Podcast

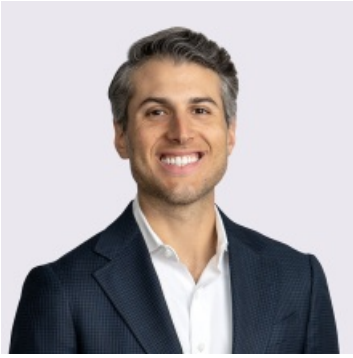
We Get Privacy for Work — Episode 3: The Increasing Importance of Data Mapping

By Joseph J. Lazzarotti & Damon W. Silver

May 22, 2025

Meet the Authors

Joseph J. Lazzarotti
Principal
908-795-5205
Joseph.Lazzarotti@jacksonlewis.com



Damon W. Silver
Principal
(212) 545-4063
Damon.Silver@jacksonlewis.com

Related Services

Privacy, Data and Cybersecurity

Details

May 22, 2025

Knowing what data your organization is collecting and from whom is critical to effectively respond to cybersecurity data breaches and prepare for future incidents.

YouTube

Listen on Apple Podcasts

Transcript

INTRO

To effectively and immediately respond to cybersecurity data breaches – and remain compliant with the constant bevy of new data privacy laws – you need to know what data your organization is collecting and from whom.

On this episode of We get Privacy for work, we discuss data mapping, the most efficient way to keep track of the information your organization is collecting and storing.

Today's hosts are Damon Silver and Joe Lazzarotti, co-leaders of the firm's Privacy, Data and Cybersecurity Group and principals, respectively, in the firm's New York City and Tampa offices.

Damon and Joe, the question on everyone's mind today is: What is data mapping, how do I implement it and how does that impact my organization?

Damon Silver

Principal, New York City

Welcome to the We get privacy podcast. I'm Damon Silver, and I'm joined by my co-host, Joe Lazzarotti. Joe and I co-lead the Privacy Data and Cybersecurity Group at Jackson Lewis. In that role, we and other members of our team regularly receive a variety of questions from clients, all of which boil down to the core question of how we handle our data safely.

In other words, how do we leverage all the great things that data can do for our organizations without running headfirst into a wall of legal risk? How can we manage that risk without unnecessarily hindering our business operations?

Joseph Lazzarotti

Principal, Tampa

On each episode, including this one, Damon and I are going to talk through a common question that we get from clients. We're going to walk through it in the same way that we talk about it with our clients, meaning that we're going to try to be as practical as we can. We're going to think about what the legal or other types of risks that we face are, what options may be available to manage those risks and what we should be mindful of when we go to execute on where we will land after that question.

Today, our question is, what is this thing called data mapping, and do we really need it? A lot of companies are wondering about that, particularly following the CCPA, the California Consumer Privacy Act. But it's probably something that really goes beyond that. At least it's something to think about beyond just the CCPA.

Damon, why don't you kick us off and define it, and we can go from there.

Silver

Data mapping is essentially a process for tracking what data your organization collects, who it relates to, where it came from, how you're using it, whether and to whom you're disclosing it and where and for how long you're maintaining it. It's basically your inventory of your data environment.

For many of the clients we work with, this is a concept that is relatively new. It has existed under the GDPR for a number of years in the EU. In certain sectors, like the healthcare space, it has existed under HIPAA for a while. As Joe mentioned, the CCPA and also the proliferation of data breaches and website tracking claims we've seen recently have put this issue of data minimization and data mapping into much more of a prominent place for organizations as they're trying to think through how best to manage their data privacy and security risk.

Joe, I'm interested in your thoughts here, but from my perspective, perhaps the most compelling reason for doing data mapping and focusing on the issue of data minimization is trying to better manage data breach risk. Sure, you do want to comply with the CCPA, and there are other reasons to do this, but in so many of the breaches we handle, and we are seeing definitely, at least anecdotally, an uptick in breaches in recent months, so much of the impacted data is data that the client we're working with has essentially lost track of and hadn't, for that reason, adequately secured it, or in many cases didn't really have a good reason to continue maintaining it. But nevertheless, it was there, sitting in that email account or on that legacy server from an acquisition 10 years ago. It was in the document management system, and no one touched it for seven years. Meanwhile, it's now within the scope of the breach.

Joe, I have an interest in your thoughts just on what the value of data mapping is? Why is it worthwhile to undertake that exercise?

Lazzarotti

You hit the nail on the head. Certainly, addressing data breach risk is a critical

function. It's like, if six people get emailed a document, and let's say that document is a census file. Who's on a leave of absence? Which patients' bills were paid? Who's enrolled in the retirement plan? Whatever the case may be, you get these documents that may have a ton of information that goes to a group of people, and they all save it. They all have it saved in their email account, and maybe they save it in whatever your document management system is, a shared file, whatever it is. Now, you have six documents in six different places with information on a thousand people. Those one thousand people may all be the same people, so you're not expanding the number of people necessarily, but you're making it easier for the bad guys to find it and to act on it. Just knowing that you're doing that can help to tighten the screws a little bit around managing data. That is a little bit of a crude example of how this can be actionable. That's one thing.

Also, a lot of times, we focus on personal information because that's what triggers notifications on the data breach laws. But the company may have very confidential business information or its clients' information. The same issue arises that by increasing the footprint of different data in the environment that you have, you're creating additional risk, maybe not even from a criminal threat actor outside the organization, but it could be from a rogue employee inside the organization. Understanding where that data is not only can help minimize the data, but you might also be better able to find out where it was. Because you know the map, you'll be able to speed up your investigation to figure out how this happened through a more deliberate understanding of your information system. Those are a couple of things.

The last thing I'll mention, and I think there are other reasons, but we're seeing the 23andMe bankruptcy filing came up. The California Attorney General put out an announcement, essentially saying that customers may want to delete that data. The notion of people understanding their rights with respect to data in certain states and certain countries, we're going to see that increase. If you don't know where this data is in your environment, somebody asks you to have to delete their data and you have a legal obligation to do that, or maybe a contractual obligation, but either way, going about doing that becomes difficult if you don't know where that person's data is. Then, responding to them and saying, yeah, we deleted your data. You might be like, did I really do it? Those are some of the things I think that can be helpful.

Silver

I totally agree with that, Joe. I would add that under the CCPA, there are about 20 states now that have a law similar to the CCPA and similar to the GDPR on their books, either in effect or taking effect soon. There is an obligation to provide pretty detailed privacy notices, and you're not really in a position to do that if you don't know what types of data you're collecting, where you got that data from, what you're doing with the data, who you're disclosing it to or how long you're keeping it.

On that last point, how long you're keeping it, there does seem to be an increased focus, for good reason, on record retention. There are a lot of organizations out

there that, just because it was the easiest thing to do, are keeping data indefinitely. They don't want to run into a scenario where they're looking for some email from a long time ago that happens to be relevant to a new project, and they can't find it. They think there might be some use to the data that they don't know about at the time, but it might come up later, and they want to be able to leverage the data for that purpose. In order to be able to have an effective record retention program and in order to develop these notices and to put yourself in the best position to mitigate your breach risk, you really do have to start by gaining an understanding of what's going on within your data environment.

Then, you also touched on one other point I want to highlight, which is that it's not just your internal systems you have to account for. It's also increasingly the various webs of vendors that you might be working with, providers of various software applications may have your data too. Both from the standpoint of responding to that deletion request, but also from the standpoint of managing your data breach risk, you want to identify that piece as well. That's one that comes up a lot as we work through data mapping exercises with clients, as we ask various stakeholders about what happens with the data, we learn that there are a bunch of different outside parties that have the data. Maybe the client has a data processing addendum in place with that client, maybe not. Maybe they vetted the security practices of that vendor, maybe not. That's the type of stuff you want to know about. You don't want to wait until all of a sudden the vendor has a breach and sends you a notice, only to learn that you had no DPA with the vendor. You have no means of holding them responsible. You're going to actually end up being the one who looks bad in the scenario because you're going to have to notify. Then, when the state AG is asking you whether you had a DPA, you're going to have to admit you didn't.

This is all stuff that is better to flesh out and get out ahead of. Usually, you can still fix it at that point in time. Once you're in the scenario of a government agency investigating your CCPA compliance or investigating your data breach, that's a bad time to first find out that there were all these areas of risk that you hadn't addressed as an organization because you didn't know they existed.

Lazzarotti

It's funny you say that, like maybe not a vendor, but there was a story. You may remember this also, and if I'm saying this wrong, please correct me. There was a breach that had happened, a company in the upper Midwest, maybe. Anyway, they acquired a company, brought them in, connected and set the systems up, and then nobody touched those systems, literally for 10 years. Then, this company has an incident, and threat actors find and exfiltrate that data. Evidently, there was no map; maybe there was a treasure map, and they couldn't find it. I don't know, but they didn't have a sense of what was on, or they didn't look at those systems for a long time. They wound up having to notify individuals whose data was on those systems that they acquired 10 years earlier. The enforcement was the attorney general, in two states, I want to say it's Pennsylvania and Ohio, wound up substantially fining them. The basis of that was that you had this data for 10 years, no clear use for it, and you just held onto it.

If someone had done a pretty basic mapping function, you probably would have found it and said, wow, what are we doing with this? It's not just systems. Sometimes companies have this hard copy file storage that no one even knows about. Oh, they forget. Somebody's just paying the $200 a month storage bill, and they don't know what's stored there. There are a bunch of boxes down there, nobody really knows what it is. That may not be as important because there's probably not criminal threat actors trying to hack into those boxes. The point is, it's knowing that you have it that can create unintended consequences that really can be pretty impactful.

The point I was trying to make is that even in transactions, not just the vendor point. Damon was absolutely on target, and perhaps more important, but even in transactions when you're exchanging information as part of the purchase of the sale, you want to know what you're giving up. If you're turning over equipment to a buyer, what's on that equipment? If you're acquiring stuff, you want to know what's on it because it could create a pretty big risk.

Silver

We did see, two years ago, that there were a number of financial institutions that got hit with big fines because they were disposing of equipment that had personal information stored on it, and they were not wiping it before they were disposing of it. That type of fact pattern could come up in a lot of different scenarios. It could be the transaction, disposing of the equipment or it could be the same desktop or laptop that someone in HR previously used has now been reprovisioned to someone who's outside of HR and shouldn't have access to the HR data.

As an organization, you may think, well, all of our data is stored within a couple of applications, and these applications are secure. But if you haven't data mapped, you may not be aware that people are also storing stuff locally on their laptops. They could even be storing it on their personal devices. You really do want to wrap your head around that because, to the point you raised earlier, Joe, you could, in a breach scenario, have a situation where someone lost the laptop. If you are able to say, based on your data map, that there was no PII in that laptop, great. Maybe that saves you a breach.

If there's a mass exfiltration of data, or if the encryption by the bad actor makes it so that you can't recover your data, if you have a map, you may be able to reverse engineer what was there, which may be helpful operationally. It will also be helpful in determining the scope of the breach. If you don't have that map, you're in the dark, and you just have to guess. Usually, if you're guessing, you're probably going to have to be pretty conservative in your assumptions, and that's not going to serve you well.

Joe, in wrapping this up, we've talked a lot about why data mapping is a worthwhile exercise and some of the functions it serves. Maybe you could talk a little bit just about the nuts and bolts of what this process looks like. If we're working with a client on a data mapping exercise, what are some of the key steps there?

Lazzarotti

Well, I was going to ask you that question, but I'll talk about it a little bit, and then I'm going to turn it on you. So, I'm going to leave the question. One question is, who does it? Who does the data map? Once you decide that, which Damon is going to talk about. Then, you figure out what the environment is. You have to know what are we looking to find and what do we need to look for? It's not just what you have.

If you're going to do this, you also want to think about data minimization. You want to say, well, what are we collecting? What is this? What's the life cycle of a piece of data or an item of data of record or whatever? What's the life cycle in our organization? Are we collecting it? Who has access to it? How long are we going to keep it? Where is it stored? Who do we give it to? Who do we disclose it to or share it with? How long do we retain it? What's the process of getting rid of it?

In a prior episode, Damon mentioned something that was important that can't be repeated enough. That is documenting this, being able to find it somewhere and substantiating that you know where this is. In that context, it was documenting privacy policies. Here, it's documenting that map so you really can figure it out, and be able to use that to your advantage, like in the example Damon gave about investigating a breach and knowing that the laptop doesn't have data.

Knowing the environment and then really thinking about that life cycle. Then, depending on the organization, going about each step in that life cycle to figure out this is what we collect. Also, why are we collecting it? Because you want to justify, if you're talking GDPR, you want to have a legitimate basis for that. As part of your map, you want to know why this data is here. Well, it's here for this reason. Okay. That makes sense. Right, and then you move on. But the point is really understanding how and why and where and what we're doing with data at each point in that process.

Once you do that, it will be helpful to build out that map and be in a stronger position. Damon, can you talk about, as well, who does that process?

Silver

It's a great question. I'm going to answer it by taking a quick step back, which is that the first step I usually raise when working through this process with clients is creating a list of the data assets, meaning the specific repositories of data that need to be mapped. Some of these are going to be pretty obvious: email environment, document management system, CRM. But it could also include individual workstations or laptops if people are storing information locally. You do want to be mindful of vendors. It could be web tracking technologies.

The reason I like to start with that inventory of data assets is that that's going to inform who should be involved. Because depending on the data asset, like if it's an HR IS system, you're definitely going to want some people in HR involved. You probably want someone in IT involved as well. If you're dealing with various types of web tracking technologies, you're going to want people from marketing involved. If you're dealing with a database that houses financial information,

you're going to want someone from the finance team involved. The reason for that is all of those people are going to have the appropriate institutional knowledge of what's going on within the environment. There likely is going to be a variety of people who are called on to help with doing the mapping of different assets.

Usually, the process is overseen by outside counsel working in conjunction with either in-house counsel or compliance or both. That team of inside and outside counsel and compliance is going to come up with the high-level strategy around what needs to be mapped, how to go about the process and what that should all look like. But there's definitely going to need to be involvement from various others in the organization in order for the mapping to be effective.

The next step after identifying the assets and figuring out which members of the organization should be involved with respect to the various assets is going through a pretty detailed set of questions, which we work with clients to develop to capture a lot of the information, Joe, that you were talking about earlier. What specific types of information are we collecting? What are we doing with the information? Are we, and this is an issue that's come up a lot recently, are we using AI tools to process that information? We have vendors involved in processing the information. How long are we keeping it? By going through that process, you will oftentimes learn a lot about what your organization is doing with data that you didn't know previously. That can be helpful in a number of ways.

At the end of the process, you will have this map that shows that you did diligence on this issue. You will be able to support what your privacy notices say, and support that you have a data minimization program in place. It could help you in the event of breaches. It could help you in the event you get requests from people to exercise their rights. But also, it's going to help you create your action plan to better manage risk on a going-forward basis. Because maybe you will learn that AI tools are being used to process data that you, meaning the compliance and legal team, didn't know about. What does that mean in terms of additional compliance requirements under the CCPA, the Colorado AI Act or the EU AI Act? You may learn that there are vendors involved in processing data that you didn't know about. You need to go to those vendors and make sure you have a DPA in place. You might learn that you're retaining data for way longer than makes sense. What steps should you take to remedy that? That, at a very high level, is how we'll typically work through the process. Identify the assets, figure out the right internal members of the client's team to help us sort through the questionnaires, complete the questionnaires, review the findings and then decide what the action items are coming away from those responses.

Lazzarotti

Well, that's terrific. This concludes this podcast episode. If anybody again has any questions or has a topic that you'd like us to address, please email us at privacy@JacksonLewis.com.

OUTRO

Thank you for joining us on We get work®. Please tune into our next program where we will continue to tell you not only what's legal, but what is effective. We get work® is available to stream and subscribe to on Apple Podcasts, Spotify and YouTube. For more information on today's topic, our presenters and other Jackson Lewis resources, visit jacksonlewis.com.

As a reminder, this material is provided for informational purposes only. It is not intended to constitute legal advice, nor does it create a client-lawyer relationship between Jackson Lewis and any recipient.