meetinnovators™
where the deals get done

Interview with Ian Clarke from Thoof

Adrian Bye: Okay. So, today, I am here with Ian Clarke. Ian has done a number of interesting things in the past particularly around file sharing and then, also I guess, building an interesting video site. Ian's here to tell us a little bit about that and also his latest start up. So, he does have a little bit of an unfortunate accent. So, I hope you can put up with the Irishness here but aside from that, he's a nice guy and I'd like to introduce Ian Clarke.

Ian Clarke: I will do my best to water down my accent as much as possible so that people can understand.

Adrian Bye: Understanding I'm on a bound to make fun of your accent.

Ian Clarke: Touché.

Adrian Bye: So, tell us about yourself.

Ian Clarke: Sure. Well, I'm an engineer by trade. I have a degree in Artificial Intelligence in Computer Science from the University of Edinburgh in Scotland. Probably, the initial project that really set my career trajectory was one project called Freenet which grew out of some work I did at the university on decentralized systems.

Freenet is a 501(c)3 non-profit and the intent was and is to allow people in countries like China and Saudi Arabia to exchange information without fear of censorship by their government. Freenet is kind of been an ongoing side project for me for the last 7 years. But in parallel with it, I've started several commercial ventures. The first was called Uprizer.

Uprizer developed a peer-to-peer content distribution technology. I left Uprizer in 2002 and set up a consulting company called Cematics. Probably, the best known project that Cematics worked on was with Janus Friis and Niklas Zennstrom. They had previously founded Kazaa and Skype. I worked on a peer-to-peer video distribution technology that now formed part of a piece of software called Joost.

Subsequent to that, I had always been very interested in the concept of how to allow people to monetize their creative efforts online and I realized that advertising…was not the first to realize this…but that advertising was a pretty effective way to do that and together with two co-founders we started a company called Revver.

Revver was the first online video company to attach advertising to the end of videos and share the revenue with content creators. Revver raised so far, $10 million in A and B round financing. I moved on from Revver in

December of 2006 to start my current venture which is called Thoof. The idea of Thoof very simply is to figure out what people want and give it to them and offer that capability as a service to third parties.

Adrian Bye: Alright, so that's a mouthful there. You've raised a million questions. I want to ask you about all this stuff and I think, the kind of the funny thing is that you were telling me before we got started here that you were unsure whether you're really fit for this call or whether you've done the right stuff. I think anyone listening can see that Ian's quite a stud and you've done some impressive things. So, I want to ask you about them because you've done some very cool stuff.

So, I didn't know what you did with Freenet and then learning about peer-to-peer sharing, that's actually become part of Joost?

Ian Clarke: No. So, Freenet has not become part of Joost. The thing that became part of Joost came quite a bit after Freenet. That was developed by my consulting company. I was approached by Janus Friis. This was actually shortly before they launched Skype. Janus approached me and asked me if I could build a technology that would distribute live video over a peer-to-peer network to minimize these central server bandwidth's requirements which is generally what peer-to-peer is good at in this type of scenario.

So, I built a technology which did that over the next 2 years and we handed that off in mid 2006 to the team that then became Joost. So, that project kind of is at the heart of what is now Joost.

Adrian Bye: So, you've written a bunch of code that Joost is now using?

Ian Clarke: I believe so, yes. I hope they didn't rewrite it. But yes, I believe they're now using our peer-to-peer video distribution code.

Adrian Bye: Cool. So, that was just the consulting project you worked on for year or so, was it?

Ian Clarke: Yes, that was about 2 years. I kind of started off working on that full time, then I kind of gradually got distracted more and more by Revver which was also kind of getting going around that time.

Adrian Bye: Alright, why don't you tell us a bit about Revver. I mean, as I understand it, it's basically YouTube clone with the differentiate of people putting videos up there they get a revenue share of the advertising money generated from their video, is that a correct summary?

Ian Clarke: Pretty correct. It is similar to YouTube in that it's kind of part of the latest crop of video sharing websites although, we weren't aware of YouTube. We're pretty well underway when YouTube launched. So, it wasn't quite a "me to" but, yes.

Adrian Bye: Did you start before YouTube?

Ian Clarke: We started working on it in mid 2005. So, yes, we'd kind of decided what we're going to build long before we were aware of YouTube.

Adrian Bye: When you started building it, how did you focus on driving traffic?

Ian Clarke: So, it was interesting, our approach was different to YouTube and in retrospect, perhaps a mistake. Our hypothesis was that the way to build traffic is to form relationships with the creators and this was a big part of our motivation for sharing revenue with the video creators. We felt that if we're financially incentivising the creators then we would have the best content and it's really the content that motivates people to come to your website. Subsequently, that's exactly what happened like some of Revver's biggest successes so far were a group called Eepybird created a video that you may recall from about a year and a half ago where these guys in white lab coats took the…

Adrian Bye: …the Mentos video, Mentos and Coke, right?

Ian Clarke: Yes.

Adrian Bye: Oh, man, I love that stuff.

Ian Clarke: Yes, they got on to like they were on the Letterman Show and all sorts of stuff. We got a lot of traffic from that. Lonely Girl 15 generated a lot of traffic for us, Ze Frank.

So, what we found was that because Revver was and is the website you go to if you actually want to get paid, we were able to attract these people who were semi-professional video creators. I think the problem that Revver always had was that this stuff would wind up on YouTube anyway because of YouTube's lax enforcement of copyright. So, it wasn't quite as big a differentiators as we originally expected it to be.

Adrian Bye: In hindsight, if you're starting over, how would you be driving traffic today?

Ian Clarke: I think the general approach was pretty sound. I think that we probably were a little bit late to market relative to YouTube. So, I think, really what would have helped us a lot would have been being 4 months earlier than we actually were.

Adrian Bye: Do you think it was just 4 months, you don't think it was anything else?

Ian Clarke: Well, I think that it was a combination of things. So, I think, it would have been helpful to get to market a little bit earlier. It would have been helpful to be more exclusively focused on the challenge of attracting eyeballs. Instead of focusing on the video consumer exclusively which is what most of their video sharing website did, at least from the first year of their existence; we were focused on 3 constituencies. Our focus was split between the consumer, the content creator and the advertiser.

So, we raised a similar amount of money to some of our competitors yet we were splitting our efforts between answering the needs of 3 types of user instead of just the video consumer.

Adrian Bye: Right. Hey, it's just a perspective that I have just from some of the people that I've talked to, I haven't gotten close to any of the YouTube guys but I did an interview a couple of months ago with the guy who founded YouSendIt, which I think, they were the ones that ended up hosting like the Iraqi video and all that kind of stuff which is very much virally driven. Now, he runs FlyUpload as well as knowing just some of the guys in the Bay area who really focused on viral marketing which I know that the YouTube guys did coming from PayPal, is everything they did was around virally driving traffic.

Apparently, the only thing that really worked virally was the fact that they could embed videos in web pages. Apparently, of all the different stuff that was tested that was the only one that actually worked really, really

well. But then, they had the other advantage and this is what was really interesting from the YouSendIt guy. He said that every file that gets uploaded to YouSendIt like a video or whatever, that encourage people to upload stuff. I mean, put links in forums and all over the web and send it to a friends to download. He said that the average upload, and this is average across the entire YouSendIt network, for every upload the average is 60 downloads. That's network-wide. So, that means of every file that gets uploaded, it's being introduced to 60 new people. So, that's very powerful viral marketing.

With YouTube not doing their copyright enforcement, you're then competing against Britney Spears and her video is getting uploaded, you're not allowing that kind of stuff as I understand it anyway whereas they are and so people are going and searching and finding all that stuff. So, their level of viral-ness is far stronger than yours and so therefore, their growth is obviously faster.

Ian Clarke: Yes. Well, I think, it's clearly an advantage if you're just going to ignore a copyright. Our thinking at the time was that YouTube was kind of destined to go down the route of Napster which is yes, they achieve viral growth because they're essentially achieving it using materials that they have not paid for and they're not rewarding as the people who created the material are not getting paid for it. So, it is certainly a lot easier when you're kind of able to piggyback on someone else's creativity in that regard. But to kind of do that would be very much antithetical to what we wanted to be as a company.

Adrian Bye: So, even today if you're starting over and we where back then when you were starting, you wouldn't take that approach?

Ian Clarke: I think if we were going to get to market before YouTube or at the same time as YouTube, I think, if we were aware that you could grow at the speed they grew and then acquired before anyone in the media really gets it together enough to sue you, then clearly, those guys made a lot of money and that will be a smart thing to do. But kind of given what we wanted to be at that time and given our perception at that time, that you can't just build a business that is predicated on violating copyright, it wouldn't have made sense. It turns out in retrospect that YouTube did build a business largely predicated on violating copyright.

Adrian Bye: Legal lax enforcement. I mean, they are legal but it's such a gray area that you could drive a tank through it.

Ian Clarke: Yes. So, the areas that they kind of live in is the Digital Millennium Copyright Act and you could say it's lax enforcement but the reality is that if they were effectively avoiding the distribution of copyrighted work for which they did not have permission to redistribute I think, it would have been a very different story.

Adrian Bye: Do you think you would have beaten them if they had been strict in enforcing copyright?

Ian Clarke: Yes, I think so or we would have beaten them if the entertainment industry had been a lot faster to move against them. But hindsight is 20/20.

Adrian Bye: Oh no, absolutely. No doubt about that. Why did you leave that company? Why didn't you stay on board, I mean, it's still a fairly successful company, is it not?

Ian Clarke: It is, yes. Really, my skill, at least the way I perceive myself rightly or wrongly is kind of as a guy who solves interesting and hard technology problems. I think that Revver was kind of turning from the type of company where it's got a bunch of technology problems into the type of company where it's got a bunch of business challenges that needed to be addressed. So, I felt that that kind of wasn't really my core competency and also, I'd developed an idea while at Revver which I felt and feel was extremely compelling in addressing a pretty fundamental problem and I wanted to pursue it.

Adrian Bye: That's a good lead-in but actually I want to backtrack a little bit more before we talk about that because I don't want to go into that into more detail. I'd really actually like to really understand what you did with Freenet and kind of why that came about and what it's doing today and all that sort of stuff.

Ian Clarke: Sure, well, Freenet, to give you the full story, while at university studying Artificial Intelligence in Computer Science, I think I mentioned, and from a technology standpoint, I was very interested in this area called Emergent Systems.

So, examples of emergent systems would include an ant colony or a flock of birds or any situation where you've got a bunch of individually simple components but when they interact with each other they exhibit sophisticated, complicated behavior. This fascinated me and what fascinated me in particular was that nobody had really found much of a practical application for this in technology. Nature had found plenty of practical applications for emergent systems but engineers had not.

So, I was kind of really keen to find some kind of practical way to apply this concept. At the same time, as I was learning about the Internet and you would hear people say things like, "Oh you know, the Internet routes around censorship" and all this type of thing. It turns out that's just rubbish. The Internet is very easy to censor. It's a lot cheaper to monitor people's e-mail and other communications on the Internet than in virtually any other medium. I realized this. I thought you know, even as people where living in this hippy kind of dream world where the Internet would be the answer to the ultimate tool for freedom of speech. I really started to worry that actually it could be the opposite; it could really be a means for control. I've started to think about maybe, there's some kind of way that you could layer a technology on top of the Internet that would allow people to communicate freely with each other without fear of censorship and that means anonymously. If you're not anonymous or at least don't have the option of being anonymous, then you cannot communicate freely because you could be punished for what you say.

So, I realized kind of, have this epiphany that these two things that I was thinking about, one on the political side and the other on the technology side with emergent systems could be combined. That you could build something that was almost operated like an ant colony of software sitting on top of the Internet that would

allow people to exchange information securely, anonymously in a completely decentralized manner just like an ant colony. So, kind of my two interests came together.

Adrian Bye: How do you mean exchange information anonymously like an ant colony?

Ian Clarke: Well, it's decentralized like an ant colony. It doesn't exchange information anonymously.

Adrian Bye: So, there's no central server, there's no central point to it all. All the decisions are all just made locally by the players as they go about their day.

Ian Clarke: Exactly, just like an ant colony or flock of birds. It's essential to avoid any kind of centralization because any kind of centralization can be used to attack or shutdown the system.

Adrian Bye: I've read about that sort of stuff for a number of years and I haven't seen Freenet out there. I've certainly followed your work for a long time now probably 8 years or something like that, I was seeing what you were doing. But I don't see Freenet being talked about as this big thing. What is the actual end result? I mean, does it actually work properly when you don't have a decentralized core or does it just means that everything's slow.

Ian Clarke: It does work well. I mean, it's hard to build a decentralized system with scales and this was kind of part of what was hard about Freenet so we published a paper describing essentially what we had developed which was a decentralized scaleable way to exchange information. We published that paper back in 2000 and actually was the most sighted computer science paper of that year although, at the same time, kind of Freenet was weird because we were getting a lot of attention in the mainstream press, mainly around the area of copyright, like how can you enforce copyright on a system where people can exchange information anonymously and where the system cannot be shutdown. But it also generated interests in the academic press and Freenet was an early example of what later became known as a distributed hash table which is something that you kind of hear about quite a lot these days in computer science circles.

Adrian Bye: Can you tell us what a distributed hash tables?

Ian Clarke: Yes. So, a hash table in computer science, it's a bit like a phonebook. It's a data structure where you'd give it a key, let's say someone's name and it'll give you a corresponding value which is, let's say their phone number. Hash tables are kind of a key component of pretty much any piece of software that you care to see. A distributed hash table is a hash table where the data is distributed over a large number of computers. So, instead of keeping everything in one centralized placed, it might be distributed among thousands of computers.

This whole concept grows out of something called Small World Theory. So, you're probably familiar with the concept of a Small World or the six degrees of separation and this comes from the idea that it's been discovered, that you can get from any one person to any other person in about 6 steps, just traveling along relationships between those people. This is the premise. This is why websites like LinkedIn and similar sites work quite well. But there's a whole area of mathematics behind Small World Theory and it turns out that not any network, if you just take a bunch of things and randomly connect them together, they won't necessarily have this Small World property but human relationships do have a Small World property. This whole kind of concept of Small World Theory and the mathematics behind it is the underpinnings of both Freenet and the other distributed hash table technologies that followed it.

Adrian Bye: So, what does it mean in the real world like, is Freenet, is that HTTP and FTP and other types of services that browse that work on top of the Internet that I can then use them in an anonymous fashion?

Ian Clarke: Yes. So, you can use HTTP with Freenet. So, you can surf Freenet just as you would surf the World Wide Web and you can actually use an ordinary web browser to do it. It is slower than the World Wide Web because there's a lot of cryptography going on and generally, your computer is talking to other people's computers directly which is slower than talking to, let's say Google's centralized servers, but yes. So, you can use HTTP over Freenet and people have experimented with SMTP, the e-mail protocol over Freenet and others as well.

So, Freenet, in itself a platform and you can implement other services on top of Freenet and people do.

Adrian Bye: How does it compare with Tor?

Ian Clarke: It is addressing a similar problem to Tor, however, its approach is quite different. So, with Freenet, any participant, anyone who installs the Freenet software becomes part of the Freenet Network. With Tor, there is still a separation between clients and servers. So, Tor has a list of…I'm not sure what it is, the last time I looked, I think it was about 600 servers around the world. When you connect to the Tor network, you're connecting to one of these servers. So, Tor is more of a client server model than Freenet.

Adrian Bye: Okay. Let's say, I'm in China, I want to browse stuff that's against the Chinese government, is that a safe thing to do, to sit there and do that over Freenet?

Ian Clarke: It depends on your definition of safe. We don't guarantee that you are 100% secure but what we do is we say that it's safer than the alternatives. So, typically, what people in China do when they're not using a system like Freenet is they will connect via short lived HTTP proxies where they really have very little protection from the Chinese government. As soon as the government becomes aware of these proxies, they can either shut them down immediately very easily or more insidiously, they can kind of sit there and watch who is connecting to these proxies.

So, we don't claim that Freenet is perfectly secure or that it is impossible that anyone would discover what you're doing with Freenet but what we do say is look, here are the things that people are actually doing today and these things are incredibly unsafe and incredibly unsecure and Freenet is a hell lot better than that.

Adrian Bye: From doing what you did with Freenet, how does that come that the guys from Joost came to talk with you, I mean, you hadn't done video at that point or is that you had started with video then with Revver?

Ian Clarke: No. So, I got to know Janus and Niklas really because we've both created peer-to-peer technologies, so we got to know each other that way. Nobody had really built an effective peer-to-peer video distribution systems. So, there really wasn't much precedent for doing what we wanted to do but I built peer-to-peer systems before and at Uprizer, I built a peer-to-peer content distribution system. So, it wasn't that big a leap from there to video.

Adrian Bye: I mean, in the real world, does that really work, watching TV and having it distributed across other machines? Does it work well enough to be usable on the day to day basis?

Ian Clarke: I believe it's working well for Joost. I think, in fact, I mean, I've been pretty busy so I haven't followed their progress that closely but I believe that they had some problems a couple of months ago where part of their network went down. It actually turned out that it was the part of their network that was being served from central servers went down and the part of their system that still worked was the peer-to-peer aspect of it. So, I think so far, Joost is having a good experience with the peer-to-peer approach.

Adrian Bye: I guess, maybe it worked better because it's broadcast and it's not transmissions back and forth. So, there's less stress on the system, latency becomes less of an issue.

Ian Clarke: Well, yes. So, it's broadcast. It means that the clients and the system can kind of build up a buffer of data. So, they can buffer like 5 or 10 seconds of video which means that if a piece goes missing, they have time to try to replace it. If they can't replace it, it's not the end of the world because most video protocols are tolerant of loss, tolerant of data loss.

Adrian Bye: Why don't you tell us a little bit about this start up you're doing now and the kind of problems that you're working on solving today?

Ian Clarke: Sure. So, really the basic idea behind Thoof is to figure out quickly what people are interested in and present them with things that will appeal to them. I did some work on this type of thing at Revver where of course, we were very interested in figuring out what kind of videos people were interested in and presenting those videos to them. I actually developed two collaborative filters which is kind of the classic way to solve this problem. I developed 2 collaborative filters for Revver. One of which we licensed to a company called Reddit which was later purchased by Condé Nast. I realized while kind of working on this problem at Revver that actually, collaborative filters were not a good solution. That kind of made me realize that there was a real opportunity to create a type of recommendation engine that addressed what I perceive to be the key problems with collaborative filters.

Adrian Bye: That's totally contrary to what I'm seeing, I mean, one of the best example I can think of is Pandora which started out with its music recommendations based around the content of the music and finding music with similar rhythms. I think, they've made it, I don't know, 6 to 12 months ago that they realized that in fact that wasn't going to provide the best recommendations and they're moving more towards a collaborative filtering model. If collaborative filtering isn't the best way, what is?

Ian Clarke: Well, the key problem with collaborative filtering…I'm not saying that collaborative filtering doesn't work; I'm saying that it has a key flaw which is that collaborative filters require a lot of data about users before they can effectively recommend to those users. Now, in some applications, that's not a serious problem. So, for example, if you're Amazon and you've got a very large user base and you can collect a lot of behavioral information about each of your users, then that's plenty enough information for a collaborative filter to figure out what those users are interested in, but most people are not in Amazon's situation. Most people have relatively little information about individual users and for them, collaborative filters are of limited use. Typically, collaborative filters, they'll do one of two things, so they'll use a user based collaborative filter which works by identifying similar users and then identifying, looking at what those users like that you haven't seen yet and recommending those things to you. That requires a lot of data before they can effectively find users that are similar to you.

The other kind of collaborative filter is actually quite quick but the reason it's quick is that if pigeon holes the user. So, it will kind of assign properties to the users. So, Pandora's previous algorithm, as you said, they admit that there are problems with it, fell in to this category. So, in Pandora's case, it might decide that you like female vocalists and that you like this type of beat etc. The problem with that is that it'll kind of rapidly get into the ball park of what you're interested in but then, it'll essentially pigeon hole you…Sorry, for the mix metaphor, but you get pigeon holed inside a ball park and you're essentially stuck.

Adrian Bye: So, I can't handle the tension here. Tell me why is it better?

Ian Clarke: So, the difference with Thoof's approach is that it's able to figure out what your interest in based on relatively little data and based upon the type of data that, let's say, a website will have about a user right from the first moment that user visits a website or very shortly afterwards.

So, it's kind of like a meat grinder algorithm where you can give it whatever information is available to you about users and it could be anything from; this guy is a Firefox user, this guy is running Window's XP, this guy got here by clicking a link on Boingboing, this guy's geographic location is this, this person just purchased this thing, this person this was the first thing they clicked on.

So, literally, any kind of information you happen to have about the user you can feed into this algorithm. The algorithm will spot patterns in user behavior and use those patterns to come up with an initial picture of what your interests are. As that user continues to use your website, you can feed in additional behavior that you collect about this user that will refine the systems idea of what that user's interest is.

Adrian Bye: Is this predictive analysis type of stuff like scorecarding, that kind of thing or is it something totally different?

Ian Clarke: It's related to predictive analysis. So, basically, it's an inference engine of sorts. It looks for correlations in user behavior and combines that with an anthology of information about the things that are being recommended.

So, we actually built a news website as a way to kind of test this algorithm and demonstrate it.

So, for example on our news website, let's say, you click on a story about Hillary Clinton then, you click on a story about Barack Obama, what it would do is not only kind of do the naïve thing which is okay, this person is more likely to be interested in Hillary Clinton and Barack Obama. But it also looks at, well, who is Hillary Clinton, who is Barack Obama, what do they have in common? Well, they're both 2008 Democratic presidential candidates and it would make the inference that what your actually interested in is 2008 Democratic presidential candidates.

So, it's using the knowledge that it has about metadata and how the metadata is associated with other metadata in order to very quickly come up with theories about what it is you're interested in.

Adrian Bye: So, in advertising contexts this could work. An interesting thing I've been seeing people talking about on the net lately is good advertising is advertising that you don't notice. I think the prime example of that is Google's pay-per-click ads. I mean those ads when you see them, they add value to what you're doing and so we want them there. Generally, I think if people have the choice between seeing those ads or not seeing them, I think, generally, people will choose to see them because they add value.

So, that kind of context is trying to be applied to behavioral targeting on the web. So, that as you're browsing the web, all the banner inventory that you're seeing instead of a retarded punch the monkey type ad, that you're actually seeing ads that add value to what you're doing and relate to what you're looking for.

So, there's a company called Blue Lithium that has gone into a fair direction with that and they've been acquired by Yahoo.

Ian Clarke: Yes.

Adrian Bye: Is what you're talking about something that can be applied in that kind of context.

Ian Clarke: Yes. So, what we are doing, it's kind of a horizontal technology. So, it doesn't really care what it's recommending but in a scenario where it's being applied to recommend advertising for users, it is a form of behavioral targeting and so it's similar to Blue Lithium or TACODA or Tern or a number of these companies.

What's distinct about our technology or at least we believe it's distinct about our technology because these guys don't exactly come out and tell you exactly how their stuff works. So, it's hard to say anything definitive. But what I've noticed with a lot of these behavioral companies is they still kind of make this assumption that we've got a fair amount of information about users and the way they justify that assumption is that, they say, well, if we're widely deployed, then we can tie users together even if it's the same user but on different websites and so we'll be able to collect enough information about users to recommend effectively or to target effectively.

So, in a sense, it seems like they may have the same problem that collaborative filters do which is that they kind of have a boot strapping issue in that they need to be very widely deployed in order to get enough information about users to recommend effectively.

So, our approach is tackling…when applied to advertising, it is behavioral targeting. The difference is that we're kind of humble in the sense that we don't assume that we're so widely deployed that when we see a user, we've got a bucket of information about them. We assume that we may have very little information about them and our technology is optimized to taking that small amount of information and turning it into a useful profile of what is this user most likely to find interesting.

Adrian Bye: So, in a case where we do have a lot of information, for example, we're Amazon.com and we're launching some new type of behavioral targeting of some kind, the fact is even though we don't have much

data today, we can turn it on and tomorrow, we're going to have a lot of data. So, in those kinds of instances would you suggest that traditional collaborative filtering is enough, but in a case of a start up that's just getting going, it doesn't have a data, then your solution is appropriate?

Ian Clarke: I think our solution is more differentiated in a scenario where you don't have much information. I think it's hard because we are pretty young and so we don't have solid numbers; comparative numbers to point to but I do believe that our technology will bring something new to a scenario where you do have a lot of data.

So, I believe that it will be extremely competitive with existing behavioral targeting solutions even in scenarios where you do have a lot of information. But I don't yet have the data to prove it.

Adrian Bye: Alright, okay. Can you talk about exactly how you're basing that or is that the secret sauce?

Ian Clarke: Well, it is the secret sauce. So, I can't go into too much detail, but what I can say is that what distinguishes us is firstly the fact that our system has a pretty good understanding of what is being recommended.

So, I gave the example in the case of news, how it would actually understand that a news story was about Barack Obama and it would understand who Barack Obama is and its able to draw inferences on that basis.

Now, in the case of let's say, product recommendation, the metadata that you would have about the product perhaps its color, perhaps its price, its manufacture obviously and other relevant characteristics about the product. You would give this information to our algorithm and it would actually be able to use that metadata instead of just looking at the product as kind of "here is item 387545" and the only thing you really know about is who bought it and who didn't.

So, I believe we're bringing something new to a scenario where you have a lot of user behavioral information as well but we're still too young to prove that. But certainly in the case where a licensee of our technology needs to deal with users that they may never have seen before and they want to recommend effectively to users that they have not seen before, that's really where our technology is primarily differentiated.

Adrian Bye: So, if I were to come to you and license your technology, how does that work?

Ian Clarke: So, we have an API (XTML RPC API) which is, we've worked hard to make it very easy to integrate with. You sign our license agreement and we give you access to this API and then, whether your site is built in Java or PHP or Ruby on Rails or whatever it is, all of those platforms are very easy to integrate with XTML RPC. Very simply, you send us whatever information you have and you send us information about whatever it is you're recommending to users; be it product or be it ads or be it news stories and for a given user, we will send you back a list of the things that that user is most likely to find interesting along with…

Adrian Bye: Do you end up teaching it things within different niches or is it just generic?

Ian Clarke: It will learn within different niches. So, when you deploy it on your website, it might take an hour or two to learn about your user base and learn about the products that you're recommending.

So, it adapts to particular niches. It's a generic technology in and of itself and it adapts automatically. We don't have to go in and tell it all about you as a licensee, it will learn about you and how you…

Adrian Bye: In the example you used earlier, like clicking on a Hillary news items and then an Obama news item, how does it learn then that they're both democrats, the contenders for the presidency?

Ian Clarke: Oh, I see. So, in that situation, we tell it, yes. So that kind of anthology, we call that our anthology and we have a huge database of data about data and literally, any topic you can imagine, our system will understand that, will know what it is. If you give it the name of a tree, it'll know what the tree, it'll know what to file them in, it'll know what trees are similar etc.

So, we already have a massive database of information about every topic under the sun. That's kind of what's used and then, the system will kind of augment that and adapt it on the fly, but we've already got that.

Adrian Bye: Okay. So, let's say, I come to you then and I want to work with you, I understand the technology, how do you charge me, how much does it cost?

Ian Clarke: Well, I can't disclose our charging structure but that's kind of later in the process and it kind of depends on the specific vertical as well because different verticals have different economics but the basic approach is we would charge you much like, this may or may not be a flattering way to describe it but we charge you in much of the same way that a cellphone carrier charges you. So, it's on the basis of usage and then, components of the charging structure is on the basis of how effective the technology is. So, our people are actually buying the products that we're recommending, our people are actually clicking on the things that we're recommending.

So, it's a pay-per-use structure but I can't get into the actual numbers.

Adrian Bye: I guess you're just getting started. You're obviously a funded start up but at some point, would you be open to licensing the whole thing out as well or is it all going to be pay-per-usage?

Ian Clarke: Well, as you say, we're pretty early in the game so it would depend on the deal.

Adrian Bye: Alright, cool. Anything else you'd like to add, Ian, in closing that we haven't talked about that we should?

Ian Clarke: I don't think anything springs to mind.

Adrian Bye: Cool. Well, Ian, thanks very much for making time.

Ian Clarke: Thank you very much.