Salisbury Road

Speaker Key:

JR   JP Rangaswami

UM   Unidentified Male/s

UF   Unidentified Female/s

JR   I've been doing in adverse circumstances, where popularity of the firm or the role wasn't the [inaudible]. I've been a supporter of web science ever since I heard the term. I'm a trustee of the Web Science Trust almost since inception.

And what I wanted to do today was to give you a perspective on why. And what I've learned about Web Science while working in the commercial environment.

And I'm going to use my last role, the Chief Data Officer and Head of Innovation at Deutsche Bank, to frame it.

00:00:43

I want to speak for not longer than the next 25 minutes, so that we've got a clear half an hour for questions. I will frame something to stimulate the questions, and then throw whatever you feel like at me.

I'm retired, and I call myself a grandfather in training. I have two grandchildren already, and I'm spending a third of my time on boards, a third of my time doing pro bono work for which this counts, and then the remaining third doing things that would just make you jealous, so I won't share them.

Okay, Deutsche Bank. I was there for about three years, eight months. To get an idea of what the landscape there is, we had about 140,000 people working for us in over 13 countries at scale, but over 80 countries if you include just the small branches or individuals.

We were the world's largest in terms of foreign exchange, in terms of international trade, in terms of payments that were non US. In terms of the derivative position, we had worldwide... In fact, we were considered the riskiest because there was something like a 1.7 trillion book.

And we were also by liquidity Germany and credit card Germany, so national dominant shareholders in certain segments of the market. And globally number one in a number of other markets.

We also paid about 14 billion in finance relating, and I say this quite clearly, to activities from at least five years before I joined them if not ten.

00:02:43

A lot of the excesses were as margins eroded in the lead up to the financial crisis, and to me they were the consequences of some of the way the economies interacted with regulation and how people behaved under stress.

But when people say, no bankers were arrested and stuff like that, don't you worry. There have been enough. It's just that you don't say these four traders were done for this, because you're expecting to see Group CEOs go.

But the two banks I worked for had the relatively unusual distinction of having their CEOs assassinated.

So I've worked for two banks who's CEOs were actually shot, but long before I joined.

The size of Deutsche from a data perspective was maybe 38/39 petabytes of operational data with maybe 120 petabytes of data that would still be used because of people drilling in for disclosure purposes, because of regulatory questions asked, investigations because of market abuse accusations, anti-money laundering and anti-financial crime tests.

So when you look at that kind of scale, it held above 18,000 databases across four major centres, but about 15 data centres and held across 44 operating systems when I joined.

00:04:35

By the time I left, that number was something like 29, but the decay rate was fairly slow with a lot of end of life equipment, so access wasn't necessarily easy.

But the first sense I wanted to give you was a large corporate that is like a small country. You are able to learn things about web science and how you apply it. What kind of questions get asked, what kind of environment you are operating in.

The technical complexity had to be at least shown to say, this was some level of scale. This wasn't a mom and pop store. We were actually dealing with a large quantity of data.

I built a lab of at peak... I don't know what it is now because I left in June, but at peak it was 85 PhD and MSc as data scientist, all operating out of Dublin because I had to find somewhere as a catchment area that I could get, especially with the looming threat of Brexit at the time, being able to get a pan European environment.

And Deutsche Bank being headquartered in Germany also meant I had to make calls around it.

But the data lab was really operating out of Dublin. The structure for dealing with the data was almost instinctively three layered, so we had one data lake for the firm.

00:06:28

I had somehow managed to prevent proliferation of data ponds and swamps and whatever else you want to call them. All the data of the bank that I was interested in doing longitudinal studies on or trying to look for patterns I would not be able to otherwise get was held in one data lake.

The three tier model was to say that my team ran the lake, ran generic tools access for it, but then overlaid that with the people who wanted to get insights on cyber or insights on anti-money laundering, or insights on marketing. People would then create much more segment specific tools that ran at the layer three.

And part of my job was not just to improve the ingress tools and the ability to provide the right view of the data, but more importantly do that while not breaking any data protection laws that I was getting exposed to.

So that's the intro for what I was dealing with as data.

The firm spent upwards of a billion on tech and ops. So, the cost of running the tech infrastructure, building the services on top of it and operating all of it worldwide was in excess of a billion.

Again something that gives you the size and nearly half the manpower of the bank existed in that space.

00:08:18

Banks don't physically make anything, so you're taking data in, you're taking data out. And along the path providing a set of value added services, if you accept that they are value adding, which we can have a separate debate on.

I don't think that banks are actually in the business of money. They're in the business of trust, which is interesting because it's like we could say banks don't get trusted at all. But I haven't seen that many people choosing to move to Facebook or to Amazon, rather than the banker.

There can be arguments about whether the investments were in the right shape or form, whether the regulation was powerful enough or the people right, etc, but banking is not what is in doubt, because the ability for people to say... When I say a bank is about trust, I say that because even if you take the word bankrupt, the merchants of Lombard knew the words for banker. They got their name because they used to sit on benches called banko.

And if you didn't keep your word, then the other merchants came along and said, hey mate. You're toast. We're going to pick up your bench. We're going to break it in two. You can't sit with us anymore. Your banko was rupto. That was the rupturing of the bench which became bankruptcy.

So the failure was a failure of trust.

00:09:48

The earliest signs, you know, most of you are well aware of things like friend of a friend structures with an identity, the first FOAF that I saw in operation was actually between the London discount market and the London accepting house committee.

There used to be in the early London money markets going back 300 years, a basis that said, if I was a trader and I had a bill which was owed to me, I could go to a discount house and give them the bill and they would give me 92% of it, or 95% of it as an advance, charging a small fee for the process. Because they were saying, yes, we understand that that bill is good and it will get paid, and we're going to help you by giving you a cash advance before you get it.

That was fine if the person doing the owing of the bill was trustable, and that was true only if it was within the domain that the discount house knew about.

What happened if the bill was payable by someone in Hong Kong or Singapore or Madras or whatever?

So suddenly the London money market did not have the capabilities, but they had these trading houses that had opened branches in all these areas. Even like the HSBC groups are out of those trading houses, and the Standard Chartered group is from those trading houses. So the Williamson, the Omega, the McNeil, quite a lot of them were actually Scottish expats during the days of empire who built these trading posts.

00:11:33

And they said, well, you might not know them but we do. So a mechanism emerged that said that a signature on a bill by an accepting house which would levy a fee for that would then get paid by a discount house because they didn't need to know the payer. They only needed to know the accepting house.

So this was practical use of friend of a friend models 300 years ago.

The reason I bring that up is these things about identity and traceability of data became really important when I took over the bank because the best way I could describe the concept... And I say this not just of Deutsche but anywhere I go, is people were used to data like mother's home cooking.

The data used to come out of systems that were in the basement of the building you were in. The programmes were written by people who lived and worked there. The data that went in there was sourced by people doing their jobs within the building, and so you knew the recipes. You knew the kitchen, you knew the equipment being used and therefore you had high trust in the food that came out.

By the time we had things like citizen publishing rights in an enterprise like Excel and PowerPoint, I used to get shocked at the number of people who thought something was true because they had seen it on PowerPoint.

00:13:15

And who assumed something was accurate because they'd seen it on a spreadsheet.

I have absolutely no idea of problems or relatabilty or even the skill level of who was collecting the data. But it had become official because they were used to a culture where the only things they saw were reports that came out of completely proprietary, within local control systems. The ability to question the data or to have anything that looked like data literacy and numeracy was completely missing. Tough.

Now in a networked world, where firewalls are porous, you have tunnelling all over the place. You buy certain things from third parties and you buy even data from third parties. You have mobile devices, smart devices involved. So the collection points are varied, the collection parts are varied, and the skill level of people doing the collecting is extremely heterogeneous.

People trusting the data that gets aggregated together is to me like saying, I'm going to have street food blindfolded, hands tied behind my back with no knowledge of where it came from. I'm going to use trust levels that were used for my mother's home cooking, and I'm going to ask why I get ill?

00:14:40

Well, teaching people about provenance, teaching people about the immutability, getting an understanding of why data identifiers become important starts getting really difficult to get the cultural transformation.

So as I was starting to get asked questions about data in the bank, I had to actually make more than anything else cultural moves in order to aggregate.

So a regulator would come to me and say, I want all the information to do with your dealings with Client X. Client X could be a Fortune 50 company, and I would need nearly 200 permissions, 198 permissions to aggregate that data because of Federation of Data Protection Rights into companies, sub-companies, departments, and individuals within those departments. And all of them would have a very strong inertia to say, it doesn't matter what the question, the answer is no.

So as much in an enterprise as in anything else, you land up with spending 80-85% of your time while you are apparently in charge of the data trying to determine how to get to that data, and finding that very few people know anything about the data you've just got to. How fresh is this, because time and date stamping aren't disciplines that are normal? Or, use of common identifiers.

You would land up with saying, again... The data points are here to help you, as Les was saying. If you have a perspective that even in large enterprises you learn things about websites that you can then apply outside.

00:16:44

And I hope that will trigger some responses in you that will help your own research. But there were things like over 30 different client identifier databases.

So then you say tomato, I say tomato, kind of problem, where an average bank has something like 800 critical systems. And these things tend to have grown topsy. So some organically grown, some without the space to have full meaningful client identifiers because they were written in the late 70s and you didn't have the actual memory to be able to do that. So they were an N digit number as the only way of identifying because you could compress that reasonably, whereas others had completely free form text.

The choice of those identifiers, even if you just dealt with the retail... And Deutsche had 22/23 million customers, so again not trivial when you're trying to look for how they are named.

When you get to the commercial where we only had about 160,000 customers, it got even messier, because 160,000 customers could have four million accounts. Because you could have the group as a customer which may have 300 subsidiaries, and the subsidiaries would be reshaped within different jurisdictions, sometimes because of the tax structures that people would want to go to. This would not be 100%, so you would have to aggregate very differently based on is this 100% owned subsidiary? Is this a subsidiary that has a majority control and you have 51%?

00:18:46

And you had to take off that kind of structure somewhere between 25 and 30%. When you're trying to associate the risk, an associated company of even 10% or 15% holding may be important for you to hold.

So you landed up having to hold hierarchy information about the clients. So not only did you have proliferation in the naming structures, and that was constrained by systems and architecture as much as anything else. But then you made it worse because you had geographical dispersion, because people wanted to optimise for jurisdiction and taxation.

And then between the systems constraint and the financial constraints, you went further and turned around and said that the reason why you would want to roll up the aggregation was for different purposes. And that meant that you also needed to have tools to be able to say, I want to pick. I'm doing X for risk management, I'm doing Y for financial reporting. I'm doing Z for legal reporting.

So the same companies would drop in and out of your catchment slips. You couldn't have predictability in any kind of good form for that.

A lot of my work was being spent just creating for want of a better phrase data identifying services.

00:20:21

Client instrument because the complexity was mainly in the trading functions. So things like Thompson Reuters or Bloomberg naming conventions for the bonds, the equities. The symbology that was used, and making sure those variations on that symbology... Because again if you want to buy a share, you will find there are 16 ways that that share could get named because some of them are bonds, some of them are ADR type issues in a different market. The same stock is traded in multiple markets and not quite the same form. The currency, the databases, the interest regime under which it operates.

There are bank holidays that relate to it. It is a mess. So you needed to have services that standardised on the labelling. Client and instrument are just two. You also needed to have some equivalent of it by turning around and saying, I also need some product and pricing information for how to touch market data and connect it onto the instruments validly.

All these required considerable investment just in metadata management.

And this is with people who aren't data scientists, this is with people who wouldn't know problems or lineage if you hit them in the face. And I tried to explain, look... A trade gets copied in a firm like Deutsche 80 times.

00:21:53

So these people... The culture of the place is to live and breathe ETL, and there are probably a hundred things I would rather do in data than to allow ETL to proliferate. But it's the way these architectures emerged.

The reason I'm sharing all this with you is the kind of questions I would ask saying, show me all the communications done by person X during the period A to B? Because we need that answer in order to have disclosure for a law suit.

Finding out that a person can have more than one mail account, because they were powerful enough to. Understanding that employee numbers got recycled, because people didn't realise that that was not a good idea.

Creating cultures where the end product of a transaction cycle was editable, with plugs that went in over the top, because you had time constraints for when you would report. So you would plug it and then go back to it later... That you had an ETL that was decayed.

I had people making adjustments and they were making the same adjustments every day. I said, why do you keep making the same adjustments every day? They said, this type of trade doesn't come down to me. It's missing so I have to put it in. I have to do it this way. Why does it not come down to you? I don't know.

We were moving somewhere in the order of 2.25 billion rows of data every month through ETL.

00:23:44

And that meant that even getting some understanding of where the data screwed up... Someone could have said, oh for this particular set of extract transforming load process, please take the following six transaction types. But the person who wrote that would have no knowledge of the fact that someone had added a seventh transaction type which was forever missing. And then you created a small industry.

And when I say a small industry, we had about 5,000 people whose sole job was to adjust it.

So when you start seeing the scale of the data, the complexity of the architecture, the lack of skill to do with the data, the complete lack of any understanding of data science etc, then you land up with questions saying, I want to know. And most of the I want to know were perfectly reasonable.

It's just that the context in which they were being asked, no one had an idea of why it took so long to get an answer. Then you throw in the corrective cyber in a network world where the problems of money laundering kind of issues, even today... Everyone turns around and says, banks must be really crap. The amount of financial crime that goes on that no one can catch.

00:25:04

But basically if you imagine a system that contains all the banks in the world, and no willingness within the banks to collaborate, all that a bad actor has to do is to learn to operate within more than one sub network of banks because no one can see the whole picture.

The malefactor has the advantage and you're trying to say, looking at a small piece of the puzzle, I'm somehow going to find ways of saying, this person is up to something.

The problem is that you can only share the information after it has gone bad. It's like an Interpol notice, or they're on a watch list.

The way any financial crime works is you do three things. You do sanctions which is a government or a state factor has said, no dealing with this. You can't deal with Iran or you can't deal with Syria or whatever, and now you've got to unpack to make sure someone is not pretending to be somewhere else in terms of how the legal structures have been noted.

The second is nameless screening. When you actually say, this is a country he represents and he's in the same bin. You're going to stop. So you're looking for a match.

When you have 40 databases with client names and then names may not be names. They could be numbers. You don't know that necessarily because you've got to validate it.

00:26:40

And validating it with street ID etc, if you've ever seen what messes you have to go through opening an account etc... You guys mean well, right?

When you actually deal with someone whose sole job is to try and pretend to be someone else, it becomes a whole different ballgame.

And then the third is politically exposed persons. So now you're saying, this person is a friend of someone who happens to be a politician, and you have lists of those.

So sanctions could be nameless screening and politically exposed persons being the way you deal with that.

Each of these is fraught with traditional challenges to do with data. So when I tell them, you know, sorry 80-85% of my time is wrangling the data, I'm not spending anywhere near as much as I would like to on the actual models of capturing the insights, on learning from the insights.

I felt in coming and talking to you guys, if you understood that the issues to do with infrastructural complexity, the problems of the lack of scalability, the population of the world that's generating the data, the lack of information actually to do with how the data was captured and under what conditions. The weaknesses in any metadata matching... The problems then, because we have come now a long way quite correctly, about data protection and rights.

00:28:12

But how we've implemented it creates absolute nightmares. I remember even 25 years ago, I was screaming at regulators, saying, in America is was required by law to scan every email looking for P words.

Law... I went to jail if I didn't do it.

In the UK, I could only scan mails if I had cause.

So I needed to have evidence that I could stand up in a court of law which normally I would just go to the compliance officers etc, but it had to be filed to say, like a mini warrant. I had reason to believe because this had happened, that is why I went into someone's purse.

In Germany, it didn't matter what the person did, I could not get in.

Now what would happen if I had a German national working for a UK company, visiting the US for a fortnight?

I had to break something. And they don't even understand that the internet is a copy machine. They would turn around and when you say, do you know what actually happens for someone to access an email while in another country?

And they don't understand even the basics?

00:29:24

So they would say, no just ring fence them completely. How do I prevent a copy of their mail at least coming onto that monitor?

No, just don't allow them to see it. Well, what's the point? You don't want me to allow them to even log into the system? No they have to. Then how am I going to prevent what you're asking me to prevent?

And like the people who come up with back doors etc, there's a lot of tension about people who set up regulation not even understanding the basics.

So the summary I would give you is my interest in web science where I believed that the kind of insights I could get from data grew exponentially when the data came from a networked world that the metric actually allowed me to light up relationships that I could not otherwise.

So you could turn around and say, if everybody switched off location services and timed services on their mobile phones, quite a lot of the digital giants would die because without those, they don't have the joining capacity for a lot of things. We don't need law as much as saying, everybody just switches those off forthright from all services.

00:30:48

But the new insights are possible because we're allowed... Data is just a number. When you associate time and place, it starts getting some value. When you associate identity and moving to a real graph, who that person or thing knows, you gain a little more. Then you throw in the labelling, the part number, the skill, the pharma code, the stock code, it doesn't matter what.

The ability to generate insights comes because you lit up what was just a number in some place or the other. A whole lot of contextual information will then really make it visible when you start joining.

It may sound simple but most of my life in data has been spent trying to explain to people that we are gaining exponentially the amount of data that's available, but not necessarily gaining exponentially the amount of value and information that's coming at that time.

People don't have even a simple understanding like the distinction between correlation and causation. And they don't understand the importance of immutability. They don't understand the importance of lineage. They don't see why the transforms they carry through are destructive. The lack of consistency in labelling makes it very difficult.

And with so little knowledge of the conditions in which the data was captured, and no one to talk to about those, very little time series or longitudinal data, because conditions keep changing. So you have to almost synthesise any attempt to work it back.

00:32:33

So when I was hearing about web science and I understood that the constraints were much more to understand the social, political, legal, cultural use of that data as much as the engineering aspects of it, I felt very attracted to get involved because I felt that these things were fractal. The things I was experiencing in industry and in enterprise were things that were as meaningful.

That in corporate open data it mattered, there were things that we could only do in an industry if we agreed to have the same labelling structures.

Right now one of the companies I advise provides time as a service. Traceable time, because people don't even understand that most data centres have system clocks all over the place, even in a single data centre, and they use open source NTP second granularity stuff. And I'm dealing in markets that are doing 20,000 price changes in a second.

So I need microsecond if not nanosecond to be able to do some of my big checks.

I'll stop there. I've taken eight minutes more than I wanted, but I started a few minutes late. But I hope it's given you a feel for why there are things that can be learned about policy, about structure, about approach to being able to form questions and then to answer questions.

00:34:08

That the barriers you find in your research are barriers that people and enterprise will also find. And then when you scale up and say, it's multicultural, it's multi time zone, it's multilingual, each one of these adds their own complexity to it.

So with that, throw whatever questions you like. I hope you found this useful. The intention is still to say, the problems of this world cannot be solved... Today's problems tend to be multi-time zone, multicultural, multilingual... Whether that's climate change or nutrition or water or energy.

We can't solve them within a single well bounded location. Our ability to collaborate is weak until we start getting the right disciplines about joining data together. Those disciplines are poor.

And web science represents the best shot we could go at being able to create the disciplines and come out with rational, unemotional proofs for the kind of things we want to learn which may be counter to our selection and confirmation bias that we operate with.

Okay. Yours. Yes.

UM   Fascinating talk, thank you. I've got about ten questions, but I'll just stick to one for the moment.

JR   Sure.

00:35:28

UM   You mentioned about the one billion was it dollars of investment on tech and ops?

JR   Yes?

UM   So when you're talking to the board...

JR   It was Euro.

UM   Oh, Euro. A significant amount of money, how do you articulate the value or the return on investment to the board in order that you can justify that one billion of investment on tech and ops? What are the value metrics? What are the benefits metrics, and how do you measure that in terms of investment?

JR   Well, you land up with effectively a three tier model of one set of things that... If you want to look at the tiering, it's always worth reading a guy called Paul Strassmann, who wrote a lot of these Harry Potter sized tomes called Business Value Computing in the late 80s. He used to be the head of the Department of Defence, which was the world's largest spender on technology.

        And the elevator pitch answer was he could not see any business value in computing.

00:36:42

        However by the time it got to Nicholas Carr who was the next guy to get involved in that debate, and I know Nicholas reasonably well... At the turn of the last century, he became famous for writing a Harvard Business School article called, Does IT matter? And then a book that came out saying, IT doesn't matter.

        And he came up with a theory that said, you know, the smartest thing to do in tech in industry is to be a fast follower. And the first time I met him, I said, yes great Nick. Follow whom?

        And then people who come up with follower mindsets don't understand there has to be an incentive for someone to be the leader. And if you take that incentive out, the whole market collapses.

        Going back to your question, the way I try to run it is that this idea of discounted cash flows or ROIs etc is completely wrong if it's applied homogeneously across everything. It's not a cookie cutter.

        So there are some things that we would do that are no regrets moves, because they are part of the cost of doing business. You only have to find the best way of doing it.

        It's like saying that if I have a regulation and I have to conform to that regulation, and the smallest iota of that would be saying, VAT used to be 17.5%. Someone decreed it's going to be 20%.

You have to make some changes, and those changes have a cost.

When I was at BT, I found the cost was higher than I anticipated, because someone had very cleverly designed that the VAT rate was captured as pixels on a screen in design and not held as a parameter in a file.

Now, you know, great. That was quite expensive because everywhere they put this, it was actually an image and you had to edit the image across thousands of systems. And so I think, well we'd like to do that once, but we're not going to do that. We're going to change it all so that it's actually driven out of an engine.

But I'm giving the example to say, the no regrets move is where there is a regulatory or market move whereby the choice of doing something or not doing something isn't there anymore. All you are trying to do is find the optimal way of doing it, in which case a business case validating build versus buy versus open source, turning around and saying... The payback is not something that you have to calculate in quite the same way because you've removed choice. You're forced to do it.

Or B) of the no regrets moves is you're still doing it but the payoff is a revenue payoff because you now have a new product being launched in a market that you're able to measure or test against and you make some projections. And they could be right or wrong, but at least you're saying, I am doing this project for this gain. And I will find some way of tracking that.

00:39:56

The second type is options. Because too many people spend all their lives saying that the only way you run the economics of an IT project are to look at the ROI, and I'm saying not true. Because ROI is just a useful measure for comparing between multiple options, rather than the core decision to do something or not.

I use option theory where I want to be able to price something where I don't know. So the textbook kind of example is a VHS versus Betamax kind of thing. The market is not yet stabilised at some point, there are different ways of doing it. You can't afford to be out of the market because of the business you're in, but you need to know enough about the options so that as the market does mature, you have kept the cost of being able to get into that market as low as possible.

So you turn around and say, there are a small number of things where I have to have my options open because that's what is prudent advice to my board or what prudent execution of my job looks like.

And the third element is big bets. These are things where you could fail gloriously. You know, you are backing something with a very high expectation that this is almost like a venture capital mode.

00:41:23

Seven out of ten will blow up, two will become lifestyle items. You're looking for the hockey stick with the tenth one.

Those are much rarer and in most companies when you go back to this billion, I would say people will give you ten million out of the billion to take the big bets.

And if the market really gets stuck because you're having to compress costs all over the place, then the discount rate kills anything to do with the big bets. Sorry guys, for the next few years we have no big bets. Okay?

In the worst case, the only things that you can do, and you cannot expand that because I gave that as investment. But you can do the same thing with operating costs.

You'll end up with saying, okay, what's the first thing? Turn the lights on and off. Keep the show on the road. Keep the trains running on time.

So what are the costs that are just to keep the show on the road costs?

Because those things are almost volume insensitive.

Then you turn around and say, a second set are volume sensitive and you can reduce your costs of operating by adopting new models of what the volumes would be.

00:42:35

But what you've landed up doing is saying, I can cut my debt costs because I won't be making a new investment on anything. But you still have to keep something that looks like a new investment. Because again in tech, who is going to actually deal with the patches that come up from the manufacturers or the providers of the software that may deal with vulnerabilities from a cyber perspective?

Because you've got to have funded the ability, the number of places where the IoT scared me... I won't finish the rest of the phrase, but shall we say it changes the colour of my trousers is when realised how much of it is unpatched.

Lennox distribution up the wazoo, all over the place, no standardisation with a whole lot of patches that no one is interested in applying, creating daisy chains where you could drive a bus through the security environment.

So how do you actually get that? Well in a large corporate for certain, you're setting aside money to say, this is version upgrades. These are patch and flash level changes. These are people who...

You may not be able to have a black hat team, because that becomes an affordability cycle, but both for investment as well as for operations, you'll end up with these are the things I go to jail if I don't do.

00:44:14

These are the things that I would like to do but they are variable based on some market responsiveness. And these are the things that were yesterday's necessities and have become today's luxuries.

None of them tries to use the phrase or word ROI, because you've got to be almost simpler than that.

Because ROI should only be a differentiating group to say, do I take a bus or do I take a taxi, or do I take a train? Or what other constraints am I operating on? Which is the best way that I do that?

ROI is not saying, should I go or not?

The should I go or not decision is a business decision you've got to be able to take. And if it's a mandate, then you've got to do it. All you're now trying to do is the best way to do it.

Too many people start using ROI as a decision to do something or not, and I'm saying, you're nuts. You've got to have said you can either afford to do it or not afford to do it, because of the economic context you're in.

You've got to know where you have to have your options understood and where you don't. And the idea of an option is you've got to tear it up.

00:45:17

It's a big question which is why I've given a longer answer, because I see too many people treated lightly and suddenly go around saying, I can solve managing all this by spreadsheet. It requires some commonsense as well.

Yes?

UM   Would you say that the new banks have a significant advantage then?

JR   Well I know Anne Boden at Starling reasonably well. We've sat together. I was on the World Economic Forum steering co for fintech and for AI, and she joked with me saying, it takes you 29 steps to open an account at Deutsche. It takes me two. Okay?

Until you show me how your 29 is going to become my two, you aren't going anywhere.

The problems are more to do with how many of the challenger banks are not actually providing the risk side of the banking service, okay?

So the 85,000 guarantee scheme that's there in the UK. What is the equivalent in other countries? So how much of your deposit is protected by law?

And if you look at a full banking licence, you will have to give all of it.

00:46:40

Then the next stage comes a bit like what you saw in the car industry. What comes for free? We've seen it with the airlines as well.

So yes, British Airways looks really useless and expensive, but you pay for your cup of coffee and you pay for your hand luggage, so the stripped down, unbundled cost is not necessarily a like for like compare.

I think that... When I started work I was told the paperless office is about as likely as the paperless loo. And I sort of believed that.

In a similar way, rumours of banks dying have been carrying on from the time I studied banking at university and I was looking at a few hundred year's worth there.

Banks will not die, but they will constantly merge, get taken over, change shape etc. People talk now about Facebook becoming the central bank for the world if they can get their crypto currency there. But the people who believe that actually believe the whole world trusts Facebook.

And yes, there are people who actually believe that. I don't know what world they live in, but the answer is the Monzo, Revolut, Transfer Wise, and Starling... Each of them comes with some new learning because they solve a particular problem really well.

00:48:20

And trying to become a full service across multiple locations etc causes some stressors. I think they are good for banks, because usually they focus on removing friction that has become a monopoly rent that shouldn't be there. So nearly 80% of all banking activity is actually just payments and foreign exchange transactions. And way too much money is made out of those for what should not be a monopoly rent.

And the challengers will start going after that, and then you get PSD2 type implications to say you can aggregate better across. But very soon you hit the same problems I talk anywhere else, saying again... You guys are all too young for this. Les might qualify... But when I came to the UK, and this was 1980, the only way I could draw any money out of an ATM was to put a card in an ATM that belonged to my bank.

A few years later, by the late 80s, small networks meant Link... Link was the building societies' link with Midland, NatWest, and TSB as it was in the case. Lloyds and Barclays were isolated, but you started seeing small networks when you could now draw in ATMS belonging to other banks, but limited choice.

A few years later by the early 90s, I could draw money at any ATM in the UK, but only in the UK.

00:48:20

If I had to go to the US to draw money, I would have to go inside a branch, still put my card into a machine, but provide my passport and use my pin as a second factor of identification, and they would give me cash.

I couldn't spit money out at an airport. So that kind of interoperability that starts in small networks and then expands beyond that is actually the same way the telecoms industry did it.

So you create networks of networks and then you aggregate and you agree interoperability. Even now, the interoperability is based on the number and not the value.

So if I'm allowed to draw 300 quid in country X, I'll be allowed to draw 300 off other currency. And it took a long time before they said, no. If I'm in the wrong country, if I'm in Venezuela, what am I going to do with being told, it's 300? The system is not capable of giving me a small enough note to achieve it.

So again, a long answer but the feeling is the same issues come, because you need standardisation. You need interoperability, you need labelling, you need better ID. And the constraints tend as much to be social and cultural than technical.

Anything else? Yes?

00:51:21

UM   This relates to the last point as well about mergers, and it popped up on BBC News a couple of days ago about the Commerzbank/Deutsche Bank merger.

JR   I voted for both.

UM   Excellent, you were well played. So given the data architecture that you've described to us, you're talking about three tier data operations if you like. The data specialist analytics and the standard stuff, from a technical perspective on the one hand, and a cultural and regulatory perspective on the other, what are the complexities that you would face within an organisation like Deutsche Bank when you merge with another organisation? One, from a technical perspective.

How do you make the data work together across those organisations?

Do you do that? Do you even do it?

And also... Culturally, how do you combine those?

00:52:28