuhan, Barnett-Neefs ,Cecil, et al. A perspective on data sharing in digital food safety systems. *Critical Reviews in Food Science and Nutrition.* 2023;63(33):12513-12529. doi:10.1080/10408398.2022.2103086
36. Manning L, Brewer S, Craigon PJ, et al. Artificial intelligence and ethics within the food sector: Developing a common language for technology adoption across the supply chain. *Trends in Food Science & Technology.* 2022;125:33-42. doi:10.1016/j.tifs.2022.04.025

37. Liao C, Luo Y, Zhu W. Food Safety Trust, Risk Perception, and Consumers' Response to Company Trust Repair Actions in Food Recall Crises. *International Journal of Environmental Research and Public Health*. 2020;17(4):1270. doi:10.3390/ijerph17041270
38. Gai K, She Y, Zhu L, Choo KKR, Wan Z. A Blockchain-Based Access Control Scheme for Zero Trust Cross-Organizational Data Sharing. *ACM Trans Internet Technol*. 2023;23(3):38:1-38:25. doi:10.1145/3511899
39. Cue R, Doornink M, George R, et al. Data Governance in the Dairy Industry. *Animals*. 2021;11(10):2981. doi:10.3390/ani11102981
40. The Data Governance Basics. The Data Governance Institute. Accessed April 29, 2025. <https://datagovernance.com/the-data-governance-basics/>
41. Al-Ruithe M, Benkhelifa E. Analysis and Classification of Barriers and Critical Success Factors for Implementing a Cloud Data Governance Strategy. *Procedia Computer Science*. 2017;113:223-232. doi:10.1016/j.procs.2017.08.352
42. Sivaraj KP, Goel L. Data Governance Strategies in Healthcare: Ensuring Compliance with HIPAA Standards. 2025;14(01).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [S1.Interviewguide.docx](#)
- [S2.CodeSaturationAnalysis.docx](#)
- [S3.Binarycodedmatrixdata.xlsx](#)
- [S4.ThemesandThemeAspects.docx](#)