Webs of Suspicion – transcript of presentation video

Good afternoon everyone, energy levels high, ready for the beer! One more to go it's us – Webs of Suspicion. We are a mixture of some good consultancy, some good legal work, some top techs, some top banks and some really great academia, and we've got an idea. Our idea is that financial crime is a virus and we want to treat it like a virus, like a computer treats it like a virus. We want to give everyone a chance to crowd source a cure and then we want to spread the cure through the industry, treating it like a virus and wiping it out like a virus.

So, if typologies are a vaccine, what's the problem right now? Well there are a couple. The first one is that real banks don't have the same data modules and they don't describe typologies the same way, that makes sharing hard. It can be tough for FIU's and law enforcement to get the full picture. Smaller institutions can struggle to get a seat at the table to talk about typologies, and in general it takes too long to develop, and crucially, to update typologies. The up-shot is, crime runs across the market, like a virus, faster than the typologies that could prevent it. Some of this is to do with how typologies are developed, so typically right now, a bank might bring a couple of anonymised cases to a public-private forum and there would be a bit of discussion, and you would develop a red flags paper or something similar and then the banks tend to take that paper back into their own organisations and develop their own typologies and their own transaction monitoring approaches off the back of it. We want to turn that process for developing typologies on its head.

Create a cure. A cure is being created pretty much in every bank for the viruses they get themselves, they have their own data dictionaries, they are good at standards banks and actually when we take that a little bit further, there is lots and lots of inter-banks standard for lots and lots of things but there is not one for a typology. So, we can't at the moment syndicate that, so we want to make this a really low barrier entry, based on typology standard, common standard for the banks to work on.

What then happens? Is that you can take a typology defined in standard turns, and very easily, put that back and ingests that into a banks data model. The bank is then in control of the decisions it takes to implement that, and for a large institution that might be via machine learning AI for complex risk management, for a smaller one, it could be simply recalibrating their rules for more traditional transaction monitoring systems. Crucially, we can do all of these without touching and transferring and moving around any of the underlying customer and transactional data, so that is how the vaccine moves. So, how do we get it around? This is just to quickly show you that we have developed a data module for a standard, during the TechSprints this week, and you can just see here, that we have been able to have a standard typology, we've got a central data dictionary up there and we've got a bank data schemer and we've used that to cover some of our TechSprints data this week. So, moving around.

It's tech stupid, I mean I'm a believer in tech and I believe that we can use tech to solve big problems. At the moment, this is a very manual problem, but if we use common standards, base protocols, open standards, off the shelf technology, we can build up a way of syndicating our virus and our cure. So, bank A over here has submitted a typology to our central data store, with its central dictionary, which you have just seen, the typology database is updated and then, it spits itself out across the network, hitting all the other participants in the market who are interested in it and vaccinating them and of course one of the participants in the market may well be an FIU, enforcement agency and they will get the information as well and they will know what types of crimes are out there. We built this during the TechSprint. So, we just took really good, off the shelf technologies, we wired them together and we built a distribution network, not only do we have on the one side coming in type of typology, but we spat that out to the other organisations, and we were able to use it.

Security is important, although we are not sharing personal data here or private data here and we don't have any reason to believe that there is a legitimate problem with the sharing of this data. We have put in a secure protocol to make sure that the right data get to the right people and that only the participants of this closed network actually have access to the data. So, we have done all the work around that to prove that case out. The end of it, we've also created a transparent audit-trail using the block chain, that audit trail keeps a fingerprint of everything that is going on, who has access to all the pieces of data and at the end of the day, who had access to everything and individual users at all points are verified for their levels of access.

So how do we do that? We create, what we called verified actors and they are defined in terms of their role in an organisation, be it a bank or FIU, and then we also classify the sorts of data that sits inside a typology, and we give those individuals entitlement, and this is just a little animation, one that we have built for the TechSprint, and this uses and aligns back to UK government data classification.

So, what have we done with that this week and what can be done? So, this a schematic representation of a fraud typology, it's the network collection typology in the harbr data and this is the thing we showed you the model for earlier, and we have been able to express that as a standardised typology and we've handed that over to Allejandro, he is not here this afternoon but he is our Data Scientist and deployed that by mapping the banks data model in the TechSprint data against that standard, and what happened? Hey presto, we've found some crime, so we have found our virus, this is just a diagram that shows you one arm of a fraud scheme that we found.

So, we had a whole load of victims across all 6 banks, this is one particular victim in Focus South, and this is where everybody else sits, and our main fraudster, and this is a guy called Gerald, up there in Northern Uniform Bank who has targeted 17 victims over the course of 2 years. So, this is a really good

way to just show that, we can get this thing to work in a real world. So how do we make it more accessible?

Well, its low cost, that's the big thing right! That's going to be the first question on anyone's lips is how much does this cost? Everybody has a mechanism for doing this, so whether you've got a human based mechanism or really sophisticated top of the line tier-one banks mechanisms, we allow them to interoperate with each other, so however you discover the typology, you can submit it to the network, and because of the common standard it distributes it, so its commoditised, it's really important that there is commodity everywhere. If you've got a system that uses AI in one way, and you've got a person who does this thing, and investigating in another. We will let them interoperate with each other, so guys that have sophisticated tools can interact with guys with less sophisticated tools but we inoculate the entire industry.

We are suggesting that this is a public-private facility, much in a way that SWIFT kind of is, and there is a subscription charge at the beginning, just to help maintain the system and at the end of it, we've also built a system which is massively extensible, the protocol is open and we allow people to mix and match jurisdictions and come in and join in the network at will.

And its super important as well that FIUs and law enforcement are nodes in this network, so that does a couple of things. Obviously, it helps them to improve their detection and conviction rates. It means that they can also feed typologies into the networks for banks to test out, and equally it's a really good check and balance to help us manage some of the insider risks that might occur, when we start concentrating typology data, we do not want our consortium to get hacked obviously or miss-used. So, we plan, as we said, co-operative, centralised governance covering ethics, data security and risk management, and a public-private consortium can make this work, and it's all possible within current laws and regulations.

Boot strapping, proof of concept, we are kind of here. Dictionary base dictionary done, language done, network created. Early adopters, let's get some definitions into the virus database, it's beginning to provide value to all the earlier adopters perhaps more sophisticated firms. And then we can really take off, once we've got the main market involved we start to get herd immunity for some typologies and we start to get some slightly more interesting applications, and in the long run, we think we can get the whole market in. The data dictionary and the cooperative modules are first.