SM So, you’ve heard a little bit about the project. So we’re researching social media data mining, and feel free to ask me any questions if you want at the end. So I have a series of questions to see what data mining is involved with [Media org 1]. So one introductory question is what does your role entail at [Media org 1]?

IR4 As Head of Data Planning and Analytics my role is to broadly extract value from data across all parts of our business, and that’s sales, commissioning, scheduling, marketing, CRM, content management, rights management, all those elements. So it’s all about how do we as a business extract value from data.

SM And how did you come to be involved in this type of work?

IR4 My background was I spent a number of years at Dunnhumby, and Dunnhumby is the company that monetise Tesco’s Club Card data in retail and my expertise there was relevant to the role here and that’s how I sort of started off in this role.

SM And regarding the data mining at [Media org 1], can you tell us a little bit about the history of social media data mining?

IR4 Yes, so it’s been relatively new I’d say in terms of the social media aspect of it specifically. I think in the last two and a half to three years is when we’ve really started to look at it more seriously, and I think we first started using tools like Sysomos and Radian6 just to do some blunt mining and looked at sentiment analysis. I think the big challenge we found as we started that is a lot of the existing tools in the market are largely developed for single brand companies, our challenge is we have multiple channel brands, so [Media org 1], [Brand 2], [Brand 3], [Brand 4], we have multiple programme brands, so [Programme 1], [Programme 2], [Programme 3], [Programme 4], so we have lots of those brands, and the problem with a lot of these tools is they’re great if you’re having a single brand and you’re looking to understand sentiments, but they’re horrible if you’re looking at multiple brands or multiple hierarchies of brands and that’s where they all seem to fall short.

SM Okay, so you previously used the types of tools like Sysomos and Radian6 to mine directly from the host then?

IR4 Yes.

SM And so for how long did that take place?

IR4 Probably about a year and a half, two years, but I think the mining was very, very basic, I think it was more about just understanding volume of conversation, type of conversation and sentiment. And even with sentiment we’re finding it very, very flawed because it doesn’t understand truly language context and nuances in language, you get very flawed sentiment results. And the other aspect of it is our shows and the nature of our remit means that we don’t just want people necessarily always talking positively about our shows, sometimes a negative response to our shows is equally valuable to us because a lot of our content will cause a reaction, it will incite a feeling or a sentiment that people will have about our content and so it’s not always about positive sentiment, it’s also about what reaction has this caused and that sort of thing, so we’re not like a traditional brand perhaps that’s always worried about ensuring that we always have positive sentiment because that would mean our programming isn’t cutting through and isn’t challenging or thinking, isn’t bringing you something new, isn’t provoking you in any given way.

SM And these insight tools like Radian6 and Sysomos and so on, so the kind of data that was extracted using them was bland to the point of being irrelevant almost?

IR4 Yes, I would say so, I think where we started, it gave some basic reporting so everyone would go well, what’s happening, and what you often find is the commissioners, when their shows are being transmitted they’ll often be on Twitter looking at what people are saying just anecdotally, you know, is it trending, is it this, is it that, as we started looking at this we found that it gave you some basic reporting but beyond that what do you do with it, how do you action something, how do you do something differently. That’s where I think it became quite problematic. We then shifted gear a little bit because one of the biggest challenges has always been how do you understand the conversation and the shape of the conversation with TV ratings, so as the show is being transmitted minute by minute you can understand audience flow, so what’s the peak audience, this sort of thing but how do you understand the conversation overlaid onto that.

So we then switched from those tools and we started working with a company called SecondSync, SecondSync was doing at the time a lot of Twitter scraping, specifically around making it relevant to production companies and then broadcasters, so a lot of it was all around our shows and our content, not about just general conversation. They do a lot of filtering, they plug into the Firehose and do a lot of the filtering, and then the most recent development is they now overlay the conversation with the linear transmission. So we now understand the shape of the audience that are watching, and also the shape of the conversation, and those have been quite key kind of insights.

SM So you’re still collaborating with SecondSync then?

IR4 Yes. And we help them develop their next generation ((?)) of products and services and all those things as well.

SM You spoke of other companies, such as Sysomos and Radian6 being problematic in terms of actionable findings, can you explain a little bit about that?

IR4 Yes, I think the whole thing with social is what’s the ROI, what’s the return on investment, and I think we’re still grappling with what it is. We understand the application effect of conversations and we see that obviously as important, but I think it’s still quite hard to measure the actual tangible ROI and I think that’s where you get into with all these reporting tools, it gives you an understanding of okay, what’s the most talked about show last night, what’s this, what’s that, but it doesn’t help you diagnose well, what would you do differently in the show next week, what’s the action that you would actually perform as a business as a result of understanding that this content had the following types of conversations.

So from that perspective I think that’s where there’s limitations, I mean I think anecdotally there’s a few things, so for example [Programme 2], there was a lot of conversation when it was transmitting as to why it was a men only cast for example. So that obviously got picked up by the press and so that caused a bit of a conversation in the media around what’s the next episode going to look like, why is that. So I think from that perspective it’s an interesting listening perspective and it allows you to respond to what the market’s talking about, but in terms of what does that mean for us as a business, how do we then commission or schedule some of our content differently, I think there’s a bit of a disconnect around what it’s telling you and how we might use it as a business.

SM I’m just looking at the questions because we’ve covered some things already and I don’t want you to repeat yourself. So would you say that the social media data mining tools, so Radian6, Sysomos, gave a harness representation of what was out there but not quantitative enough or whether also problems of that order as well?

IR4 Well, one of the biggest problems is that very few of the tools even today don’t give you a view of who’s passively interacting, so who’s just reading the timeline, versus who’s interacting. So all the numbers that get reported it’s about people that are interacting and I would argue there’s probably far more passive viewers of Twitter or social than there are active users that are actually interacting and posting and all the rest of it. And I think that understanding is quite key for us because that’s where you get a real measure of amplification and there’s nothing that’s currently giving you that at the moment. And that’s a real flaw, because if you think about traditional broadcast TV it’s about reach, it’s about scale, and therefore if you don’t have a true understanding of the amplification of a conversation you’re looking at half a metric and half a developed metric and you don’t understand that last piece around not only have you got this active conversation but you’ve got this huge base over here that are passively engaging but nevertheless they’re engaging in additional content, additional perspectives that’s happening on the timeline. And I think that’s a big limitation.

SM Yes, so it includes cases where conversations on social media might be skewed because of an over active kernel ((?)) of viewers.

IR4 Correct, and I think we see that all the time, I think we tend to skew quite young, our market, so a lot of our users are active on social but if you think about the conversation it’s representing the most vocal, almost the early adopters that are maturing, but that’s a subset of the entire TV viewing audience, let alone the UK population, so it’s not a representative nature of what’s happening. It’s a very vocal sort of conversation point but I think you just have to be mindful of how you use it and what you do with it.

SM And regarding what I was talking about in the email exchange, for example the fact that you now try to obtain direct first party permission from your viewers to have more quantitative data, what underpins that sort of decision?

IR4 So I think for us as a business when we started capturing data about viewers we believed strongly that the trust relationship and the value of exchange with our viewers had to be absolutely clear, so we launched what we call a viewer promise so it’s an [Presenter 1] video and it talks about two key things, one’s about transparency and the other’s around control. So on the transparency while we fully accept you can plug the Firehose and it’s all open data, it’s all open source, it’s all openly available, they’ve already opted in on the Twitter side of things, our concern is how do you explain that very openly to a viewer without it feeling like a Big Brother kind of exercise. So because we felt [Presenter 1] to explain our viewer promise, if [Presenter 1] can’t explain it very simply to the viewer and explain back what the benefit to the viewer is how can we justify doing that exercise. And that’s why I think we specifically said we’re not going to pull in just the Firehose data that way and try to understand it, instead we’re going to develop a consumer facing proposition that allows them to take their [Media org 1] registration information and voluntarily opt to give us a permission base, their Twitter handles. That becomes more compelling and I think what we’re trying to work through is what’s the consumer proposition and the value of exchange with the consumer that warrants that type of activity.

SM And so if I understand correctly, correct me if I’m wrong, with using SecondSync you both process information obtained from registered users who have volunteered their social handle? No?

IR4 No, SecondSync is just like a reporting tool, so they plug into the Firehose, all those limits ((?)) and they just give us a very volume based view of what’s happening during a live transmission. Nothing from our registered users is going into that environment.

SM So you’re combining the SecondSync information with data analysed by yourselves or registered users?

IR4 If we can start to bring those things together but what we’d rather do is we’re asking our viewers currently to start to voluntarily give us their social handles but before we can get that plugged in correctly we need to develop a consumer proposition. So the question to the consumer is well why would I give you my Twitter handle? What do you need that for? Why does that help me as a viewer when I’m watching [Brand 5], what’s the benefit? There’s no feature or user case that brings that to life.

SM Okay, so the kind of shift towards the viewer promise and going that way, it’s rather recent then I gather?

IR4 It’s when we first started the whole strategy about two and a half years ago, we were very clear that it had to be a compelling viewer proposition and it had to be with the openness and transparency that’s there as well.

SM Has it been rather successful up to now in terms of obtaining data?

IR4 Yes absolutely, I mean we have now over ten million registered users, we have over 50% of the 16 to 34 year olds registered with us in the UK so that’s a very hard to reach audience but we’ve hit ten million in 22 months.

SM Again, sorry just looking at the questions… So you’ve talked me through the process that is now used in terms of data analysis and on the registered users by volunteering handles and so on. So is that analysed by you in-house then?

IR4 Yes. So we have a team of data scientists in-house and they’ll build the predicted models and all the analytics in-house.

SM And what sort of things are they looking for in the analytics?

IR4 So they’re using machine learning to build predictive models around how do you predict what the audience is, how you predict what their interests might be, so things like recommendation engines, personalising and tailoring the onsite experience. So we do a lot of work in that area, and then we’re looking at how do you optimise our marketing, so for example if you receive an email from us how might we change that based on our perception ((?)) of you as a viewer. So if you watch comedy do we introduce you to more comedy or do we introduce you to some factual content?

So how do you start to tease that out a bit more and how do you tailor the experience to the consumer. We’re using it for advertising sales, so we develop a lot of predictive models that they can then use when they’re selling audiences commercially and that allows us to drive incremental revenue. And then we’ve started to use it with marketing where marketing are trying to answer and understand how they optimise who gets what content in terms of promotional air time when we’re promoting our own shows, so do you get an advert for [Programme 1] or should it be an advert for [Programme 2] or should it be an advert for [Programme 5] for example. So what content do we promote to our users is where marketing starts to play a role in the analytics.

SM Okay. So to summarise, one part is about using SecondSync in terms of the volume and the social media stuff more broadly, and the second…

IR4 Yes, I’d say it’s quite a blunt tool though, it just gives us an understanding of by audiences that are watching content what’s the intensity of the conversation, and that’s largely at best is what it gives us because we don’t understand a true understanding of amplification because it doesn’t understand how many people just passively reading versus actively interacting.

SM And if you were to compare Sysomos with other companies you’ve used previously, you’ve briefly mentioned that, but can you tell me in a bit more detail what SecondSync has as an upper hand with in terms of…?

IR4 Because they’re aligned with TV ratings and television and they made the product very specific to the media and broadcast industry it’s much more relevant for us as a business, whereas the other tools, they seem to have developed listening tools for anybody but they didn’t understand two key things, one a multiple brand hierarchy, so channel brands, programme brands, genre brands, talent brands as well, so [Presenter 2] versus [Presenter 3], and there’s no way to do all that easily in those tools because they’re designed for a single brand where you can plug your brand in and you can understand what’s happening with your brand. They don’t have a brand hierarchy kind of understanding, whereas SecondSync focus solely on broadcast and media and I think they tailored the proposition, they worked with us in a co-development approach to make the product much more relevant to us as a broadcaster by understanding how we are trying to understand viewer behaviour.

SM But the analytics done in-house then on the base of registered users, how does that compare to…? Well obviously you can go in a lot more detail but I’m interested in hearing what are the definite advantages that they offer as opposed to insights tools for example?

IR4 Insight tools tend to just give you reporting and not much else, the analytics tends to give you a set of models that we then deploy back into the business and get the business to change and deliver incremental revenues. There’s a very tangible outcome that happens with the analytics, whereas insights often can be that’s nice to know, that’s interesting, and there’s very little action that comes off it. So I think we’re very mindful that we’re talking about defined business outcomes and changes in the business rather than reporting that’s nice to know, interesting to know, but it doesn’t help us move our business forward.

SM Okay. And in terms of viewer information, I briefly read the viewer promise available online so I got a glimpse, but what does it encompass in terms of information you obtain on viewers?

IR4 So the information capture is first name, last name, date of birth, gender, that’s it. Oh and the email address. And then we optionally ask for a postal code and that’s a great example of our openness policy in that when we ask for a postal code we tell our viewers specifically that this is going to be used to help us deliver more relevant advertising and 75% of our users are voluntarily giving us a postal code. That’s the information we capture, however in the analytics environment none of that is stored in the analytics ((?)) environment and what we do is we convert that to like an anonymous identifier, like a 15 digit number, and all we have is just that number and then all the viewing history. So we don’t actually keep any of that personally identifiable information, none of that’s available in the analytics, it’s only used for communications, that’s sort of thing, nothing in the analytics. So we use an anonymous key basically and you can’t take that key and reverse it back to any of the personal identifiable information because it’s a one way method.

SM And I also read that should the user want to opt out of this…?

IR4 Yes, they can opt out of everything so we give them full control to opt out of matching with third parties like Experian, from first party modelling, so if they didn’t like the way we were doing some modelling for the advertising benefits they can opt out of that, they can opt out of tracking, so there’s all these different controls we give our users to opt out of everything we’re doing.

SM And can I ask you to provide me with an example of how data mining has informed decisions taken at [Media org 1]? And if you could be specific whether that decision was taken on the basis of SecondSync or in-house analytics on viewer registration.

IR4 So an example of in-house analytics is an advertising product we launched, so we built a predictive model, we took registration information, who you are, all of your viewing behaviour, and we built a set of models that would predict your social economic class, so ABC1, C2DE, housewife status, housewife and children. So these are all of the demographics that regular TV has traded against, so we wanted a quite similar product on our digital estate, so we built a predictive model and what we found was the model was we had an independent company, PricewaterhouseCoopers to validate the model and they found the model was upwards of 90% accurate and then we’ve taken that product, we did some trials with advertisers and the product was twice as effective, far more efficient than any other product on the market and we now are trading on that product and because it’s more effective and more efficient for advertisers it’s a more premium product that we charge for. So that’s an example where it’s modelling that we use, first party data, we’ve delivered some commercial benefit and now that’s delivering incremental revenue into the business.

SM And can you think of an example using SecondSync kind of data?

IR4 In terms of an outcome I think the way SecondSync has largely been used is to understand the depth of audience conversation and the popularity based on conversation. There’s no immediate use case or sort of business decision that’s come off the back of that, outside of sort of passive reporting of understanding what content is performing particularly well. I think we have a greater understanding of what genre of content tends to illicit a deeper conversation and a higher conversation, so comedy does particularly well for example whereas some of the docs do well but there’s just difference. We also better understand the shape of the conversation, so for example dramas tend to have a spike before transmission and then a spike after, whereas comedy or quiz shows for example, they’ll have a different conversational pattern so we’re understanding that in greater detail. But beyond that there’s not an action that’s come up that says right, therefore this is how we need to kind of structure some content because we want this shape of conversation.

SM But I guess it complements to a large extent the analytics done in-house.

IR4 Yes, absolutely.

SM Now, you’ve talked a little bit about this already but can I ask you again what is the data you mine used for in terms of business decision, and broadly speaking I guess it’s partly to inform what programmes should be put on, things like that?

IR4 It’s used for sales so they’re delivering new ad products as I said earlier, it’s used for marketing to optimise promotional airtime. It’s used for CRM to optimise outbound email and all of our communications to make it more relevant and more personalised to our viewers. We use it within the online team to tailor the online experience, so when you come to [Brand 5] all of that’s now tailored so the user experience and the user information and the user flows are all tailored to you as an individual based on all of your viewing history. We’ve started to work with scheduling to understand how we might support them in what they’re trying to achieve and we’ve started to look at things like rights and acquisitions. So it largely gives us benefit across the entire business but we’re working through in a very incremental gradual way delivering that benefit. And the way we’ve worked that is we first ran with proof of concept trials first, sort of R&D, and if we can get the trial to work we then switch over and then we scale the project up.

SM So when you are talking about work in terms of R&D first, you pioneer an approach based on the data?

IR4 Yes. In a low cost, low risk way we see that we can deliver benefit and then we sort of justify a bigger scaled up version.

SM How useful do you think the data mined is?

IR4 From our perspective it’s been tremendously valuable, it allows us to deliver incremental revenue which is a great benefit. I think we’re starting to look at other use cases in terms of how you optimise and tailor the online experience, I think from that perspective it’s been quite critical. As we go further across the business and we start to think about the creative side of our business I think that’s where it’s going to be quite interesting to see how it might support the team there in terms of what their challenges are and what they’re trying to achieve.

SM And on the creative side it hasn’t moved too much in that direction yet.

IR4 I mean they use a lot of bar ((?)) reporting, a lot of insights that way, but as yet we haven’t applied some of the analytics on some of that. We’re working with those teams to understand some of those challenges, things they would need in terms of forecasting content performance and that’s where we’re starting to do some early R&D trials with them at the moment.

SM And in terms of delivering incremental revenue, obviously there’s the example you just gave me of the ((0:22:10?)) being a predictive model that can then be modelled. What other ways can it deliver? Advertising obviously is another choice but what else?

IR4 Marketing, so within marketing the way we’ve tailored a targeted promotional content, so when [Programme 6] was coming back as a show when we targeted that at certain users it delivered eight times more incremental VOD viewing than an untargeted approach, which means the more viewing the more advertising, the more advertising the more revenue, the more revenue we can put back into our content proposition. So it allows us to plough all that money back because we run as a non-profit company ((?)) so all of the incremental revenue we create commercially goes back into investing in our content with our independent production companies.

SM And so do you think that data mining is beneficial then, and if so, to whom? Viewers? [Media org 1]? Other stakeholders?

IR4 I think to all parts of our value chain to be honest, I mean internally to us because we have a deeper understanding of our viewers and therefore their interests and their likes to media buyers and advertisers because their ads are now targeted more effectively, more efficiently. To viewers because they get to have a much more tailored and relevant experience but yet to discover content that they wouldn’t ordinarily discover and they get to have a much better experience and they get access to all of our archive material as well when they register so there’s a lot of additional free content that becomes available to them when they register with us. So I think all parts of those key stakeholders receive benefit.

SM And what other methods do you employ to obtain feedback regarding [Media org 1] and its programmes, apart from the digital kind of data mining side?

IR4 We use a lot of research studies, we have our own panel, [Panel 1] and that’s a panel of 10,000 viewers, so we poll them quite often. We have an image tracker every year that runs, Statement of Media Content Policy, and that tracks with key stakeholders how we’re performing, what they think about our content, how innovative the content is, how much it cuts through, how much it challenges, so all of those elements, and then we have other brand tracker and other research studies that we run throughout the year that allows us to understand and report. We also use AI scores, audience appreciation scores, to understand how content’s performing and we also look at obviously the ratings for bar ((?)) to understand how content is performing from a data perspective.

SM The reason I’m asking you is because I’m wondering how you think data mining compares to these other ways of gaining feedback.

IR4 Well data mining and all of the reporting just gives you raw numbers, it doesn’t give you any view of how they felt about the content, what emotion it provokes, any of those things, that’s why it’s really key to bring together data and analytics with research because the data analytics gives you what people are doing but we don’t understand why and we bring those two worlds together in audience technologies and insight because we bring together what people are doing behaviourally with why they’re behaving that way and what they think about content and what their motivations are, their needs, all of those elements you can’t infer easily from data and it’s really important that you get that through survey based approaches.

SM Okay, so they largely complement each other then?

IR4 Yes, absolutely.

SM And can you think of any ways in which data mining offers some distinct advantages or is superior to other ways of gaining feedback, and conversely inferior to?

IR4 I think what it’s allowed us to do with ad sales is delivering committal revenue which I think is quite a key tangible benefit, I think where it’s flawed, as I mentioned earlier, it doesn’t allow us to understand why people are behaving that way. So you can tell the what’s but not the why’s and I think it’s really critically important that you’ve got to bring those two things together to have a true understanding of who the consumer is, otherwise I think you’re always flying with half an understanding or half a perspective of what’s happening and therefore what you need to act on.

SM It’s quite interesting you say that because there seems to be one rhetoric at some point that said that you don’t really need the why question because that speaks for itself.

IR4 No, I completely disagree and I have seen that at a lot at conferences ((?)) and I think that’s a very short term view because I think in the data ultimately you can infer certain things, you can infer interest and intent and all these things, but I think it’s really important that you bring those two worlds together and it’s complementary and it’s fusing those two things rather than data will tell you everything. Because fundamentally you just won’t have, and you won’t have the ability to understand the human nuances, and at this stage given the data we can capture that’s not quite feasible at the moment and that’s why the two things need to play complementarily versus the rhetoric around data mining’s going to replace research, you don’t need research anymore because you absolutely are. And I think the beauty of research is it’s the representative nature of the panels that you develop and how representative that is of either the TV audience or the UK population, whereas in the data mining world we have quite a skewed view because if you look at Twitter for example it’s the most conversational people that have taken to Twitter, it’s not representative of the TV viewing audience. If you look at our own base it’s largely users that are heavy video on demand watchers or very light video on demand watchers and that’s again not representative of the entire UK watching public. So that representativeness I think is absolutely key and I think we need to bring both of those two things to the table.

SM And obviously things like sentiment analysis technically present ((0:27:26?)).

IR4 Yes exactly.

SM Have you encountered any difficulties in making use of mined data?

IR4 I think the biggest challenge has been, you know, the data’s available, the technology’s available, the people has been the challenge in terms of just hiring and finding highly skilled data scientists. So we have now partnered with [University 4] and we develop five year PhD programmes with them and we on board talent at the start of their careers and we bring it through that way. And then the other challenge I think is not so much the analytics and the complexity of the modelling, the biggest challenge is understanding organisation agility and understanding what things need to change in the organisation to actually extract value from it. So you can build the best model but if the business isn’t using it the model is largely irrelevant and flawed. So it’s about understanding the speed at which the organisation could react to something and the processes you need to change in order to bring that change about.

SM So that would encompass issues such as colleagues’ attitudes as well towards data mining?

IR4 Absolutely yes, all of that change from an organisational perspective is absolutely key, because if you don’t get that right for all the sophistication, for all of the modelling, for all of the data, all the technology, it just doesn’t deliver any value.

SM And given that it has delivered quite incremental revenue, your in-house data mining practices, is there an organisation or rather views towards data mining from colleagues and so on who are not entirely convinced, because that’s what we found in many organisations that is one of the challenges.

IR4 Yes I think well now that we’re delivering the revenue and it gets reported through finance I think more people are convinced that this area can deliver value. So I think it’s more a case of how much more can we do, what else do we do if we were to scale this out across the entire business, I think that’s the opportunity now for us.

SM And what are the challenges that organisationally present themselves to that kind of endeavour?

IR4 I think it’s the mixing of the digital world with the linear world, so in retail it’s offline and online and we have exactly the same challenge in that we know exactly who’s watching down to the individual on [Brand 5], we don’t know exactly down to the individual who’s watching on linear TV. Yes, it’s a weighted panel, it’s 5,000 homes, but down to an individual level we don’t know. And that’s exactly the same with retail, you have online stores and they know everything about you, when you go into SainIR4ury’s in store they don’t know anything about you unless you have a club card or one of those things. And it’s a similar challenge that we have, so fusing those two worlds together is probably the single biggest sort of challenge and opportunity if you can get that right.

SM Are there are difficulties involved also in making sense of the data sometimes?

IR4 Not so much, I think it’s more the focusing on business outcomes and let the business outcome drive what model and what data you’re going to need. I think all too often people get excited about the data, build lots of complexity and then they secondarily think about the application, we start with the application and the outcome first and that drives everything else that we do.

SM And sometimes do you come across findings from data mining that are genuinely interesting but hard to put into action?

IR4 No, what we haven’t done in our team is we haven’t said to the analyst team go find some interesting things, we don’t work that way at all, we start with the business outcome first and how to deliver benefit, we then use that to map back to the projects. So we don’t just start with just go and look at the data and tell me something interesting because that never works and I’ve built these teams before and it’s a flawed way of thinking or operating.

SM You had experience in building things that tried to do that before and didn’t quite deliver or…?

IR4 Yes, I have in the past, I mean I’ve built inside teams like this and data analytics teams in the past quite a few times, both offshore, onshore, in organisations, outsourced and I think the thing you always come back to, there’s this belief in business that let the analyst look at the data and they’ll come up with this magic, it doesn’t work that way, you have to have a business outcome otherwise you have a huge investment, a cost in people, technology and data, after three, six, nine, 12 months they say well what value is it delivering, and inevitably they shut down these initiatives because it isn’t delivering value, or they outsource it or they do something else with it. And it’s because they expect the analyst to come up with magic and it just doesn’t work that way, you have to have a business outcome that you’re very clear on, understand all the change that has to happen in business to realise the benefit, let that then drive the data and the analytics and everything else otherwise it’s just perpetual R&D, it’s interesting but nothing happens to it.

SM Oh okay, interesting. What are the advantages as opposed to disadvantages of data mining? We’ve covered that briefly but…

IR4 Yes, I think we have, so it all comes back to outcomes and advertising revenue and all the benefits you can deliver and without it we couldn’t do any of those things.

SM And the change in approach you followed from the previous kind of insights companies towards doing in-house analytics so just a feeling that data mining from the host ((?)) might not be ethical. Is that right, and can you explain [Media org 1]’s thinking here?

IR4 No, I think you’re right, I mean I know it’s on the right side of the Data Protection Act and those kind of things, but it comes back to our view of promise, you know, if [Presenter 1] can’t explain why we’re pulling everything out of the Firehose and trying to understand all about you and there’s no clear benefit back to the viewer why would you do that? Yes it’s legal, yes it’s on the right side and the ((0:32:42?)) are all open and all this, that and the other, but it’s just questionable when you start to talk about that and you try and explain that to a viewer in a very open and transparent way, it’s questionable why would that be of benefit to the viewer and what’s the value exchange for us doing that. And I think that’s the piece that the company should focus in on, you know, what’s the actual benefit to the viewer for you to get that data and for you to have access to that information. And that’s why we have pulled the Firehose in internally and started to analyse that ourselves because we just don’t see how that sits with the viewer promise.

SM Would you say that the decision to move to registration and in-house analytics has been the right one then from the perspective of data obtained?

IR4 Absolutely, I mean I think it’s allowed us to ensure as we go through this journey and we develop trust with our viewers that in no way any application of the data is in any way misused, and I think that ensures that we’re always checking against the viewer promise and we put the viewer promise at the heart of everything we do.

SM And would you say also that perhaps it’s the peculiar status of [Media org 1] as a non-profit and so on and its kind of characteristics that allow that to work out in a beneficial way like this?

IR4 I mean I think there’s elements of that that I think are definitely true, I think ultimately it comes down to what are your brand values and how do you as a brand, as a consumer facing brand, want to interact and have a relationship with your viewers. So I think our history has been we’ve had a broadcast relationship with our viewers and we’re now going down to a one to one relationship and I think as we do that we realise that we had a huge amount of trust equity as a brand and we had to ensure that that trust was always paramount and always key to us. I think ultimately for brands to succeed longer term they have to have and embrace a trust relationship with their viewers otherwise yes it’s legal, yes it’s all these great things, but if you’re not starting that journey of trust now you roll forward and I think very quickly consumers are going to be disenfranchised.

SM Let me just check whether I’ve covered everything. Yes, there was one thing I forgot to ask. Why did you decide in the first place to move towards data mining like three years ago when you experimented with insights companies?

IR4 A few reasons really, I mean right from the very top […], our CEO, declared that data would be the new oil, so I think there was an inherent interest in this and I think he saw the vision as TV as consumed more on an IP based delivery mechanic there was going to be a return path of data and with that return path of data we should be able to leverage value from that data and that’s where I think the genesis of our strategies started from. And then from there it was a case of okay, what is the capability you need, how’s it going to deliver value, where’s it going to deliver value and that’s when we started looking at those things.

SM Well I think I’ve covered all the questions, do you have any questions?

IR4 No, not from me, thank you.

SM No, okay, thank you.