

## Navigating Issues in the Data Monetization Journey

January 21, 2021

### **Announcer**

Welcome to Mayer Brown's Tech Talks Podcast. Each podcast is designed to provide insights on legal issues relating to technology transactions and keep you up to date on the latest trends in data, digital, outsourcing and software, by drawing on the perspectives of practitioners who have executed technology transactions around the world. You can subscribe to this series on all major podcasting platforms. We hope you enjoy the program.

### **Julian Dibbell**

Hello listeners! Our topic today, data monetization. Data monetization in its broadest sense is the process of using data to increase revenue. In practice, that can mean a lot of things including using data to help a company optimize internal operations, improve existing products or services or even, in the most direct sense, create new data driven revenue streams. The bottom line though is that implementing a data monetization strategy is a journey and what we'll be talking about today is how to navigate through the various issues that can arise on that journey. My name is Julian Dibbell and I'm your host. I am a senior associate in Mayor Brown's Chicago office with a practice focusing on technology transactions. Our guests today are my colleagues Rohith George and Marina Aronchik, both partners in our Technology Transactions practice. Marina is out of our Chicago office and Rohith joins us from our San Francisco office. Thanks for joining us today Marina, Rohith.

### **Rohith George and Marina Aronchik**

Happy to be here. Happy to be here.

### **Julian Dibbell**

Glad to have you. Big data and data monetization generally are hot topics lately. Some of the hottest IPOs from 2020 are data companies like Snowflake and Palantir, Microsoft, IBM, Oracle, Salesforce and others continue to be heavyweights in data analytics. We're hearing every day about companies putting data to interesting uses that can help build revenue. What do we think is the driving force behind this emphasis on monetizing data? Rohith, let me start with you.

### **Rohith George**

Well, you know, companies have always gathered information and analyzed it to improve their business, whether that's, you know, a grocery store keeping track of which cereals sell fastest to try to manage their inventory or a fast food chain tracking which items on the menu are at the highest uptake the highest profitability so the idea of monetizing data is definitely not new.

I think there are a few things that have changed though. First, the amount of data that a company collects has increased exponentially. The grocery store is not only tracking which products are selling

fastest but it now has a rewards program that it uses to track which customers are buying which products and what the profile is of those customers, and, then because, you know, most commerce and much commerce is done online or over to your phone the grocery store may know what times of day you browse, it may know which products you look at but don't buy, what you search for on their app, and so on. An exponential increase on the amount of data that's being collected.

The second thing I think that's changed is that is the companies are increasingly interconnected and they have access to not only the data they collect but also the data that other partners collect. So sticking with my examples the fast food chain is not only tracking which items on their menu have the highest uptake but they may be doing online delivery through a delivery aggregator that's willing to provide them some additional data about which items are popular in the region, and at what times. The final thing is that the tools are changing as well. The two tools that are available to aggregate to analyze data they've improved by leaps and bounds where before customer service data was separate from sales data which was separate from manufacturing data, now you're able to aggregate that data and use sophisticated tools to analyze it to generate insights. It could be, for example, that customers that buy a particular product they keep coming back for customer service about a particular issue and there could be money being made there by expanding your product offerings to address that issue. So, I think those three things across industries. I ended up using some very basic examples but this is true across all industries, banking, insurance, pharma, healthcare, across all industries companies have access to more data and they have access to better tools to put that data to use. So, I think that's what's driving the kind of the focus, the hotness in the area.

### **Julian Dibbell**

All right. So, summing up, we've got more data, we've got better ways of analyzing the data and just the incentives, than to put all that data to use. so, what kind of uses are we talking about now when we talk about data monetization? You've given some examples but let's break it down a little more closely.

### **Rohith George**

I tend to categorize it generally into two buckets, internal use and external use. Internal use is analyzing available data in order in order to get a better understanding of your products, get a better understanding of your customers, ultimately to find ways to reduce your costs or increase revenues. Let's say you sell farming equipment. You could install IOT sensors in the equipment in order to continuously monitor its operation and then send that data back to you for analysis Then using that data, you can track when the equipment is starting to show signs of wear and tear, starting to show signs of requiring preventative maintenance, and then you can notify your customer that they need to do x y or z to improve the life of the equipment. That's you know a win-win for everyone. You the company avoid a costly warranty issue with a more minor fix and the customer avoids the hassle of a breakdown. That's an example of using data that you're collecting internally in new ways in order to improve your offerings to reduce your costs. External use, on the other hand, is taking data or insights that are generated from the data and sharing it externally for some economic benefit.

You could be directly licensing the data or insights to a third party, you could be partnering with another company, let's say it's a tech company, to combine your data with their platforms in order to create a new data driven business. Different companies have different data strategies. Some are focused at this stage entirely on leveraging data and proving how they use it for their internal use and others have discovered potential for new lines of business, for using the data externally.

**Marina Aronchik**

Julian, if I may, to add a little bit more color to this, what we typically see is that a company that's in early stages of digitalization tends to be more focused on internal use cases that Rohan was just describing. And, then, as that same company advances in a digital digitalization journey, these data related initiatives really tend to move to the external or revenue producing category. I think it's also important to mention that these external use cases typically do not involve the licensing of raw data even if it's cleanse data. We've seen that in some instances and in my experience that really related to cases where a company was unable to gather the raw data and then that same company might be partnering with another company to help it gather the data. But, again, that's pretty limited, that's not the common scenario. I think the more common scenario when we're talking about external use cases is really turning data inside into a product, or, even more likely, and the sense to drive more revenue, it's continuous access to insights through some sort of a platform and then that offering becomes a new product or a new service. Now these relatively complex collaborations are platform related deals where a company has a mass significant amount of data and that company is now partnering with a tongue company. That really seems to be the new frontier and I think an interesting aspect on that is that part of it is the result of technology companies moving into new spaces. So, for example, we might see what was previously a tech company start to dabble in what was previously a pharmaceutical space, they're pretty high barriers to entry and so, by that same tech company now partnering with a pharmaceutical company, together they able to create a new product or a new service that Rohith was describing, that really neither company could do on its own because the pharmaceutical company might lag the technological know-how and the technology company probably has tried and failed to enter a new regulated market.

**Julian Dibbell**

All right, so lots of potential kinds of uses and innovations in this space. Let's say you're a company that is sitting on a trove of data that has been collecting for a while but hasn't really taken the first steps on this journey towards really maximizing the value. It could be data about their customers, about their products. How can they go about turning that data into economic value in some of the ways that you're talking about?

**Marina Aronchik**

Well, of course as a tech transaction lawyer, when you mention a company collecting data whether it's about the use of products or its own customers, I want to start delving into the various legal issues that arise when you start to consider those scenarios. But let's first talk about the business perspective or the data scientist perspective of what the journey looks like. So in order to apply data analytics to data and to generate these valuable insights data first has to be made accessible and probably structured. So, described earlier, in many organizations data sets in separate silos so a manufacturing company would have finance and accounting information on one system it's maybe an ERP system and then invoices are stored somewhere else that might be in the ERP system, they might be in emails, they might be in a different independent system. And, then there's R and D data for a particular division that might be somewhere in the cloud let's say with Amazon. And then on top of it you've got manufacturing equipment with lots of sensors now because of course manufacturing is heavily dependent on sensors to increase reliability and those sensors are now generating high volumes of data on system performance. And then, probably, the M and A group at the company has separately licensed data from SNP and private company deals. So, you look at all of this data and some of it is structured and by that I mean that it's organized in tables in a systematic way, but a lot of that data is what data scientists would call

unstructured. So, if you think of emails, there's not a particular way in which you organize content in the emails. And that's true for things like invoices as well and when you look at all this data collectively it comes in all shapes and sizes and it's structured, semi-structured, whichever label you want to attach to it. So now what we're talking about is in order to analyze the data you need to be able to access all of these different silos and you do need a way to search or analyze that data.

How can you do that? About a decade ago this probably would have meant that you got to take all of this data and put in the same place and organize it in exactly the same way. The reason that we're talking about the types of initiatives and data monetization opportunities that exist today is because new technology allows you to do this through what's called a data lake where you can skip the step of organizing all this data in one place which, of course, is impossible because people keep writing emails right and generating invoices and you can really combine data from all these different data sources. You don't need to make it all structured in the same way but collectively you're able to create a consolidated, single source of truth and then make the data available for all sorts of data analytics and various users. Those users might be internal or later in the monetization journey they may be external.

**Julian Dibbell**

It sounds like there's a lot of prep work to even get to a state where you can meaningfully use the data you have. Do companies typically build these capabilities in-house or do they go to third parties to help them with this?

**Marina Aronchik**

It's almost universally third parties. Theoretically a company could develop it on its own but time is precious and typically companies don't have time and they don't have the resources and there's really no need to invest in the types of resources that it would take to build your own infrastructure for data lake and to build your own data analytics tool. So typically we see a lot of these collaboration deals or just services agreements with cloud providers without servicers managed service providers. Companies that provide data lake related services to put all of this in place.

**Rohith George**

I'd agree with that. Occasionally a company is already a data company and has been developing these tools in-house for a while. For example, a company that you don't usually think of as a data company as an insurance company. Their whole business is analyzing data and assessing risk and so some of them have developed sophisticated tools to aggregate and analyze new data. But, we're just right for the most part. We see companies going to the experts. They go to the service providers, the software providers, tool providers that do this as their core business, that's dedicated and they have the real brain trust necessary do this in the most sophisticated manner possible.

**Julian Dibbell**

Got it. Okay. So let's say you've built all the infrastructure necessary to put your data to use and we've been talking so far about all the interesting value that can now be generated out of that data but we're all lawyers here so let's talk a bit about risks. What are some of the key risks to watch out for in this area?

**Rohith George**

I'll start. So there are the usual risks around data security, data privacy cross-border transfer. Companies are used to dealing with those risks around data, they understand them well. However when you start

aggregating data analyzing it you start trying to generate value out of it in this new manner there's a need to evaluate how those activities might impact some of these traditional data related legal risks. For example, data security risk changes or expands and you start copying data from their primary databases to secondary databases or data lakes, the primary database may have been in your own secure data center that you had complete control over and the data lake may be in a third-party provided cloud that you have less control over. Generally speaking data being in more than one place effectively increases your threat surface so that's a different risk than what you had previously. It doesn't mean it's not it's insurmountable but it's just something to take into account.

Similarly, data privacy laws, cross-border data transfer laws, even export control laws all need to be looked at closely when using third-party cloud providers for data when they'll be kind of transferred to from you and an external third party of data. Those are all traditional risks but I think it's a matter assessing how they may change with the new activities that you're undertaking.

### **Marina Aronchik**

I'm really glad you brought up cloud solutions. I think those are critical to the data monetization journey. Earlier, Julian, we were talking about the various third party including cloud providers that a company will engage to help it create and manage the data lake and to apply data analytics and drive the data insights that in turn drive revenue. I think a key issue to consider when dealing with these relationships and these arrangements is the treatment of data, the rights to data and the ownership to data. A lot of times when this issue comes up there is a propensity to try to use IP laws to protect data and to rely on IP laws to preserve and allocate those rights. It's important to keep in mind that intellectual property laws for a number of different reasons really provide meager protections with respect to data. It's typically not protected by copyright at least in the US. Trade secret protections tend to be very narrow. Now there are some exceptions to that under US law but in general you really cannot rely on protection of data and data rights. Particularly ownership provides some derived data under IP therefore strong contractual rights are really critical to allocating rights and preserving rights to data. Recall earlier, we said the data is becoming critical and it's becoming a competitive advantage to many companies so in order to execute on that properly, it's important to preserve the rights to data as you're dealing with these various third parties. A lot of agreements with the with the third parties might expressly address data and data rights and those are a little more easier to address. I think it's also important to keep in mind that many agreements may contain what I'd call self-out licenses. For example, an agreement might say that a provider may use the company's data so the customer data to improve services or for other business purposes. Recall earlier we talked about tech companies trying to enter new lines of business. So a decade ago when you were talking about a provision like that everyone more or less understood that, yes, the data might be used to improve managed services. But today, we're seeing a lot of companies enter new non-traditional spaces and you're seeing industries sort of converge. I think Rohith has talked about that. So when you have something as broad as using data to improve services, does that mean the service is provided customer does mean launching new businesses that might compete with the same digital business that the customer is looking to enter into? That's important to address and surprisingly or perhaps not surprisingly, it's not always that once you raise the issue the clarification is immediately added because, in fact, it might be that a particular provider that you're dealing with is looking to use that data to launch a new business so that becomes an important business issue. It's also, these self-out licenses might also look like an agreement where a customer has to provide data that they have regarding x y and z and so that might be another way that the provider is getting valuable data out of a seemingly unrelated arrangement. Of course, if the contract is silent with respect to data on the rights, that's less than ideal as

well because possession is nine tenths of the law and once the provider gets access to the data there's nothing in the contract saying they cannot use data in a certain way. So it's important to think through all of these provisions now that the value and the use of data is really changing.

**Julian Dibbell**

What about the concept of open data which is somewhat related to these ownership concerns?

**Marina Aronchik**

I'm really glad you asked that question Julian. Open data is interesting because it's really the polar opposite of trying to protect your data. If you think back to software, it's the concept of proprietary software versus open source and here it's really a similar idea. We're seeing these ideas gain more momentum in agriculture industry and in the pharmaceutical industry. That has been in the news more recently, but I think it's important to note that even though there are these ideas of open data out there, the data and data thoughts that would be captured potentially by these open data initiatives are still very limited and then it becomes important to specifically address data sets that are subject to open data initiative but it would not mean that, for example, a pharmaceutical company is making all of its data subject to the open data initiative.

**Julian Dibbell**

So another concern that I've seen in some of the deals I've worked on is what are the sources from which you are obtaining your data? Rohith is that something you've seen as well?

**Rohith George**

Yes. I think that the issue is obviously companies obtain data, you know, I think we were talking about earlier, pursuant to many different contracts, from many different sources not just the data that it collects directly. There is the data that it gets from end customers themselves under terms of use or terms of service but there's also data from contracts with data aggregators or data resellers or licensors. There's the data they collect from supply agreements where the data kind of comes with the product that was bought. There's data from strategic alliance agreements or joint ventures and a host of other agreements. The issue is that data scientists are trained to find data and use it. They're not trained necessarily to find data licenses and comply with data licenses. So compliance is not easy to comply with the applicable license terms under which you obtain data. Each actual item of license data has to stay continuously linked to its source into the specific terms under which it was obtained and, you know, unfortunately, this is often not tracked at all or it's lost when data flows into a database or it flows from one database to another. The danger of course is that you wake up at some point to find that the data is in a general database being used in ways and for purposes that were not contemplated by the original license under which you obtain that data.

**Marina Aronchik**

I think these use cases raise some interesting issues with respect to interpretation of language that previously would have been viewed as non-controversial. For example, licensing data for internal use only. What does that really mean when a company is now creating a use case for data or perhaps insight from that data right so derived data is used in connection with a data modernization initiative? I think when we're thinking about external use cases and I should clarify, Rohith let me know if you disagree with this, but I think all of the concerns that you just described they really apply for both internal and what we're calling external use cases. Certainly when you're dealing with an internal use case the risk profile is a bit

lower and post maybe less concerned with tracking up all of the various licenses attached to data sources but the further along you move into the external use case the more of a risk that becomes because now you're making promises and commitments and risking incurring pretty significant liability in connection with breaches or failure to comply with these obligations under the third party licenses. I think when we're thinking about external use cases it's also important to consider a different or changing regulatory landscape. Are there any laws that might prohibit the intended data license? So, for example, there's some state laws that prohibit telecommunications carriers from reselling internet search activity and there might be other laws like that that might prohibit a company from entering this new business, at least without addressing those regulatory concerns. There's also an issue of does this external use case and a new product or a new service that a company is offering, does that make the company subject to new laws with which you didn't have to comply previously, such as the fair credit reporting act, for example, And as I mentioned earlier the further you move in the in the food shade of your youth use cases the further along you are with external use cases the more difficult it might make for you to comply with the contracts that govern original data sources. So, for example, you might have made a commitment to a third party to destroy its data under certain circumstances. You now need to slow down that obligation and probably related obligation to customers or partners to whom you're providing data otherwise you may not be able to comply with that underlying obligation. This really raises the question of can you manage a data business and I think it's important for companies to recognize that these use cases and initiatives so we're talking about really our new business and it requires new controls to be put in place or these new activities ranging from auditing to verifying compliance.

### **Julian Dibbell**

Okay. Well I asked you to describe some risks and you've definitely delivered a lot of risks potentially involved here but obviously there are ways to address these risks or nobody would be doing this, right, so what recommendations, Marina and Rohith, would you make for a lawyer working in this area and looking to protect their company from some of these risks?

### **Marina Aronchik**

Maybe I can get us started here. I think the first and critical component to this is making sure that lawyers are part of the data strategy team. A lot of the risks that we were just discussing here can be addressed, they may not be addressed fully but if you follow the 80/20 rule you're able to make a lot of progress with addressing all these risks or at least the more critical, the more material risks. The issue is that what lawyers might advise a company to do is probably very different and maybe the opposite of what the data strategy team might be inclined to do on its own. For example, data scientists might prefer to create one data lake that creates the single source of truth we were talking about earlier in order to make it easier to use data. Well for regulatory reasons, you might need to create a couple of different data lakes or manage the single data lake properly including the permission so that it's not anyone in the company that's able to access their data but its particular sets of users. I think that as companies move along the data monetization journey all of the companies will come to this realization that it's really important to manage the data sets and the idea that you can have a single source of truth with everyone able to access all of the data. It's probably not the ideal, probably not the ideal solution from the compliance and risk perspective but it's important to think through these cases ahead of time to the extent that you can predict them. And I think that lawyers can really help data scientists find the optimal solution to pursuing the business strategy while mitigating and accounting for risks given the reasonably foreseeable use cases. Another thing that I think lawyers can do is really update the form templates and agreements to contemplate data modernization initiatives. What we're seeing is that almost every deal that's done now

includes a data component and it comes in all different shapes and sizes but if you have those quote unquote standard provisions or your preferred positions and some acceptable fallbacks and those have been reviewed by the data security team, the data privacy team and other related specialists, and that the transactional lawyer who is working in a deal knows the range of acceptable compromises, my experience is that they can really help you address some of these scenarios that come up on those deals that are coming up today in the technology landscape. And finally, I think it's critical that companies begin to think about and to form teams that are really focused on data licensing. Collecting all of these documents alone is a very significant undertaking and of course in this day and age it's not just the company's own data, it's increasingly third-party data so really thinking through ways in which you can collect and manage the vast volume of agreements and terms that govern this licensed data is critical to the company's ability to then move quickly when data monetization release related initiatives and opportunities are coming up.

**Rohith George**

You know I think Marina covered a lot of the recommendations pretty comprehensively there. I'll just add a few things. As you're as you're negotiating new agreements or amending existing ones obviously you know that that gives a great opportunity to add the consents and the licenses required for not just existing uses but anticipated future uses as well. Also investigate possible problems in a legally privileged manner when they arise. Don't wait for data trolls or regulators to attack just because something's wrong. Just hiding it not looking at it doesn't make the problem actually go away so investigate those in the legally privileged manner as you can. And then finally treat data related contracts, data integration, data licenses, data analysis contracts with your data vendors, treat those contracts, those deals, treat them as high risk contracts, treat them as strategic contracts, dedicate additional legal resources dedicate the appropriate attention to it. I think if we do all that I think you're in a in a good place to attack this, to kind of be a successful and productive team member who's actually driving towards the monetization, towards the actual value that's derived from data.

**Julian Dibbell**

Well thank you Rohith and thank you Marina. Great insights from both of you on the promises and pitfalls of data monetization. Listeners if you have any questions about today's podcast or anything else related to technology transactions and the law please email us at [techtransactions@mayerbrown.com](mailto:techtransactions@mayerbrown.com).

**Announcer**

Thank you for listening. We hope you enjoyed this program. You can subscribe on all major podcasting platforms to learn about other Mayor Brown audio programming visit [mayerbrown.com/podcasts](http://mayerbrown.com/podcasts). Thanks for listening.