

Supplementary material for the article:

*Preparing Future Business Data Sharing via a Meta-Platform for Data Marketplaces:
Exploring Antecedents and Consequences of Data Sovereignty*

Appendix 1: Overview of participants of our twenty semi-structured interviews

Table 1 Overview of participants of our twenty semi-structured interviews

ID	Category	Job description	Overall Work Experience
I-01	Data provider	Director of innovation	28 years
I-02	Data provider	Security solution manager	18 years
I-03	Data provider	Product owner of a data platform	14 years
I-04	Data provider	Head of Standard Business Reporting	23 years
I-05	Data provider	Project manager	10 years
I-06	Data provider	Commercial director	24 years
I-07	Data provider	Chief data officer	12 years
I-08	Data provider	Technical innovation manager	28 years
I-09	Data provider	Data protection specialist	3 years
I-10	Data provider	Head of architecture, innovation, and tech.	16 years
I-11	Data provider	Senior strategy manager	32 years
I-12	Data provider	Product owner	11 years
I-13	Data provider	Risk manager	5 years
I-14	Data provider	Business consultant	22 years
I-15	Data provider	Associate director	24 years
I-16	Expert	Technical researcher	9 years
I-17	Expert	Deputy studio director	13 years
I-18	Expert	Data science director	12 years
I-19	Expert	Project manager	10 years
I-20	Expert	Project manager	9 years

Appendix 2: Semi-structured interview questions

1. Opening segment

Questions about background information.

- **Question 1:** What position do you hold in this company now, and for how long have you been in this position?
- **Question 2:** Could you please tell me briefly the nature of your job?
- **Question 3:** How familiar is your organization with business-to-business data sharing? Has your organization known, experienced, or thought about it?
 - *[if yes, probe for]* Could you explain how your company conducts B2B data sharing?
 - *[if yes, probe for]* What is your role in decision-making about B2B data sharing?
- **Question 4:** Could you please tell me briefly your experience related to data marketplaces?

2. Middle segment

Questions about exploring potential value proposition and hindrances of business data sharing via a meta-platform for data marketplaces.

Figure 1 illustrates the presented use case to the participant:

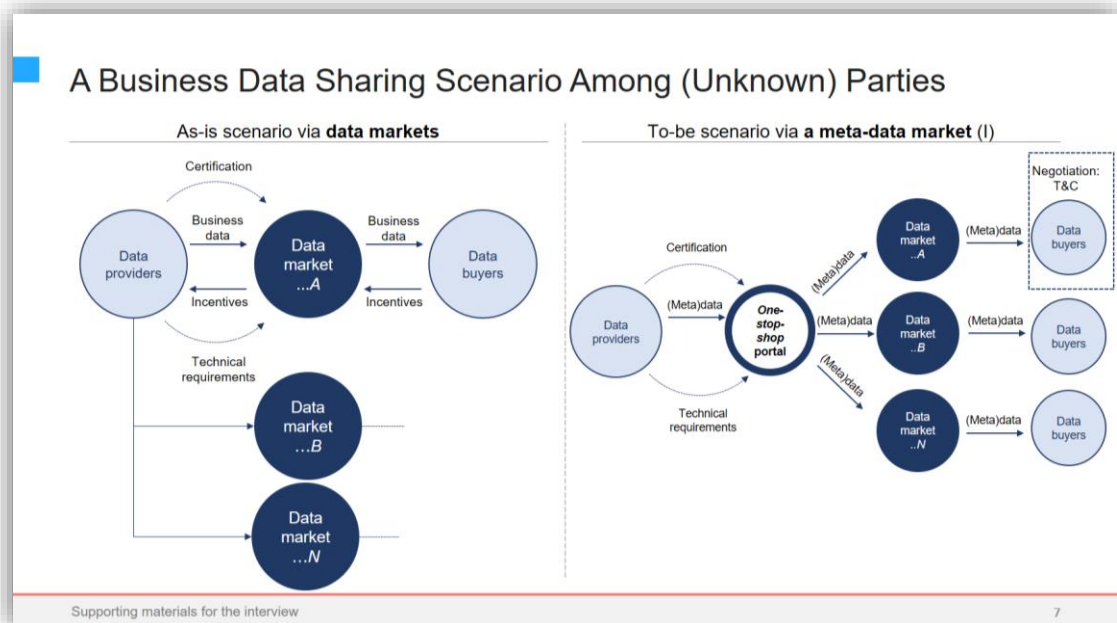


Figure 1 To-be business data sharing scenario via a meta-platform for data marketplaces

- **Question 5:** Do you have any questions related to the use case of a meta-platform for data marketplaces?
- **Question 6:** To what extent do you think a meta-platform for data marketplaces can facilitate B2B data sharing among organizations?
 - *[for greater extent, probe for]* Could you explain the reasons why it is the case?

- *[for less extent, probe for]* Could you explain why business data sharing via a meta-platform for data markets may not work?

3. **Closing segment**

Questions about data sovereignty in a meta-platform for data marketplaces.

- **Question 8:** What kind of control over data do you want when sharing your business data via a meta-platform for data marketplaces?
 - *[probe for]* Could you describe some examples or scenarios?
- **Question 9:** What are the primary causes of losing control over data in B2B data sharing via a meta-platform for data marketplaces?
- **Question 10:** What are the consequences of losing control over data in B2B data sharing via a meta-platform for data marketplaces?

Appendix 3: Unconstrained categorization matrix

Exploring the antecedents

Table 2 The antecedents of data sovereignty concerns in a meta-platform for data marketplaces

Category	Sub-category	Codes
Data governance mode	Untrustworthy data marketplace participants	Bad intermediary / Mafia involvement / Malicious access / Outdated security / Unknown operation
	Opportunistic behaviors of meta-platforms	Imbalance privilege / Transparency issue
	Unclear division of roles and responsibilities	Unclear responsibility definition / Uninformed actors / Unreliable meta-platform
Data ownership and access	Data policy divergence	Different work rules / GDPR applicability / Jurisdiction applicability / Legal measure variation / Liability differences / Non-EU regulation
	Data type variations	Aggregated data / Constantly updated data / Continuous data stream / Data latency requirement / Data usage variation / Derived data / One-off dataset / Personal data / Repetitive use
	Vague data ownership definitions	Complementary ownership / Data ownership complexity / End-user consent / End-user ownership / Inaccurate meta-data / Legal framework inexistence / Unclear data usage definition / Uncertain transferred ownership
	Unclear use cases	Ambiguous data license / Data usage agreement complexity / Unclear data sale purpose / Unclear data purchase purpose
Data usage	Data provenance difficulties	Blindspot / Complicated data access models / Data usage invisibility / Data tracking limitation / Data storage challenge / Lineage chain limitation / Two entities / Technical barrier / Unidentifiable dataset / Uninformed end-users
	Data withdrawal issues	Personal data removal / Processed data / Machine learning model
	Unverified buyers	Less trustworthy buyer / Unethical organization / Unknown buyers / Unknown criminal record / Unserious buyer

Table 3 The connection between code and excerpts (the antecedents of data sovereignty concerns)

Sub-category	Code	Excerpt	ID
<u>Category: Data governance mode</u>			
Untrustworthy data marketplace participants	Bad intermediary	“So if I want to evaluate good data markets, to upload my data to discuss with them....”	I-01
		“But I’m mainly interested in finding a good or a set of good data marketplace to upload my product....”	I-04
	Mafia involvement	“...you're able to distinguish between them [data marketplaces], or bad intermediary.”	I-04
		“If a channel [a data marketplace], for instance, is ruled by mafia, you will try to avoid it....”	I-12
	Malicious access	“...so one is the security of malicious use or access. So, are you [marketplace operators] able to prevent people accessing the data?”	I-12
	Outdated security	“...but and then they have don't have their own security up to par.”	I-01
	Unknown operation	“...because there are no establishment marketplaces where you can know the terms and conditions of running this marketplace and transacting with it.”	I-01

Sub-category	Code	Excerpt	ID
		"... I think that the metadata marketplace will not only have to create different APIs to each marketplace, but it has to understand the different rules of operation...."	I-01
Opportunistic behaviors of meta-platforms	Imbalance privilege	"...that the [meta-] platform doesn't have that privilege of seeing what the data is...or selling data that belongs to one marketplace to another marketplace or things like that."	I-05
	Transparency issue	"I'm not sure about a middleman [a meta-platform] already. I think you when it comes to data, it should be as transparent as possible?"	I-03
	Unreliable meta-platform	"Like your meta-data place, how reliable is it?"	I-03
Unclear division of roles and responsibilities	Unclear responsibility definition	"It's definitely tough...who is responsible for providing a certainty about the data, right? So who is responsible for providing the data limits? Who's responsible for providing proof that the data is secure, that you know the data is of quality, that it's that you know that the ownership is correct?" "It will be much heavier: How do you proof and how do you secure it and how do know where it responsibility lay? That's the most important thing. Who is responsible for making sure that it's doing OK?" "So for example, if you're a meta platform and the data marketplace gets it from you and then sells it to our customer and then that data marketplace has security issues or goes down, or the data is, let's say uh corrupted or something like that, then that's also a part of like the security issue in the chain, and then the question, of course, is again who's responsible for that? Is it the data marketplace itself? Is it the meta-marketplace?"	I-12
		"Where can I go with regards to questions?"	I-03
	Uninformed actor	"But there's also a risk of loss. And the less clear it is where people can go on. Unless so, there you know, remediating the risk or taking the responsibility for the risk."	I-12
<u>Data ownership and access</u>			
Data policy divergence	Different work rule	"And also, in several parts of the globe, work rules are different."	I-01
	GDPR applicability	"In some GDPR applies, in some other data other policy applies, I don't know."	I-01
	Jurisdiction applicability	"...I think we're talking about jurisdiction applicable to data, so sometimes even hard for us to know when does the payment fall within the [country] jurisdiction. For instance, you talk about cross-border payments."	I-13
	Legal measure variation	"And it's a really big blocking thing for this because if you want to do it Europe, you still have those translations between all the different legal stuff between countries... So how are you going to handle them?... and it's really difficult to get in place."	I-08
		"...if it's in European platform, there is still a lot of difference in the European country on how they handle data."	I-08
		"...but that's difficult then to even translate that to a third party and trust them to understand what they can and cannot do something with the data."	I-13
	Liability difference	"Whenever we have to negotiate with a [country] company, it can take months to get the contract signed because there are all the GDPR clauses and other liability things differences between	I-02

Sub-category	Code	Excerpt	ID
		[countries], which makes it difficult to sign these kinds of things.”	
	Non-EU regulation	“So for us, it would be OK if we sell this information within the European Union, for example, but it already becomes more difficult when we start selling it, probably outside of the European Union.”	I-02
		“OK, but I think the sharing data inside of the European Union will be as free as possible, but sharing data outside of the European Union will be very complicated.”	I-09
Data type variations	Aggregated data	“If we are sharing only metadata, aggregated data, then this concept is OK.”	I-09
	Constantly updated data	“And this becomes more difficult when data are constantly updated, for instance, every day, every week, every month....”	I-01
	Continuous data stream	“You could also be buying a data stream, so it's not a well one-off, so you get a constant flow of data.”	I-03
		“Yeah, maybe an example that it's also, let's say, bit familiar in regular more retail type of platforms is that if you have, let's say as a platform owner, you have multiple, let's say outlets, websites and well, we all know let's say the network effects, but if you have insights in the data streams between, let's say, a data buyer...”	I-15
	Data latency requirement	“It could be that one of the requirements of your data is latency, so you want the data in a specific period of time”	I-03
	Data usage variation	“I think that the examples from the hotel or the driving are not exactly equivalent...But these are very small, I would say variations and options.”	I-01
	Derived data	“Another thing I'm always thinking about with these data sharing or B2B data is derived data. So the derived data. Yeah, An analysis on top of the data which adds new input.”	I-13
	One-off dataset	“At the end of the day. In many of the cases, it's not just a one-off buy, but it's a repetitive buy or a repetitive connection.”	I-01
		“Uh, if you want to buy a one-off data set, maybe you want?”	I-03
	Personal data	“If we are sharing personal data, then this concept has some problems.”	I-09
Vague data ownership definitions	Repetitive use	“So the incentive that goes beyond the one-time use of the data.”	I-05
	Complementary ownership	“Or even more. If there's more parties involved in the payments, like in another bank, whose data is it? And who can share it?”	I-13
	Data ownership complexity	“It's clear, but it raises still a lot of questions. So, but it's not really not really direct questions, but I really considering, and then I go back to the role of the authority competition. Who's the owner of the data? The market power and then it becomes complex because for example, in the chain that you showed for the hotel and consumer. In a certain time it was booking getting too big. Uh, you really have to look into what all those competition authorities also already doing on their own data.”	I-16
	End-user consent	“If you have a meta-marketplace and you are handling customer data, then every time you put a contract in your meta-data marketplace, you will have a customer saying, yeah, I agree, you can use my data in that meta-marketplace. And that means if you want to have a nice commercial offering and say we can give you a prediction market of our customers. We as a bank have to ask all of our customers.”	I-08

Sub-category	Code	Excerpt	ID
		<p>“So I have to be informed if you want to pass this thing to sell personal data on data market platform.”</p> <p>“Yes, from my perspective, it's OK. If we are sharing personal data, then this concept has some problems. So if we are sending, so if you're sharing personal data and then I have to inform clients or customers....”</p>	I-09
	End-user ownership	“I always doubt that we have any data at all. Because we are more or less maintaining the data of our customers. We have dare, uh, details we have dare usage with their activities, uh, and uh, you can ask consent, but that's always then a grey area.”	I-03
		“I personally believe that will be very difficult when it comes to our customer data. Yeah, which is still relate to...Look, the data that is really our customer data.”	I-13
	Inaccurate meta-data	“Short description, some metadata like longer description varies explained. What does it mean? Some data ownership and so on?”	I-07
	Legal framework inexistence	“Even there, you're not always sure it's your data to share, because a lot of one-off contract now state something about data ownership, which is weird because I'm not sure if you looked at the law, but data ownership is sort of in the in the in the, if you look in the whole IP landscape right, so what IP rights you can claim to data, there is not a lot there, it's only a database right.”	I-13
	Unclear data usage definition	“So for me also when I will be in the role of the buyer of the data I need to understand exactly what is in the data and how can I use it in to models?”	I-07
	Uncertain transferred ownership	“Will it be still be owned by the provider or does he sell it to the platform and then it's not this data anymore?”	I-10
		“What happens to the ownership of the data? Is the ownership transfer to the platform or is it kept? Is it still owned by the data provider?”	I-10
Unclear use cases	Ambiguous data license	“Is it still part of the license?...Or it doesn't become the independent property of that party?”	I-13
	Data usage agreement complexity	<p>“But I wouldn't put them there as a mediator because I don't think you can set up the agreements up to now that easily yet.”</p> <p>“...has an agreement with the data market. The market has an agreement with the meta-data market and in the meta-data market needs to do it with the information provider. The more parties you include, the more difficult it can get.”</p>	I-03
	Unclear data sale purpose	“Yeah, and especially if you make money out of it. So if I'm selling these details and I don't know why I'm selling them.”	I-03
	Unclear data purchase purpose	<p>“No, for me, the purchase of data starts with the purpose.”</p> <p>“But I think it should always start with the question what's the purpose of the other side?”</p>	I-03
Data usage			
Data provenance difficulties	Blindspot	“Umm, to be, let's say to be able to be transparent towards our customers, I want to show the full pipeline. For example how their data are used. And for example if I will have this, let's say a blind spot that I shared this data to some metadata platform.”	I-07
	Complicated data access models	“Okay, that's it then comes the more complicated business models like, Okay, this guy only have access to up to three times to the data set. And this is already the fourth time, I should deny the access and so on.”	I-16

Sub-category	Code	Excerpt	ID
	Data usage invisibility	"But me, as [the name of the company], I have no clue what the other side will do of their data. They just bought it, it's theirs."	I-03
		"Do I still know why the customer is using that data? What they are doing with the data, and when they are doing something with that data?"	I-08
		"For me, at least, for me that there is at least a degree of confidence or trust I want to do by having some insights in how data travels."	I-10
		"I mean, this is very simple data sovereignty because the provider keeps the data all the time and then he doesn't know she doesn't know what happens when the data is transferred."	I-16
		"So many parties can be there without knowing where their data is going and what are they doing with this data."	I-19
	Data tracking limitation	"...And it cannot be tracked anymore."	I-17
	Data storage challenge	"...as a provider, that you know or have an idea, at least, where your data is residing..."	I-10
	Lineage chain limitation	"You don't even know a lot about the lineage chain, what's the data behind it?"	I-03
	Two entities	"Who is responsible for providing the lineage from supplier to buyer if you have two stops, which are two separate entities?... We have two parts in the chain."	I-12
	Technical barrier	"Well, if you build such a thing from scratch, it would probably be possible to think of ways to make sure that that you still keep control over the data that you are sharing, even if it's even after it has been shared. But I think this is the main challenge, and I would be the most difficult thing to implement."	I-02
		"Then you can't have a broker (a meta-platform) in between because it will be too slow."	I-03
		"In technical terms, it's very hard to do that"	I-17
	Unidentifiable dataset	"And there are many good practices that have been developed and that the data markets community doesn't know. For instance, a completely basic principle is to have persistent identifiers. Now, you need to be able to somehow identify datasets in order to be able to for them to be uniquely identifiable within in network"	I-18
	Uninformed end-user	"A data subject does not have a clue to whom the data will be shared."	I-09
Data withdrawal issues	Personal data removal	"It's easy to erase data, for example, in the bank, but if those data are ourselves. If, for example, to mining company in the EU, [a country], so I have to contact this company. Please erase my data. I must have a clue which company is that."	I-09
	Processed data	"I no longer want to do it, so I removed the datasets. Would that involve, would it mean that even that these kinds of rights, for instance, help you to revoke the rights in the in terms of the people who might have already used the data."	I-18
	Machine learning model	"A lot of the times datasets will be traded so that machine learning models can be built on them... (If data is withdrawn)...Are they required to retrain these models or not? I have no idea. I don't have a solution to this."	I-18
Unverified data consumers	Less trustworthy buyer	"Yes, then that could be a potential risk here. For example, if an insurance shares data that is also from my medical records or	I-17

Sub-category	Code	Excerpt	ID
		something, yeah, and then they share it with less trustworthy people. Yeah, it might become a problem, yeah.”	
	Unethical organization	“If [a government] would be interested in where the most crowded places are in [a country], then it is going to be a problem.”	I-02
		“I’m not willing to share, for example, my customer data with... or the organization that I am considering unethical so to have some let’s say also the certification of the buyers....”	I-07
	Unknown buyers	“Because you don’t have just mining company. You have hospitals. You have telco companies from all over the EU. We are talking about maybe 100,000 (buyer) controllers, for example.”	I-09
		“So many parties can be there without knowing where their data is going and what are they doing with this data.”	I-19
	Unknown criminal record	“Because you don’t know if these parties maybe in the past any criminal offenses.” “...check if this company that is demanding data has no criminal records.”	I-19
	Unserious buyer	“I will say from the seller perspective, quality of buyers. So you want to you want to make sure that you have serious buyers that will use the data....”	I-05

Exploring the consequences

Table 4 The consequences of data sovereignty issues in a meta-platform for data marketplaces

Category	Codes
Knowledge spillovers	Commercial secrets / Competition / Competitor advantage / Data disclosure / Data loss
Legal liabilities	Big player disadvantage / Dispute / IPR violation / Legal action / Legal consequence / Legal violation / Unclear legal liabilities
Data investment loss	Dataset loss / Data misuse / Data reselling / Data reuse / Dark web / Free dataset / Illegal market / Stolen datasets
Reputational damage	Bad press / Big company reputation / Credibility loss / Company closure / End-user concern / Reputational concern

Table 5 The connection between code and excerpts (the consequences of data sovereignty concerns)

Sub-category	Code	Excerpt	ID
Knowledge spillovers	Commercial secrets	“Big companies are the ones that are less reliance of this kind of data set of data markets I think because they are the ones that have more secrets.”	I-19
	Competition	“And for example if I will be if I will have this, let’s say a blind spot that I shared this data to some metadata platform, it will not be very transparent because out of it competitors and everybody considers the data....”	I-07
	Competitor advantage	“But don’t share the content because maybe other parties it could be used for competitors or for yeah other type of insights.”	I-15

Sub-category	Code	Excerpt	ID
		"My competitors, no, I don't wanna competitors to benefit from my data."	I-19
	Data disclosure	"...I will put your data on meta-platform and this meta-platform will sell will provide your data to an unlimited number of customers."	I-09
		"...but I don't trust a some mid corp somewhere playing around with our data set to know what is going on."	I-13
	Data loss	"Otherwise, without this protection, if you put the data set in the market, then you will lose it the very next day, and somebody else will benefit."	I-01
Legal liabilities	Big player disadvantage	"It's you know it's the bigger player in the market that's always going to then bear the brunt of it."	I-13
	Dispute	"In some cases, for reasons for the dispute, it would be good."	I-01
		"I don't care if for any dispute, of course, the data marketplace should assist in the dispute."	I-01
	IPR violation	"As long as I'm getting paid and as long as the IPR is not violated, I don't see my product being stolen and sold for the benefit of another person."	I-01
	Legal action	"They can be, you know, there can be legal actions."	I-18
	Legal consequence	"But when there is a personal and sensitive data element in it, then I think it becomes extremely risky for the organizations that own the data to do so, so that they don't expose themselves to other kind of legal consequences."	I-18
		"Also because at third party probably doesn't have the resources to actually pay damages or do something about it."	I-13
	Legal violation	"I think, it should be, uh, there's lots of work to ensure that this concept is GDPR compliance."	I-09
		"Yeah, so they had maybe a check on OK? Is this all? Are you not violating EU privacy rules or have I ticked all the boxes here?"	I-11
	Unclear legal liabilities	"And then I'm a lawyer. Someone who's liable? You know if something goes wrong? Where in the chain is responsibility."	I-13
Data investment loss	Dataset loss	"Otherwise, without this protection, if you put the data set in the market, then you will lose it the very next day, and somebody else will benefit."	I-01
		"But there's also a risk of loss."	I-12
		"The less control you have to the higher risk that it is that that your data will end up somewhere, first of all."	I-02
	Data misuse	"...the data has been, you know, misused or whatever."	I-05
		"Is there a contract as well that can call back that information or that can punish a buyer that has misused your data?"	
		"Because of data misuse potentially, yeah."	I-10
	Data reselling	"As long as I'm getting paid and as long as the IPR is not violated, I don't see my product being stolen and sold for the benefit of another person."	I-01
		"You don't want data to be sold like, it's yours, see what you do with it."	I-03
		"So let's say they're the adversary the threat model you're trying to address. It's just someone who's going to resell your data."	I-16
		"How do you ensure that they don't just resell it to someone else?"	I-19

Sub-category	Code	Excerpt	ID
	Data reuse	"How do you prevent the data from being reused?"	I-03
		"Let's say you uploaded data set and making sure that it's only used, basically no one copies it around and sends it to somewhere else and retains that it's their data set."	I-17
	Dark web	"Well, the thing is, of course, it depends on the type of data, but once you share something, if it ends up on the dark web and it is available free of charge for everyone."	I-02
		"We believe that if we exchange data with him, yeah, he's not gonna put it on the dark web."	I-17
	Free dataset	"But also, from a revenue perspective, if the data is available for everyone for free, we will not sell it anymore as well."	I-02
	Illegal market	"And you know that the black market out there is proliferating...."	I-01
	Stolen datasets	"Data easily stolen, data can be difficult, at least at this point watermark, or safeguards in a very prominent manner, you can get portions of data and sell it, you do not need to say the full set of data."	I-01
		"How can you know what part of your data has been stolen, so you need to have secure mechanisms to protect and identify the sovereignty of the data."	I-01
		"Okay, being sold everywhere, so everything is stolen. So we have to find ways to protect the sovereignty of the data."	I-01
		"If a channel for instance, is ruled by mafia, you will try to avoid it because you will be robbed or you will lose your credibility."	I-01
Reputational damage	Bad press	"So that's bad press, and I think, that's because we should never forget who generates the data from the beginning and how we generate the data from."	I-13
		"They actually got a lot of negative press about that up about location data. So they got a lot of bad press for it, right?"	I-13
	Big company reputation	"Yes, they really want to like the big companies are always hesitant on their reputation"	I-03
	Credibility loss	"If a channel for instance, is ruled by mafia, you will try to avoid it because you will be robbed or you will lose your credibility."	I-01
	Company closure	"Yep, so effectively a company that misuses data of citizens will be forced to close down"	I-18
		"They can be the company to close down and dramatically damaged the reputation of the company that does this."	I-18
	End-user concern	"That indeed [company] sharing the location information for the scope it is monitoring. The first thing that was on the radio, of course, was people discussing. The privacy and whether it was actually personal data or not, or whether you would be tracked now or this kind of things."	I-02
		"That's me or you making a payment online, right? So that's the data. I don't want to be impacted by that. I don't want other parties snooping around and knowing where I'm putting my grocery shopping. I don't know, right? And I don't want to be directly advertised, maybe for that. So what if and I think that's not allowed even, but what if it's third party starts doing that right? And I find out my bank is shared as shared my data. Uh, and this party wasn't allowed to do that, but they still did. Who am I going to look at, right?"	I-13

Sub-category	Code	Excerpt	ID
	Reputational concern	“There is also reputational damage. For example, there is a [name of a company] case from other we, I think it's 2014. This case was very was brought up in 2020. [This company] was spying on its customers. So this breach damage its reputation”	I-09
		“And from a reputational point of view, I'm going to look at my bank because I'm like, you know, you're the ones who should have guarded my data. I'm not going to look at the third party, you know.”	I-13
		There's always this and like I said, there's also think with the reputation there's always this thing that it's a one way street.	I-13
		“They can be the company to close down and dramatically damaged the reputation of the company that does this.”	I-18