Detecting and preventing fraud and scams Transcript

0:04
good morning everybody and welcome to the final demo day of this uh
0:09
digital sandbox pilot um uh demo sessions which are bringing us to
0:15
the end of of of the pilot season if you will um today's focus
0:20
is on preventing fraud and scams the final of the three use cases and just to remind you this week we've
0:26
already had uh the vulnerability pilot and the sme landing pilot we're
0:31
recording all of the sessions and they'll be available on the digital sandbox pilot website um so if you're if you haven't been able
0:37
to have a chance or you know people who would like to have a look um then please do um please do kind
0:43
check them out um i'm really grateful to all the teams today for for coming along and prevent uh
0:49
presenting their demos i'm really excited to to see to see what they look like and an opportunity to kind of ask some
0:55
questions of them as well i'll spend a few minutes setting the scene and and and um reminding you all of uh what this has
1:03
all been about and then we will leap very importantly into the demos if you do have questions please pop them
1:10
in the chat and we'll be curating those as we go along um i know the teams would really
1:15
uh welcome welcome your questions and your reflections so so please do to use that um
1:22
and um we will we will create those as we go along so too easy if we could go into the next
1:29
side please thank you so the the digital sandbox
1:35
pilot has uh has been a really exciting uh venture over the last few months
1:41
we accepted 28 teams out of nearly 100 applications to take part in
1:47
this inaugural pilot program and it's really been aimed at helping to support
1:52

and further augment innovation within financial services as i mentioned at the start and this
1:59
pilot between the financial conduct authority and the city of london corporation has really focused on three specific use
2:05
cases vulnerability sme lending and the focus of today's session um for uh
2:11
uh fraud and scams we've had uh 12 teams that have been developing
2:16
solutions uh in relation in relation to this the pilot officially closed uh uh on the fifth last week
2:23
and the teams have had just under three months have had 10 weeks to develop their solutions and today is really an opportunity as i
2:29
said for us to really have a look and see what they have been up to teresa next slide please
2:38
so the purpose of this pilot has been to test really uh several hypotheses uh of ins from insight that we have got
2:45
from our wider innovation offerings and as you all know we run a text print program we run obviously our
2:52
very well-known innovation services around the regulatory sandbox and one of the pieces of insight we
2:57
identified particularly from the text print program was uh there was a space to enable and
3:03
assist innovators to really take that next sleep uh from from proof of concept through to back proof of value
3:09
um and are really an opportunity for the wider ecosystem to observe and to uh and to
3:17
sense check and engage with the the the offerings as they are being developed we often had that one of the key
3:24
features of text prints was the data that we made available to teams to test and and iterate their solutions through through
3:31
the phase of those text prints but when the text prints came to an end those uh that that data was closed down
3:37
and we knew that this was uh what we were increasingly hearing was this was a missing piece in the puzzle to really enable
3:43

um those those teams and those solutions to make that leap from proof of concept and to scale through
3:48
to proof of value so one of the pieces of that we have been really keen to focus on with this pilot has been
3:56
developing synthetic data for the teams across those three use cases um to test we also knew that there was a
4:03
real opportunity as i said to really engage the wider ecosystem whether that was other regulators
4:08
incumbents vcs to really um seek to be able to understand uh
4:15
observe and engage with the solutions that were being developed and so a key element of this pilot
4:20
has been to create spaces to collaborate and observe throughout the life cycle of the pilot
4:25
and that's been something that we will we will talk more about later um trees are the next slide please
4:33
so i've mentioned a couple of uh significant features of the digital sandbox pilot already as i mentioned access to high
4:40
quality synthetic data sets and these were principally developed from a data sprint that we hosted
4:46
last summer a three week data three week data sprint um and the the data that we have
4:52
developed and in terms of scale and volume is significantly more than we would uh
4:59
have for a normal tech sprint and one of the important pieces that we are evaluating as part of this pilot is the efficacy of that data and what can
5:05
we learn more about to really refine and enable greater use and engagement with that data
5:11
as i said collaboration has also been a really key uh element that we wanted to test through this and so an observation
5:17
deck to really enable interested parties as i mentioned such as regulators or incumbents to observe that in flight
5:23
testing um has been a really important piece um there's been an integrated development environment
5:29

for to allow participants to really test and develop their solutions and an api interface or vendor
5:36
marketplace where they where reg techs fintechs other vendors can list their ap their solutions and apis and that's
5:43
about fostering greater interoperability and to really engage and encourage a thriving ecosystem
5:50
the next slide please teresa
5:55
so what's been happening over the last 10 weeks or so well it um the teams have all been working away not just in this
6:01
use case but across all three of them and we've had over 800 users that have registered to create accounts and sign up to the
6:08
platform we've had over 5000 unique views of the website and 600 total views of the showcases
6:14
that we've run throughout the pilot and the pilot program we've been delighted to see the amount of
6:20
engagement from um our from our uh the wider community we've had 40
6:25
mentors from across industry academia regulation and tech have all leaned in to provide expert support
6:32
and and over a hundred different chat channels have been created to to really uh collaborate and share
6:38
insight and understanding we have read as i as i
6:43
mentioned being really interested to understand the efficacy of the data we know that the jupiter notebook has
6:50
been launched over 650 times to query the data um and we've had over 800
6:57
000 api calls to the data set as well and during that time we really uh
7:02
encouraged uh with the teams to engage with surveys that we have
7:08
uh completed as part of an evaluation process to really understand uh where are we hitting the
7:14
mark where are things not where could we do better and improve and actually what are the drivers
7:19
and dependencies around some of those but my thanks to all the teams are really engaging

7:24
and participating in that process so today's session each team will have
7:30
uh sorry teresa i skipped ahead if we can move on to the next slide please each uh team will have 10 minutes in
7:36
total they will be kept strictly to time and this will consist of a six minute presentation after which teresa will
7:42
uh a bell and uh bring them to the end um a metaphorical well maybe not such a
7:48
literal one and then there'll be a four minutes of q a uh from you guys so this is again an
7:53
ask out to the audience who are watching today please bring it bring forward your questions pop them in the chat bar and
7:59
we'll be curating those um uh and uh with before i kind of introduce the
8:06
first team i just wanted to kind of remind uh us all really of why we are particularly
8:12
focused on uh on fraud and scans and we will have had uh the kind of particular raise on
8:18
detroit for vulnerability and sme lending earlier in the week we know that fraud and scams at any time
8:24
are a pervasive and pernicious problem but in a time of covid we have seen that
8:29
fraudsters have really been using this as a hook to enable their scams with over 20 new typologies and
8:35
utilizing covid that have been identified in particular the impersonation of authority bodies such as government
8:41
nhs and who have been a a have been a uh an identity uh uh a
8:47
characteristic of the frauds that we we have observed over the last uh 12 months or so
8:53
uk finance said in september that their members reported almost 15 000 impersonation scam cases in the
8:59
first half of 2020 and that was up 80 over 80 84
9:04
compared to the first half of 2019 with 58 million pounds lost um outside of covid which has
9:11
really it greatly has exacerbated the issue as i said fraud has it has a devastating impact for many in 2019 alone fraudster sold

9:19
over 1.2 billion from uk consumers and four percent of adults have lost in the uk have lost money in the
9:25
in the last 12 months to one or more scams so we know it is a as i said a pernicious and ongoing
9:31
problem which has only been exacerbated unfortunately by covert at a time when people are generally more
9:38
vulnerable or newly vulnerable as a result of the circumstances of covered
9:44
so it's really important that we're having a particular focus on this use case um at this time okay i
9:51
am going to hand over to the first team so i'm delighted to well welcome
9:57
team faculty ai and i believe catherine branter and lawrence cowton
10:02
will be giving us their presentation
10:11
my name is lawrence calton i'm a senior data scientist at faculty ai
10:16
and today i'm going to talk to you about ai explainability for financial services and uh so a quick word on on who we are
10:24
uh faculty are europe's most experienced ai and machine learning specialists and we exist to make ai real for businesses
10:29
across a number of different sectors we believe that a key component of making ai real is to make ai safe
10:35
and to achieve this we've invested significant time and energy in researching and developing ai safety tooling
10:42
so what do we mean by ai safety we think of ai safety as the application of ai
10:47
in an explainable robust privacy preserving manner so explainability enables users to
10:53
understand why the model has made a particular prediction robustness enables the model to determine when it should and should not
10:59
trust the predictions it has made such as if the input data is significantly different from that which it was trained on privacy ensures that
11:06
the sensitive information in the training data does not leak through the model and fairness ensures that all the
11:12
protected groups within the data are treated fairly so today i'm going to focus on
11:17

explainability but faculty has tooling to address each of these areas and we'll be more than happy to discuss
11:22
any of these with you if you're interested after this so when it comes to ai models there's a
11:29
commonly held belief that explainability is traded off against predictive power so you can either have a sort of
11:35
um low-powered intrinsically interpretable model or a high-powered complex black box
11:41
model however with with the right explainability tooling we believe that you can actually deliver explainable and
11:48
high performing ai models now there are many open source
11:53
explainability tools out there um but many of these suffer from a major shortcomings in their approaches
12:00
most commonly these algorithms assume the features in the data are independent from each other and this assumption doesn't really hold
12:05
for the real data sets now faculty's explainability tool doesn't actually need to make this
12:11
assumption which boosts both the accuracy and the reliability of the explanations we produce
12:17
and our research has also enabled causality or any other structure that's contained within the data to be factored into those explanations
12:24
and we also have uh methods for explaining how high dimensional data such as images in terms of a small uh small
12:31
number of human understandable features so that's about our explainability
12:37
tooling and i'm now going to demonstrate how this tooling can be applied to a fraud detection use case using some of
12:42
the data from the fca digital sandbox so the synthetic transaction data
12:50
contained within the sandbox uh contains fortune fraudulent behavior in the repayment of bounce back loans over the
12:56
course of 2020. so this figure here shows the repayment histories for a random sample of 10
13:01
entities within this data but in order to simulate some realistic fraud detection scenario we're going to

13:07
focus on a single date which is the the first of june and look at the data from the period immediately preceding the state
13:13
so we zoom in to the the first of june we can see that uh for a random sample
13:19
of ten entities um the repayment of these loans over this time period was fairly slow and the transactions that are being
13:26
um so the transaction values are also sort of relatively low
13:32
so how do we actually go about detecting fraud in this dataset well the first step is to engineer
13:38
features about the repayment of these loans uh from their repayment histories and these features might be things like
13:44
time since the last repayment or transaction value as a percentage of the total loan value
13:49
so we can then take these engineered features uh pass them through a dimensionality reduction model that
13:54
allows us to derive further features from this data and so we end up with something like a figure on the right the actual sort of
14:00
process of this is relatively uh not important for this for the sake of this talk
14:07
but once we have this sort of new data we can then um pass this through uh some classical uh
14:13
classical anomaly detection algorithm such as an isolation forest and these helps us identify these red
14:19
points here as being more anomalous than the gray points in this figure so
14:24
okay so now we have a group of entities that the model is flagged as anomalous
14:29
and now we need to somehow manually verify whether these are actually uh examples of fraud or whether
14:36
their model has got this wrong so there's two ways that we can do this and we can
14:42
either just pass a list of the ids of these entities that have been flagged as
14:47
anomalous to subject matter experts and let them start investigating uh from scratch trying to infer the elements of
14:54
the data that the anomaly detection algorithm had flagged as anonymous if the data has a lot of features or if

15:00
the anonymous behavior is hidden then correlations between some of these features this can be just incredibly hard to find
15:05
and can be a really really slow process as i'm sure many of you are aware there are also a large number of
15:11
entities that have been flagged then this task just becomes really really enormous alternatively we
15:17
can pass the id and an explanation of the prediction to our subject matter expert and this enables them to rapidly pinpoint exactly
15:23
why the model is flagging this entity and to verify that decision so if we put our subject matter expert hats on and go
15:29
back to this example here's the an explanation of why this anomaly was uh flagged
15:34
and these are things like uh the mean repayment values of the total loan or the mean repayment
15:39
value per day and so we can see that these entities are paying back these loans in large amounts over a very short period of time
15:45
and so if we look at our transaction histories for these flags anomalies we can indeed see that these are large
15:52
large value transactions over a short period of time and we can compare those to our uh transaction histories
15:58
there's a clear contrast between these two different sort of uh states of paying back these loans
16:06
so hopefully this example has highlighted the benefits of incorporating explainability and tooling into your machine learning workflows
16:11
um however as i said before explainability is just one tool uh in in the ai safety toolbox that we
16:16
that we've created so if you'd like to talk more about explainability or any of these other areas please do
16:22
get in touch after this talk and i'll leave our email addresses for me and kathy on screen now but i welcome
16:29
any questions so thank you
16:35
thank you very much lawrence um so we have had a question from lucy how do you avoid the issue of
16:42

criminals learning your red flags and avoiding them that's a really good question um so
16:49
i mean well i suppose in order for them to do that they would need to have some sort of internal knowledge of the
16:55
of the process that we go through um i think i think it'd be relatively difficult to
17:02
infer that um unless they had a large amount of data on uh which flat which anomalies we were
17:08
flagging um without yeah as i say some sort of internal information um but as i sort of described in that talk
17:15
the the i think that the key takeaway here is not so much the the algorithm that you're using to detect anomalies
17:22
um but but the fact that you can explain that that detection and that will sort of help speed up your
17:27
your process of of identifying fraudulent behavior within your within
17:33
your data set by the by the subject matter experts thank you um and another question if i
17:40
may how much more accurate would you expect this system to be this to be over a rules based system
17:49
it's it's very hard to to quantify um so sort of give a sense of
17:56
of improvement i mean i would definitely expect it to be more accurate it will be able to uh incorporate well it'll be able to
18:03
detect different types of anomalies importantly uh rules-based systems will only really be able to detect
18:09
extreme value anomalies whereas machine learning algorithms algorithms will be able to
18:14
detect uh correlation anomalies that would be much harder to detect than than just simply sort of saying
18:21
this is at the the extreme tales of our distribution for a single feature so you can see how how maybe two
18:27
different features trade off against each other and uh maybe an anomaly breaks that correlation uh that you would expect to see so so
18:34
definitely expect the machine learning algorithm to be able to detect new and different types of of uh

18:39
fraudulent behavior that you wouldn't expect to get with a um with a rules-based system
18:46
wonderful lawrence the questions are flooding in so i might just ask teresa if you can let me know how many more questions do we think we
18:52
can we can take before we start to impinge on the next team's time um but um one of the questions coming
18:58
through is can you tell us more about the process by which you acquire expertise from subject matter experts to
19:03
automatically identify red flags in ownership so these so um
19:11
apologies if this perhaps wasn't clear in the talk um the idea with this sort of uh workflow would be that uh so
19:19
subject matter experts would have a would pass on their domain domain knowledge to developers for these algorithms so so
19:25
developers could incorporate as much of that expertise as possible into their algorithms the algorithm would then
19:32
um flag you know a number of entities that they suggest are anomalous and that
19:39
information would then get passed back to the subject manager experts so that they could then verify the the
19:45
algorithms decisions and the and the the benefit of having explainable ai at that point is that those decisions
19:52
are clear and obvious to the subject matter expert and they can really zoom in on exactly why the algorithm is flagged in the first place
19:58
rather than just knowing that you know something slightly suspicious is sort of happening with this row of data
20:06
fantastic lawrence um there's been lots of questions but in the interest of time and moving on to the next team we will
20:12
uh we will move on thank you very much indeed that was really helpful um so the next team to welcome is team
20:18
elax and edgar lopez i think you are presenting on behalf of the today
20:28
thank you very much francesca i'm edgar lopez i'm the founder and ceo of relax
20:34
we specialize in advanced simulation and think crime analytics and we are happy to be part of the

20:39
digital sandbox pilot not only as a participant but also as
20:44
one of the things that help to build this so we are currently working in a
20:51
solution called synthetizer and synthetizer what we want to
20:56
bring to the financial organization's internal samples
21:02
and this internal sandbox will generate synthetic data on demand for the organizations and also
21:08
for companies like the ones lawrence is working just for you to provide
21:13
a sorry for you to understand various your
21:20
solution and so what happens is that
21:25
organizations have a a bunch of say hidden crime inside their real
21:30
transactions and every every financial institutions have a transaction monitor system in
21:36
place and this transaction monitoring system have some of the let's say rule-based
21:42
scenarios or machine learning and degeneration alerts and the big question
21:48
is the one that francesca asked how good they are i mean how much effective crime do you
21:55
find and sometimes this explainable ai helps helps you to do this but uh comparing one to another
22:02
is one of the hardest tasks that we have right now so solutions like the ones a faculty ai is
22:08
producing can really change the way we we currently do in financial crime analytics but it's
22:15
hard to actually understand if these solutions are actually better or worse so that's why
22:22
we are providing the financial institutions with a solution called synthetizer and what
22:28
we want to do is to break some of the problems that we have in this
22:34
field and one of those is a confidentiality so we want to extract from the real data so non-confidential
22:40

parameters and we want to add expertise that we have of what we know about free crime
22:46
technologies and we want to combine this with the uh with the knowledge that the institutions
22:52
have to generate a simulation environment and provide this as a service so the data scientists will be able to
22:58
actually generate different scenarios of synthetic data and these scenarios will be the ones who power the
23:06
machine learning algorithms so machine learning is one of the the same possible solutions that we have for
23:13
solving the problem or for addressing the problem of in-crime but it is required for machine learning
23:19
to have quality high quality date so in organizations like garner has predicted that by 2024
23:26
at least 60 percent of the ai in the world will be trained using synthetic data and there are a lot
23:31
of advantages of using synthetic data and one one of those is that we can actually
23:37
generate these scenarios so we can test several machine learning machine
23:44
algorithms at the time and since we have the labels we can actually benchmark them and i think that's one of
23:50
the key aspects to answer questions like the one francesca had about faculty ai
23:56
how much improvement do you have of other algorithms so the ideal situation is that we do
24:01
this in the lab we train this and we finally go and deploy and time for deployment is one of the
24:08
pains in organizations because since the time that we actually identify some of the threats
24:14
it takes a lot of time for the organizations it could be six months to one year and and this doesn't actually
24:20
help the race of catching the bad guys so if we we can say minimize the time of
24:25
deployment of ai but not only whatever way ai is a effective ai that we can
24:31

that we have been testing in the lab we will be able to inform the law enforcement authorities
24:38
with a quality information that will help them to catch the bad gas so in erlaps i think the the
24:46
digital sandbox was something that was very good for us and basically
24:53
because at the time that we joined the digital sandbox pilot we got a couple of grants from innovate
24:59
uk so one of the grants is for the project called frozen that is an optimization tool for the
25:05
adjustment of the new normal and we're working in in creating the quality synthetic data for
25:12
for ai and we're focusing on in kovitz fraud and the second brand that we got was a
25:18
couple of months later it was cp mark and cpmar focus a lot on on benchmarking this so in trying to
25:24
understand which one is better and so just to give you a little bit of overview and not not
25:31
too much what we want to do is to connect from the real data to go through all the process of synthetic data
25:37
generate parameters evaluate the controls and finally benchmark so frozen and cpmr are just part of the
25:45
solution that we call synthetizer um so just talking about the
25:50
digital sandbox is i think the the project itself is fantastic is
25:56
is one of the the ways that we have to validate our solution uh so supporting innovation
26:02
in financial services is probably the best way to describe the detailed sandbox and that's what they're doing for us
26:08
and we're particularly using uh some of the data sets that that we have to create so some of those
26:15
are the banking transaction banks in uh basin and the banking data
26:20
and we also use the synthetic entities and individuals that were provided before so basically we use this in a in a
26:27
process called bootstrapping so we we learned we learned previously the techniques

26:32
and we apply these techniques on these datasets for extracting the parameters and at the end to generate the synthetic
26:38
transactions so what the digital sandbox provide was
26:43
the possibility to do all the analytics uh to check the validity to understand
26:50
where where are the points for improvement so um i'm not gonna stop too much in
26:55
analytics i have a lot more slides if there are some questions about that and but one important part that helped
27:02
us is the injection of rather than agents with the injection of other agents we
27:08
we study a particular particular problem okay so sorry to interrupt i'm just
27:13
giving you a time check here um if you could wrap to your last slide sure i'm just gonna wrap up so with this uh injection of
27:21
slaughtering agents we can generate different scenarios with all frauds on fraud of injection time
27:26
and this would provide us actually the input for the benchmarking and the benchmarking tool will be a tool
27:32
that will help the compliance officer to understand where is the organization at risk
27:38
and the important part here is that we bring new analytics and one of those is the generation of
27:44
metrics with a hidden crime so finally just to thank some of the collaborators and one of those is
27:51
graham barrow from the there money files that he's been a mentor also in the in this field and
27:59
i'm hoping for questions thank you thank you very much edgar um so we've
28:05
had one question through um when you inject malicious behavior does this mean the
28:11
data only contains four typologies that we already know about or can it start to include unknown typologies as well
28:18
i mean the possibilities are um are quite wide so we can start using some of the things
28:25
we don't and you know artificial intelligence is quite good to detect
28:31

some of the patterns that we know but it can also create a wide range of vulnerabilities
28:36
so so basically the the we start with the concept of injecting what we know but we aim to
28:42
actually create a wide range that will later show the organization the gaps that they are exposed to
28:51
okay thank you any further questions from anybody i can't see any coming through
28:59
thank you edgar that was fantastic thank you very much indeed okay i am going to move on to
29:06
uh team cinetics solutions um and ask chris lewis if he can take the
29:12
floor and present
29:17
um again i'm here presenting what the work is predominantly done by rob bevington and luke abele and our head of data science and our
29:25
data scientists and you know i'm just here to talk on their behalf um i'm sure they'll skype me if i say anything stupid um
29:32
anyway just about synaptics we've got 28 years of experience um at the thick end of fighting ford we
29:38
host the two largest uk uh data sharing databases for the purposes of fighting forward in national
29:44
sewer and the nfi and we got the queen's award for innovation in 2019
29:49
for the work we do in machine learning so we thought this would be a nice opportunity to compare the real-life data and machine learning
29:55
models that we upgrade versus the data that's available in the sandbox you know the synthetic data specifically
30:01
um so um we first off we wanted to try and classify uh trying uh authorized
30:09
payment forward by utilizing transactional data that's held within the same stuff that lawrence faculty just spoke to in the first
30:15
presentation um we basically did some preliminary analysis and came to a similar conclusion as long
30:22
as did and decided that we actually weren't going to build the model this was largely due to the nature of
30:28
the data um the lack of vital membership functions

30:33
um and the fact that the only real variable that we could use is the amount variable we couldn't really create a
30:39
particularly predictive model um so we didn't think it'd be a particularly good benchmark versus some of our existing ones
30:44
so we decided to just move on and try and use a different one of the synthetic data sets and create an alternative
30:50
model so we move towards using the synthetic account data and comparing it versus our
30:56
precision national model which is basically a machine learning model that operates across our data sharing
31:01
consortium to score current accounts in alignment with actual applications for current accounts in the uk
31:08
um so we took the uh current account data and when comparing
31:13
it against the real world world data we saw that there was a some you know sort of medium risk type
31:18
um referrals were created by utilizing the variables that we were able to input
31:23
so it's named mainly personal details and address and all that sort of thing but we had nowhere near enough
31:30
high-risk referrals compared to our real-life model which suggested that some of the sort of introsees that
31:35
actually predict and indicate fraudulent behavior in a real-life scenario
31:41
weren't necessarily present in the synthetic data that we used or indeed our model needed a bit of tweaking to identify
31:46
those more high-risk activities um so i think that the short story will be that we weren't really able to
31:52
classify what it is that we wanted to by utilizing our current account model
31:58
that's uh so that all being said um you can see there's a very massive
32:03
variance between uh what we perceived high risk and medium this to be across the world that we operate in so we see about
32:10
five percent of all current account applications being within that high risk potentially fraudulent
32:15

banding uh versus 0.01 that we identify within the synthetic data um
32:21
and the synthetic data had a much higher proportion of low risk um accounts within it that they did compared to the uh
32:26
the information again that we hold in our national consortium uh we've got a couple of examples as
32:32
well um so on the left hand side we have a medium this referral we've taken from the digital sandbox on the right hand
32:38
side we have a real world applicant and we can see there's a strong correlation that the uh um
32:44
email mailbox field feature in age application are all having a large importance factor on the
32:49
score and the key thing here now i'm sure luke will be laughing at me in the background is that there's not very many negative
32:55
importance factors on the score which basically indicates that we're only getting sort of predicting that
33:01
ford is happening we have nothing that's predicting that ford isn't happening here which is one of the key things to basically
33:07
identify whether something's legitimate or not just not not just the stuff that looks bad but the stuff that looks genuine as well
33:13
um and in our three examples we've got you'll see that we didn't find any features uh in any of the um
33:19
applications in the digital sandbox that were representative of um a uh negative important score
33:27
which again shows there needs to be a little bit more refinement into how the synthetic data is generated and for it to be as predictive as real
33:35
data um so we've got a nice conclusion here i think the key thing for us is a lot of
33:41
the most predictive features that we have on a national basis the like staff email address telephone number um gis
33:49
um are both real world and have quite a prescribed format so it's quite easy to understand
33:54
you know what's fake and what's real uh i've got a real-life use case for example where there was a large forward ring in the
34:00

insurance world where it was a football club's name followed by a series of numbers and
34:06
lots of email addresses generated using that format it's really quite easy to then predict and all the things that fit within that
34:12
particular typology whereas here um the email addresses were basically almost
34:17
nonsensical compared to a real email address which therefore meant that that feature was useless
34:22
and would say the same for the telephone number uh so for us in lots of different formats they never
34:28
use mobile numbers in the digital first world we see mobile number only accounts being more predictive before
34:34
the ones that use um normal uh landline numbers and so again didn't necessarily represent what we
34:40
uh would expect from a actual fraudulent application application and a couple of other things you know so
34:45
all of the addresses were fair you can didn't actually align to the telephone numbers so there was no way to do any sort of
34:50
geographic analysis you know he's applying from leads but he's based in cornwall clearly that's going to be indicative of
34:56
something maybe some sort of compromise identity or something like that but when the whole thing is nonsense um it became
35:02
very difficult to then identify and classify that sort of those sorts of forms as well so we did
35:08
our first piece of work on this and if i've got time we then decided to think well we've got computer
35:13
generated data how about we create a model that uses the computer generated data to identify
35:18
computer generated data within our databases and within our own syndicate so we tried to create a synthetic
35:25
identity model uh using the synthetic data and this was actually really quite successful so we managed to successfully uh plan or
35:33
share these slides doesn't know you won't be able to lead them in time um uh we want to successfully classify
35:38

the best part of 30 impersonation fraud um using the synthetic identity model uh
35:44
across the top 15 of high-risk applications so actually the the fake data that was generated within the
35:50
exercise could be used with a little bit more refinement to help predict uh fake applications using fake identities
35:56
in a real world scenario and we think this is quite an exciting and interesting insight that we got from going through
36:01
the sandbox process so all being said we think the entire exercise is incredibly valuable
36:08
we'd love to use the method that we applied during this um entire exercise to help
36:15
the likes of um elapse define the synthetic data generation process
36:20
align it more to what we'd expect to see from real life data from the stuff that we hold in our data tax
36:26
and we'd love to take part in any future initiatives around this because we do think that our expertise in actually gathering real world
36:32
information and using it to predict actual real-world uh foraging behavior um
36:37
would be you know really beneficial to the people that set up the the sandbox and indeed you know the
36:43
refinement of synthetic data moving forward because there's no arguments that synthetic data is an absolute mandatory requirement in the
36:49
world of gdpr to test and use new technologies in a open and easy manner like we've been able to
36:56
join this exercise with the sandbox so yeah that's it um any questions please
37:04
thank you very much um uh chris and thank you very much for kind of really teasing out and offering something right
37:10
around that real world um uh fake identity uh synthetic piece we do have a couple of
37:17
questions do you have plans um to uh to roll out those fake identity
37:23
models into the real world uh is first question and if i just the second question in as well um what um um
37:31
how commercially viable do you think this this might be so i think the um absolutely we would
37:36

love to refine the synthetic data generation process to basically make the synthetic identity
37:43
model uh more predictive um i think that it's a really good output from the whole
37:49
process and we would love to apply in a real-world scenario i think for us and we basically need to
37:55
use the information that we capture the 300 or 7 million rows that we currently use and for direct generation of these
38:01
models and then identify probably some high-risk features work with the likes of vega
38:06
um and the team at elax to see is there anything that we can learn and show that help you find the algorithms
38:13
because at the moment it's not quite enough i think to be uh to demonstrate a tangible return on
38:18
investment um if it were to deploy it at one of our clients at the moment uh that's not to say that it's not a
38:23
great starting point for what could be an incredibly compelling product um i mean my first hypothesis was that i
38:29
didn't think it was even gonna work because i thought that obviously what edward has done to and the team have
38:34
done to generate the synthetic data it's not going to be comparable to what a forest has done to generate a synthetic identity
38:40
but it transpired that some of the features that actually did strongly correlate across the two um
38:45
particularly around the likes of the email address and things like that um so yeah absolutely that would be the
38:52
short answer to your question thank you thank you much very much i can't see any other questions coming
38:58
through um so i will say thank you very much chris that was a a very uh energetic and
39:07
energizing uh presentation um and uh really interesting so thank you i'm now going to
39:14
i'm going to now move and ask team callsign babesh and chris if they can take the
39:20
floor um so good morning ladies and gentlemen
39:28

uh thank you for uh attending today my name is bavish gayla vp of products i'm joined today by chris stevens head
39:35
of financial services uh solutions um we looked at this in in
39:40
a in a slightly different way so let me just give you an overview of call sign so call time was founded in
39:48
2012 by dr zia hyatt essentially what we do is we look at passive and active
39:55
telemetry uh and we use intelligence and data learning models to identify
40:01
genuine actors and bad actors and essentially what we do is with context give
40:07
friction a security friction to um to telemetry with where the data might
40:12
be bad or or we're not sure but also balance customer experience with uh with uh with
40:20
security so if we go into kind of the the issue we were looking to solve
40:25
and we were looking to solve app fraud but from the lens of social engineering
40:30
and there's three issues that we we came across as we worked on the digital sandbox
40:36
one is detection uh one is the intervention and then one is the overall experience
40:41
um so when we looked at this and we collaborated with a number of uh people within the digital sandbox and
40:48
we also were looking to partner with hsbc and get some real-life examples uh of
40:53
of of these three three three areas uh what we found was um how do you
41:01
how do you detect when a customer is at home using their device in their location and making a payment
41:07
which could be a fraudulent payment through social engineering the other one we found was uh when you
41:13
provide generic error messages they become noise and customers just then ignore that noise and just carry on
41:20
and do the payments the the other two things that we found were sportsters get very clever and
41:27
understand the customer journey from a banking side and they're able to coach the the
41:32

vulnerable customer through uh through that that journey and um
41:37
and ultimately then then are able to get the money from from the customer and
41:43
then finally how do you make sure that you're only alerting when it looks like it's going to be
41:49
fraudulent and letting everyone because as soon as you start alerting everyone uh it becomes noise so what i'm going to
41:56
do is hand over to chris and chris is going to go through uh the solution and also do a quick demo it's over to
42:02
you chris well cheers bob yeah so the courseline technology is embedded
42:08
in the user journey so we passively analyze things like the device location
42:13
behavior and we combine that with on other analytical risk feeds so doing some things like looking at the
42:20
transaction risk doing some beneficiary analysis um you know telco intelligence um assessing the
42:27
customer profile and then you as we touch on the behavioral biometrics as well which is a great way to
42:33
identify a change in the in the user behavior so we use all those bits of information
42:39
and when the thresholds are are breached we then introduce these dynamic interventions so
42:46
these are our questions and and and fraud warnings are very tailored to the specific risk that's been
42:52
identified but to baba's point we don't you do that for the majority of transactions it's the minority that
42:59
actually are presented with these warnings so when customers see them they know that something's a bit different
43:04
with a view then that either they can be um you know we can then inform the customer
43:09
and the customer realizes they're being scanned and they stop the payment or actually we capture enough information to know that
43:15
the customer is taking long to answer these questions they might be typing differently and so
43:21

we can actually infer that the customer is being socially engineered so that's kind of our approach to this
43:28
and i'll just show you a quick demo of how this works in practice so this is an example and bank
43:35
invitation so it has our products baked into it so i'm going to log into southfield bank
43:42
so i'll go ahead and type in my credentials
43:50
click login and i'm logged straight into my account so when i actually logged in there we performed a lot of analysis around
43:56
the device location and behavior and actually we align with strong customer authentication just by typing that username and
44:03
password i've actually performed three factors so the device is a possession factor that's 100 recognized for me work laptop
44:10
location as with everyone it's not changing too much at the moment and key strokes not only did i type the
44:15
correct password but the way i typed is consistent how i normally type and so that acts as the inheritance factor so if i go now and
44:22
make a payment and i'm just going to go and set up payments on the payback and put
44:29
in some account details for him uh and i'll pay him for dinner this was
44:34
quite a long time ago a bit of an overdue bill when we're allowed to meet um save that and confirm it
44:42
and the payment goes through straight away now if i repeat the process and i might do slightly higher risk
44:47
transactions say i'm paying hmrc i'm going to put in the actual hmrc bank
44:52
details and their and their account number so this is something that will be
44:58
assessed by our system and i'm going to play my self-assessment
45:06
save that and confirm i get a different user journey it's asking to step up the authentication
45:12
um and i get my sms through on my phone type this in
45:20
click next and confirm that so this isn't anything different to the you know what you're used to you know with
45:27

your existing banking news setup but essentially we're assessing the risk and we're not changing anything in particular related
45:33
to the user journey but what i'm going to do now is i'm going to log in and
45:38
i'm on the phone so what's the difference here well i'm typing with one hand for a start
45:44
so i'm typing in my my credentials and i'm being coached to to make this
45:51
you know to log in and so obviously my behavior is going to
45:56
be a little bit off to how i normally okay chris just a
46:02
type check here so i know you're doing a demo if you can go through that a little quickly and throughout the six
46:07
minutes we're getting that so i'm gonna uh i'm stepped up to facial recognition that's what we've set up in
46:13
the journey i click continue and now i'm going to go
46:26
ahead again
46:32
copy the process
46:52
so i'm going to step up to facial recognition provide my face logs in
46:59
now you can see the keystroke is down at one percent so it recognizes this one-handed typing this deviation
47:06
i'm going to go and make a payment so now i'm going to pay someone new and i'm going to say i'm going to
47:11
pay um chris stevens so i've been asked to move my money to a safe haven
47:17
um i put in the account details and i'm putting in you know an amount
47:25
so now i'm going to click confirm and i get a different user journey so i'm presented with these interventions
47:31
did i expect to make this payment today no i was on the phone to my bank they said i need to move my
47:36
money is this an unexpected pay request from bank of police or hmrc it is i click yes and then i get
47:44
presented with the warning um you know tailored warning and i knew asking you whether i want to wish to
47:49

proceed so i might click stop payment and new the payment is cancelled now this is all driven by our back end
47:56
um our decisioning component that determines you know what is the next step what questions should you ask
48:03
next and what conditions under which that that question should be asked so that's a quick kind of demo of
48:09
how our system works it's very flexible you get full control of those journeys
48:14
and you as you see new fraud attack vectors is very easy to update those warnings and the conditions under which
48:20
they're applied and yeah have to take any questions
48:28
thank you very much indeed chris we've had a few questions coming through so how does your keyboard input analysis
48:34
compensate for people who use password managers to auto fill details
48:39
so yeah we we recognize what's a deviation in the norm for a user so where they normally use a password
48:44
manager you know we pick that up um but it's very much a case of we
48:49
we also look at things like how long they they take on the page you know it's not just the password page we look at
48:54
it's all the different pages when they're navigating through the system thank you and earlier on in your
49:00
presentation you said we found on the app uh upfront on this on the rise slide can
49:06
you talk to us about how you found this in terms of outflows on the rise so we
49:13
we speak a lot to a number of banks so we in in combination with hsbc we were going
49:19
through this and we actually looked at some of the the most recent fraud trends
49:24
and so everything from the vaccine and scams to there's a big one at the moment around bitcoin you know everyone's trying to
49:30
buy bitcoin because it's going up um and yeah it's tricky if you go through coinbase so there's people that
49:36
happily help you buy some bitcoin so we've got a whole load of industry standard templates with these
49:43

questions behind the scenes that help detect all these different floor demos but ultimately it's always changing and
49:48
so that's where our clients can make these changes quickly wonderful
49:53
chris there are quite a few more questions in the chat so maybe i can ask you to turn your attention to those um
49:59
whilst i now move on to our next uh presenting team chris thank you very
50:04
chris and bob thank you very much indeed so our next team up is financial network analytics um and
50:12
brandon smith i think you are taking the floor for the team
50:21
great so we're fna myself and matteo are here and today what we want to talk about is how we use
50:27
the sandbox pilot's synthetic data to apply
50:32
two different schemes for compliance organizations whether you're looking at fraud business risk or any money monitoring to
50:39
conduct a uh basically a um ensemble-based approach to identifying
50:46
anomalous or high-risk behavior very quickly in a lot of data so uh we'll get right into it a little
50:52
bit about who we are can be seen here and if you'd like to hear more about what we're doing and other solutions we
50:58
have in other areas um of course we're here to do that uh we're heavily participating you know we
51:05
participate heavily in academia as well as the business and industry our work spans
51:13
uh academia central banks financial market infrastructures corporate banks and
51:18
uh more uh direct with some of the work that i do personally the department of defense and
51:24
intelligence communities so we'll jump right into the problem we have today which is that most the time
51:29
in compliance risk monitoring most of the data which are the cases that are generated by centralized
51:35
monitoring systems are um unproductive case volumes so it's uh it's unlikely to generate a
51:41

suspicious activity report or some sort of alert that will uh actually inform law enforcement
51:47
or government of what the actual typology of risk is so what we're simulating here to the
51:53
left is we've selected one node that was in the fca sandbox and just emanating from four degrees of
51:59
relationships with the one node you can see some of the statistics that we have so you know 37 um million dollar million
52:07
pounds plus worth of transactions and transactions ranging from 285 all the way up to
52:12
almost 100 000 uh 411 000 individual transactions across three
52:18
thousand two hundred and twenty five entities that represent seventeen distinct business segments all of the business
52:24
segments available actually in the sick code database so generally what we would say is well
52:29
that's already a data reduction we're only looking at the ecosystem around one node and there are one entities
52:35
behavior for one day and then what we say is well what if we took the traditional risk score that was already
52:41
in the data so we were able to ensemble that risk score based on the back all of the data about you know
52:47
maybe risk uh credit risk scores and things of that nature and you still have 568 entities to
52:54
consider if you just looked at the top 10 percent of the risk in this network so that's still too much for anybody to
53:01
to really dive in on why is that because compliance-based rules are are designed for uh
53:08
keeping keeping financial institutions compliant more than they're more than they're oriented toward
53:13
actually finding suspicious behavior and that criminal enterprise adaptation can outpace regulatory kind of red flags
53:21
and uh schemes that we come up with especially in rules-based monitoring to you know try to catch them in their
53:27

financial transaction behavior so what we suggest instead is that instead of focusing on that focal
53:33
entity which is what most people do today when they generate a case or they generate um let's say you're a company that wants to
53:40
underwrite this person for insurance doesn't matter what it is we use the full breadth of just transactional data
53:46
as well as the data about the people in their network such as their risk scores all the things all the data you would get from
53:52
something like companies house or another data aggregator and we suggest that you evaluate networks
53:58
uh evaluate the risk of your focal entity in this context so what you have at the end of the day
54:05
is um relationships to other people that can influence the initial score this could
54:10
be a business risk score business failure score uh this could be an aml risk score
54:16
um but you base the the risk you you modulate the risk of the focal entity
54:21
based on their relationship thereby somebody who seems very safe at first could actually have an increase in risk
54:27
or an increase in business failure risk uh or somebody who seems as though the risk is very high to begin with when you
54:34
consider the behavior in the rest of the network actually they're they're they're transacting with people in a manner that
54:39
makes sense for their network and therefore the risk can be seen as decreasing how do we do this
54:46
uh we basically use uh two different approaches the first one we just showed was creating a behavior risk score based
54:52
on the relationships and the relationship data utilizing network science tenants
54:58
mateo our data sciences here is here to answer any questions that you may have about that and then the second is that we used a
55:04
neural network that was trained to identify members of each segment that say they're
55:10
a member of one part of a business segment but actually behave as another part of the segment
55:16

prior to the fca sandbox we had tested this approach on real data from the world input output
55:22
database and we were able to find in the simple visualization members that say they're supposed to be
55:27
uh one segment but in their behavior we see them as outliers well entrenched in another segment so we wanted to then
55:34
bring this uh to bear as well as combine the relationship risk scoring uh with this approach in the fca sandbox
55:41
data so um the results of this are actually pretty good uh what we were able to do
55:47
is take a look at two weeks worth of transaction data perform a day-by-day analysis of it and
55:52
then identify day-by-day node by node um what are the most risky
55:59
members of the network given given a base node so so if you have one member of this
56:04
network because you have to consider every member of this network in your monitoring for every network in the monitoring you
56:09
can generate a list right up front of the most suspicious uh members of the network
56:15
suspicious being those that don't conform to their segment combined with how they uh permeate risk through the network
56:23
so this is what the network looks like uh by itself this is photo one as we talked about
56:28
and twenty 3225 entities uh apologies this this this little uh callout box is supposed to pop up in the third picture
56:35
photo two as we talked about this is if you just um decreased it to the top ten percent of
56:41
your normal risk scoring and as you can see by node size being the risk it's very difficult to discern
56:46
who the risk is but here in the third picture this is what we're able to reduce that whole network to
56:52
is around the focal entity you have all of the industry sectors that they represent um by their shape
57:00
you have the volume of the transactions that are going between them by the density of their links uh so all
57:06

of this is customizable and then as you can see there's blue and dark blue the dark blue nodes are those
57:12
in the networks who um say that they were one thing
57:17
but behaved as another so in this case this vertex id which is one of your um fca entities uh and organizations
57:25
uh they said that they were a member of you know sick codes 86 through 88 and health but
57:31
we were actually able to predict that they were actually a member of a completely separate sector so instead of being in sector 11 we predict
57:37
that they are in sector 2. so um hi brandon sorry i'm just giving you
57:42
a time check that your presentation time is up oh absolutely so what that looks like in
57:48
practice is we've simulated that we have uh this whole network um
57:55
here is what would it look like if you tried to reduce that giant network to just that same focal id
58:00
but the output on that focal id is actually here we can build the network and so if
58:05
you're the investigator um or you're doing due diligence on this you would build this network out you can
58:12
say what links matter to you uh you know maybe you can also do this by transaction amounts so on and so
58:18
forth and then the idea here would be um as these load because it is loading through
58:24
a ton of data um you can come in here and then say well i would like to just
58:29
know is there a suspicious actor meaning they say they're one thing
58:35
but behave as another according to their category which are now in orange and instead of the business failure score we'll take a look at the new
58:41
business failure score the enhanced so now what we have is a very quick way to say out of
58:48
thousands of entities i care most about these ones here that are non-conformers as well as the ones that
58:53
have an increased business failure score that negatively impact this focal id that i'm looking at again
58:59

for aml or maybe business decisioning so in total had you gone through all 3 000 nodes you would
59:05
have generated this list of this 28 that are in its ecosystem that that you should care about the most
59:11
and as you can see here all of the data about from the sandbox is here about each node that concludes our
59:17
presentation and we're happy to take any questions thank you very much indeed uh brandon um
59:24
so we've i think we have time for uh maybe one question so i might direct you to the
59:29
uh the chat to see if you could answer any more that come through please um how does the relationship network and
59:36
analysis work when the customer has multiple um accounts as accounts sorry at multiple banks
59:42
does it require banks to share data with each other that's always the concern we have and often
59:47
my experience in trying to improve compliance monitoring systems does include multi-bank analysis when i was
59:53
at citigroup and what we found is that um high-risk individuals are more likely to have
59:58
their behavior explained as lower risk when you combine banking data across banks um
1:00:07
the hard part about that is yes you would have to have a very targeted reason to you know kind of request information
1:00:13
from another financial institution about the same customer if the same client has multiple accounts within your same firm
1:00:19
in the network science point of view what we would do is uh just merge those entities or you
1:00:24
might be able to uh decipher behavior between let's say organizational accounts versus individual accounts so you may want to
1:00:30
keep them separately and monitor the behavior separately or combine them and get a more holistic view of
1:00:36
you know here's brandon smith's personal checking but brandon smith uh also owns the accounts that are um
1:00:42
transacting for brandon inc for instance okay thank you as i said there's a few

1:00:50
more questions in the chat so if i can direct you there to maybe pick some of those up that'd be really helpful thank you very much team financial
1:00:56
network analytics i'm now going to go to uh team like stego i'm sorry if i pronounced that
1:01:03
wrong my apologies but um janae and rob you are a leading leading the team
1:01:08
welcome and over to you
1:01:15
so my name is janet i'm here to present our proof of concept um we are a pretty new startup we less
1:01:24
than six months old so it's been a bit of a whirlwind um last three four months to get this proof of concept up and running we are
1:01:30
working together with a firm in south africa called cybrin
1:01:35
who provide core banking platforms across africa at over 300 customers and
1:01:41
we are building this for the bill and menindee gates foundation it supports their level one project about bringing financial
1:01:47
products and inclusion to the poorest and our first implementation
1:01:53
is with emergency foundation which is an open source switch and so kind of think analogous to faster
1:01:59
payments here in the uk so first up why open source
1:02:07
and we believe it's a shared problem and what we hope with axio with the product
1:02:12
is that we create a starting point for fintechs all over the world so i will dive into
1:02:20
our actual concept so this is the fraud risk management um
1:02:25
holistic concept that we have so you have um a payment being fed in or transactions fed in from merger to
1:02:32
us is on preparation we are currently doing a rules-based approach and and for
1:02:39
the rules in the typologies at the moment we have identified 270 typologies
1:02:45
from there we have an analysis outcome and the transaction is fed back into the hub and and the transaction
1:02:53
is processed so through this whole journey one of our
1:02:59

big questions and the big learning from us through the sandbox is understanding our operating models
1:03:06
so we we came from a place where we thought are we going to have to go either fully distributed or a
1:03:13
completely invaded system and it's about understanding our characteristics um that we need to be aware of one is
1:03:20
that the hub and the financial institutions are going to want to potentially do their own
1:03:25
thing maybe you need to have a trusted party in between and the sandbox has really made us understand how we would actually
1:03:32
deliver a semi-attached or a standalone system so that has been a big outcome for us
1:03:39
and our vision is probably more of a semi-attached where we have shared compliance and a trusted partner
1:03:45
or a hubble operator which allows for banks to which direct some points banks in a
1:03:51
certain way so that they can actually do certain types of investigation so what was our challenge we needed lots
1:03:58
and lots of data so 270 typologies trying to hide
1:04:03
our fraud in all of that data is a tricky problem for us and the fca sandbox has
1:04:10
helped us to do that especially because we need to run at 3 000 transactions a second and being able to scale and handle 10
1:04:17
000 so that almost ends up being a billion transactions a day what have we done in the sandbox um
1:04:26
we have had to adjust the transactional data so for instance more granular timestamps there was only
1:04:31
i think four timestamps that need to be expanded and allocating more individual um
1:04:37
data so passwords imei driver's license etc and the other part which we've
1:04:42
gained valuable help and coordination and learning is through the mentors and participants
1:04:48
so broadening our view of what is possible as i touched upon in the operating models we have
1:04:54
collaborated with siddiqi and i'll get to that as i present our proof of concept and thanks to the mentors

1:04:59
i've mentioned a few here but there's been a lot more that we have talked to if you have helped us on this journey
1:05:06
so our actual proof of concepts if you want to see the demo um i've provided and think i'll try and
1:05:12
slice a video later into our showcase but what we have done is take 20 000 users so that equates to a
1:05:20
million transactions so typically a year's worth of transactions and we fed it into our system and across
1:05:27
four typologies and the actual top parts here has shows a kind of the transactions a
1:05:34
second over the time that we played this through we went to the 10 000 where
1:05:41
our issues and what the data really has helped us with is understand where our system broke so
1:05:46
as we build up more and more historical data the system slows down so we can still achieve our 3000
1:05:52
transaction seconds but we need to be aware of this as more and more historical data is built up and
1:05:57
more analysis is done against these typologies
1:06:02
apologies scored the results look something like this so we can see the
1:06:09
highest scores is from say ashley scott playing a russell hunter but there's 102 ashley scott's and 37
1:06:16
russell hunters in the data so in the next part of this journey you have a problem
1:06:22
that you need to investigate so how do we investigate well we have partners and people and
1:06:28
solutions that can help us in this case sadichi so in the next presentation sadie she will
1:06:34
as part of their demo show how this is done i should also add it in real life you
1:06:39
don't see all of this information this is just here to kind of highlight what we have and that is another crucial point and the help we
1:06:46
need from other tours to be able to have financial institutions speaking to each other
1:06:51
and discussing in a way that doesn't um break any data and privacy rules
1:07:01

um to be able to do this investigation so what is our next steps
1:07:07
we need more realistic data synthesis and doing it on a larger scale as i said there's 270
1:07:14
typologies um and we have a lot of raw typology calibration to do this was
1:07:19
just a proof of concept we've achieved what we wanted and the mvp will demand a lot more
1:07:25
another key area for us is the security and privacy side of things um the actual engine itself
1:07:32
will be open source but you can't open source the rules and typologies you can't be as redeemed the thieves
1:07:38
cookbooks and give them two bad actors and for them to dream up new ways that they may not have
1:07:45
thought of we need to increase community participation i said this is an open source project
1:07:51
and the more participants the more interest we have in being able to help them build a better project and the tool is in everyone's interest
1:07:58
and lastly we need to also think about how our commercial model will wrap around this so that we can continue and support the
1:08:06
journey that has been started by the billionaire in the gates foundation and we want to continue this journey and
1:08:12
for that we need a commercial model and to work that through this year as well so that is me thank you
1:08:21
thank you very much indeed janae um i can't see any questions coming through so did
1:08:27
this there was always you can quit chris sorry apologies um i clicked answer live and then press
1:08:33
done because i was gonna um type a response so chris asked about the typologies um they're
1:08:38
held by the gates foundation um and one of the things we're looking at is as as janae said we'll have a close repo
1:08:45
for the rules and typologies um it's called the thieves cookbook for a good reason um if we share
1:08:50

all 270 typologies there's a whole bunch of fraudsters who are suddenly going to get new ways of trying to circumnavigate a lot of the
1:08:57
controls um the rules that we're creating will have both the manual controls and the digital
1:09:03
controls that we plan to instigate so any fintech doesn't have to start from scratch
1:09:08
but that process of vetting and giving access to that is a process we're working through at the moment
1:09:13
um chris if you are interested would love to chat to you because that's one of the big questions we didn't cover off in this demo
1:09:19
there's a whole model that goes behind it with apricot um you know it for the purposes of the demo
1:09:24
it didn't have as much value but if there is something um you know that you want to discuss i'm more than happy to discuss that with you
1:09:30
because that's something we are trying to make sure is available in a controlled way thank you rob and we've had a really
1:09:37
interesting um question through around quantum and would quantum more quantum inspired tech
1:09:42
help with the vast amount of calculations required good question if someone actually knows
1:09:48
how to answer that and wants to join we're using a basic rules engine sorry we're not advanced enough as a machine and one of the things this did teach us
1:09:54
is that we need some data scientists um it's an open source product if someone's got some ideas and thoughts in that
1:09:59
please do feel free to reach out to janae and myself um we would happily have some
1:10:04
um proper insight we've got the resources to throw out this so yeah please come and talk i mean even
1:10:10
with machine learning we have to be wary as to the sort of people and our potential users of this if this is
1:10:16
somewhere in africa they may not be able to have all the bells and whistles and so we need to have a system that can
1:10:21
cater for both sides of the market thank you an important point there about jurisdictions and applicability across
1:10:28

jurisdictions thank you both very much indeed you have teed us up very nicely in your presentation for our
1:10:34
for our next demo sadichi um welcome uh david cunningham who i think is uh
1:10:40
leading off for for team sadichi um i can see you've started sharing on the screen so i
1:10:46
assume you are ready to go david
1:10:53
so atsudici we are focused on providing world-class identity and security solutions to
1:11:00
prevent financial crime and enable commerce so we're really focused on delivering certainty
1:11:05
in this digital world in a simplified manner as possible with a really good team based uh about
1:11:12
20 of us based in the uk ireland germany belgium and tenerife and
1:11:17
entirely focused on on really delivering great solutions the work in the sandbox for us uh you know
1:11:24
was really great to get into the sandbox we were looking for collaboration with teams learning from
1:11:30
mentors uh hopefully some interested parties to use our technologies and we got all of that
1:11:35
and more uh as i'll demo in our collaboration with lex tago in a moment you'll see that we worked
1:11:42
really closely together which was a great learning experience synecdic solutions we really feel there's a lot we can do together there
1:11:48
and we are looking for some research opportunities with npc for aml uh the mentor engagement
1:11:54
from jonathan frost for us has been invaluable and also denise uh ruddich really just to lean on that
1:12:01
expertise has been fantastic the facilitators who see so many of these solutions uh and matt theresa and uh and mary have
1:12:09
been great too and the good news is we have a lot of interest in this technology so let me just move on to that but just
1:12:14
want to want to get in a really important thank you uh for this process
1:12:20
so what do what are we doing so we've got a background in digital identity but uh our focus in this sandbox has been

1:12:26
in um in with our solution which is um using privacy preserving technologies
1:12:33
to fight financial crime and particularly aml so the big problem with uh fighting
1:12:39
financial crime is that organizations if they were able to share information in
1:12:44
more granular detail more freely they could actually reduce reduce financial crime
1:12:52
but the problem around data sharing is that the data has to move or it has to be pooled and that brings all sorts of
1:12:57
problems our solution prexa allows institutions to leave the data where it is
1:13:03
at the bank or institution but allow insights or knowledge around that
1:13:08
transaction our individual to be shared between the parties
1:13:13
um without actually disclosing the underlying data so we find that the best way in order to
1:13:19
avoid leakage of data or potential compromising of data is to never move it in the first place
1:13:25
so we use this zero knowledge proof and secure multi-party computation to enable a risk score to be created
1:13:31
while the data stays in place and miguel our cto who likes to call it fancy maths
1:13:36
he can answer questions on this uh later but the great thing is that privacy and confidentiality are fully preserved
1:13:42
so on to the pilot itself so lex tago with their phenomenal capability to analyze
1:13:48
at 10 000 transactions per second were able to look through reams of data and you'll
1:13:53
see the blurry details in the background at the back of this slide is the reams of stuff that they they they
1:13:59
sent to us um and then they assigned a risk score and as they mentioned there was a
1:14:04
particularly uh high ranking uh um gentleman called russell hunter
1:14:10
who seemed to be up to no good in their uh in their in their data set and i'll show you a
1:14:15
demo as to how we we had a look at russell in a moment but the key thing is that we worked with lextego to build a framework to allow
1:14:22

the banks to communicate and this enables enables a lot of time to be saved for banks a
1:14:27
reduction in false positives and a lot of unnecessary sars being filed and ultimately um preventing financial
1:14:34
crime so we we built this uh this framework which asks questions around the payment instruction
1:14:40
data and also around the suitability of the sender so uh you can explore the demos uh it'll be on the website but let me just show
1:14:47
you it uh real quick here so here we we have two banks bank a and bank b
1:14:52
neither party shares in shares the questions to their uh to the answers the answers to the
1:14:58
questions with either party we use a secure multi-party computation to do this but each bank answers
1:15:04
questions about um about the suitability of their account holder and also about the transaction
1:15:09
details so here we see ashley scott has been trying to pay british telecom but in fact this bank account details we
1:15:16
learned from the process actually are associated with this character russell hunter um and both banks
1:15:24
really ask to answer the questions as per the framework and and the the the process is executed
1:15:30
the multi-party competition runs and what comes back is an advisory to say look there's there's
1:15:37
going to be some issues around russell hunter here because uh he has a lot of sars filed he
1:15:44
uh has um and his house his his uh account has been on hold in the past too
1:15:50
so this will come back with um with the with the details that there's an identity
1:15:55
identity and suitability issues around this transaction and further investigation is needed
1:16:02
the good thing then just zipping on here is that the bank a who who was um
1:16:09
who was in fact uh our friend uh ashley's bank they have identified that
1:16:15

there's been a lot of a lot of transactions to this account uh of of of russell hunter with these
1:16:21
account details and it seems like in particular i'm sorry i'm pressing the button here that brenda
1:16:27
core has in fact been very active uh in in transacting with this russell hunter
1:16:34
and it looks like that uh um that she may be an an accomplice to the
1:16:40
fraud that was being perpetrated by by russell hunter so uh let me just uh refresh this excuse me it's just after
1:16:47
of course live demos would would pause but uh what has happened is that um
1:16:53
that brenda and russell have in fact as we've ran our execution on on on the data in the past uh have been
1:17:01
colluding she has knowingly been sending money to to uh to russell it seems
1:17:07
there has in fact been some sars file on her in the past year but it wasn't really as obvious until we
1:17:13
had number one lex tago's great analysis of the of the uh of the of the transaction data
1:17:20
and secondly our ability to find additional information related to to the transaction uh from um
1:17:28
using our secure multi-party computation so in uh in essence really we've found the
1:17:33
the process really fantastic for uh for dealing with um for for learning
1:17:39
for testing our model and and bringing it to life and look forward to the next steps with lex
1:17:44
tago with cinetic solutions and and and plenty of the other uh organizations that explain
1:17:50
expressed interest so welcome your questions and miguel our cto is also here to handle
1:17:55
any more technical ones that may come in thank you very much indeed david that was a really comprehensive overview
1:18:01
we've had a couple of questions coming in um so someone's asked since legally compliance
1:18:06
uh legal compliance responsibility cannot be rolled over how can the data recipient bank feel

1:18:11
comfortable that what is shared is actually valid without seeing the actual data yeah
1:18:17
the um mig do you want to take it or shall i yeah that's that's a very good question
1:18:23
it's around the data governance model in in the communication so typically data governance expands to just within
1:18:30
the bank but in this case a global data governance model is required for the collaboration between
1:18:35
the banks that make sure that the quality of the data contributed to the computation meets uh basic standards so we can think
1:18:42
about audit processes in place that uh you know a certain and make sure
1:18:48
that that quality meets the standards we can also think about the algorithm making some basic checks on the
1:18:55
syntactic um interoperability for the data so to make sure that dates
1:19:00
and passport numbers and some other information meet the the specific requirements for the
1:19:06
for the computation to take place but it's definitely a problem that needs to um to you know involve the two organizations
1:19:13
or multiple organizations in the computation uh to make sure that that quality meets the basic stand-ups
1:19:19
thank you very much and stepping onto that around the kind of uh engagement between banks i mean this solution benefits when more banks are
1:19:26
involved and at a practical level how challenging is it for banks to implement the solution given their challenges around legacy
1:19:32
systems and data quality the uh very good i'm sorry sorry the
1:19:40
the good news is that the banks don't don't don't have to get permission to pool data into a central database which
1:19:46
which is really a big saving and we've designed it to be deployed on site
1:19:51
at the various banks uh thanks to miguel's uh great engineering nick you might like to
1:19:56
follow on yeah it's it's a simple sdk that it's deployed on premises and it just needs
1:20:02
access to the data but that data never leaves the system so it's very easy to to interface it to
1:20:08

existing transaction monitoring systems and legacy systems and the good news is we've got we've got
1:20:14
a a network of banks in switzerland now going ahead with a full uh proof of concept using this share with with real
1:20:21
data um which has taken us a number of years to get but we really feel that this technology is uh
1:20:27
is its time is is now coming wonderful thank you both miguel david
1:20:32
thank you both very much indeed that was a really uh helpful uh uh overview and
1:20:37
thank you very much so moving on to uh to the next
1:20:42
team i am not i have to confess i'm not quite sure how i adequately said this team mpc4aml
1:20:50
um which i think is being led by mary beth and so over to you marie
1:21:00
thank you everyone for your presentations until now i think it was very interesting to hear what everyone is doing
1:21:06
uh especially the presentations from uh brandon and david i think uh what we are
1:21:12
doing is a sort of you could see it as a sort of combination of those two so i'm happy
1:21:17
that they were first um well my name is uh maribet van egmond i'm a researcher at tno
1:21:25
which is the netherlands organization for applied sciences
1:21:31
which is an independent research institute in the netherlands
1:21:36
and we are working on a project that is called mpc for aml so secure multi-party computation for
1:21:43
anti-money laundering and this is a shared research project between tno and
1:21:49
two dutch banks rabobank and abn amro
1:21:54
well what are we doing in this project well we are researching the feasibility of using
1:21:59
secure multi-party computation for anti-money laundering and um secure multiple multi
1:22:06
uh multi-party computation or mpc as i will call it is a cryptographic technique to
1:22:12
jointly analyze sensitive data without sharing it and this technique actually enables a
1:22:18

group of banks to perform analysis on the entire transaction network so the combined transaction network without
1:22:24
having to share their individual transaction data well david already sketched the problem of
1:22:30
data sharing in such a trans transaction network very clearly i think
1:22:37
and what we actually want to do is um do an analysis using this new technique and what we
1:22:44
run into every time is that this is actually a chicken or egg problem because we have this technique that
1:22:50
enables this group of banks to perform this analysis but then the question is what analysis
1:22:56
do you actually want to perform because there's no ready-made aml algorithm
1:23:03
that we can perform because this this possibility has never been there before
1:23:10
um so our starting point was um to
1:23:17
think of an algorithm um that we can that has which has an added value
1:23:23
um of uh where npc has an added value so where where collab collaboration of these banks is
1:23:30
actually uh needed and this is what we call the risk propagation algorithm
1:23:35
and i think when i hear the talk of brandon this is really has the same
1:23:43
idea namely every account gets a risk score which can be based on cash
1:23:48
or high risk geographies or cryptocurrencies or anything and this risk score
1:23:56
is being propagated through the network which means if you look at this picture that if a risky account sends money to
1:24:04
an account that is not considered risky then its risk score
1:24:09
increases and well mpc actually makes it possible to securely
1:24:15
use these risk scores from other banks to update your own scores while keeping your sensitive data so
1:24:21
your own risk scores private [Music] let me go to the experiments so i want
1:24:28
to talk a bit about two experiments today um what we did in sandbox data which is
1:24:34

mainly mathematical analysis of this algorithm and we also performed some experiments
1:24:39
on another data set which is outside of the sandbox but i think for demo purposes it's nice
1:24:45
to show you well in the sandbox data we use the synthetic transaction data so that contains the
1:24:53
sources nation and amount of these transactions which is what we actually need
1:24:58
for risk propagation but actually to actu to validate the
1:25:04
algorithm we need some more additional features such as gas transactions or
1:25:10
money laundering patterns um which were not in this data set unfortunately
1:25:16
so that's why we also looked at the other data set and we mainly focused on mathematical
1:25:21
analysis such as convergence and distribution of the risk amongst
1:25:27
a transaction network unfortunately i don't really have time to talk about that now but here are some
1:25:34
nice pictures um well and it definitely gave gave us some more
1:25:40
insight into um the algorithm that that we came up with
1:25:45
um so let me go to the second experiment so we investigated the effect of this
1:25:51
algorithm on some patterns that were included in this data set
1:25:58
which are mainly getter scatters scattergather and cycles so you could imagine a pattern such as this one but
1:26:05
for the demonstration i want to focus on the so-called gather pattern
1:26:10
so imagine we have five accounts that are distributed amongst three banks then
1:26:17
if the accounts of bank a and bank b have a high risk score for example because of cash transactions
1:26:23
and they all send money to an account in bank c then the account in bank c cannot see this
1:26:31
because they the account or bank c cannot see the risk scores of bank a and bank b but using npc we can
1:26:40

securely send these risk scores from bank a and bank b to bank c and bank c will see that his account is
1:26:47
suspicious without actually knowing the scores of bank a and bank b
1:26:52
so that's what the mpc solution is about for now we just look at the effect of
1:26:58
risk propagation on this pattern without the division on banks
1:27:03
so then it looks something like this we have start situation with these four suspicious
1:27:10
nodes and there's this triangle node that we actually want to catch um but then our our research
1:27:18
question was like what happens if we perform this algorithm well then you see if we do one
1:27:24
iteration you see that the score of the triangle increases a bit and
1:27:29
um if we do two iterations it increases even more and three iterations more um and
1:27:36
then you see uh here you see the same thing again in a small demo
1:27:45
and what is our main observation of this is that it is possible in this case to
1:27:50
detect this triangle account um but you also see that the initial risky nodes
1:27:56
they their score drops but if you look at the scores relatively
1:28:03
then you see that that it's quite even so that means that we we need to add some
1:28:09
some kind of scaling to this algorithm yeah so just to go back to the situation
1:28:14
of the three banks you see that here we
1:28:20
achieve actually what we want if we would do this in a secure way namely that bank c sees that his account
1:28:26
increases in score without actually seeing the scores of the other bank
1:28:33
because they are kept private because of the use of secure multi-party computation
1:28:39
um yeah so that was my story um our conclusion is that this at least
1:28:46
for this pattern this risk propagation seems useful and our next step is to
1:28:52
build a proof of concept um where we implement this algorithm in a privacy surfing wait
1:29:00
yeah that was it thank you very very much indeed murray

1:29:06
beth um we've had a comment through from an attendee uh kind of uh reaffirming the importance
1:29:13
of the question you raised about um uh uh kind of the the the
1:29:22
compliance responsibility and i can i can see someone is is leaping into to answer and engage on on that topic so i would i
1:29:29
would point you to that um as well any other questions coming any questions coming through
1:29:34
for marie beth on her presentation
1:29:39
i'm just double checking the time we do have a couple of minutes if uh if there are any questions coming
1:29:47
through uh from marie beth how do you ensure that the banks use
1:29:54
standardized ratings um i think if um
1:30:01
if i understand correctly this question you you are talking about um ah okay yeah i i think i know what
1:30:07
you mean like um the banks so if one bank says risk score is 0.5 does that mean the
1:30:14
same thing as that another bank says 0.5 um
1:30:20
well i think that has to be discussed uh very um
1:30:27
that has to be agreed on in advance but uh now in my story i think these risk scores are very general
1:30:34
um but in when we want to use this these risk scores will be more specific maybe there will
1:30:40
also be like a factor of risk scores where one is for example about guest transactions and the other one
1:30:46
is about high-risk geography and so so the definition of these risks course
1:30:52
should be more specific than the way i present it now and then hopefully this will be aligned
1:30:58
in the right way but uh it is an issue of course it is uh something we should think about
1:31:04
yeah lovely all right thank you very much that i think has brought us to time thank you
1:31:10

very much indeed for your presentation marie beth thank you i'm going to come now to norblock uh we are
1:31:16
in the in the final run of presentations uh north block is the first of four left to go um and uh we have uh
1:31:24
manos who is leading the team there i believe
1:31:33
hi it's actually sorry sorry simon no worries um good morning everybody um
1:31:39
i'm simon and we're norblock um we're on a journey to uh sorry let me just get our my
1:31:46
screen up um and we're on a journey to redefine kyc through our onboarding and
1:31:54
data sharing uh utilities and so the demo that we're going to be running for you today
1:31:59
is designed to showcase how our fetus kyc data sharing utility which is built on blockchain
1:32:04
can help prevent fraud and scams and allow institutions to be more uh product proactive so the first use
1:32:10
case that we presented uh back in december's demo day is based on utilizing the kyc ecosystem
1:32:19
to both enhance the customer onboarding experience improve the quality of the compliance data that's being captured
1:32:25
and then also um being able to share suspicious transaction data without um sharing proprietary or
1:32:32
sensitive uh competitive data and then still respecting uh privacy regulation so in our first demo
1:32:40
day we looked at how two institutions with the same customer can share the suspicious transaction
1:32:46
data and basically ensure that they can secure customer accounts if
1:32:53
there happens to be a transaction that that's flagged through the ecosystem and so today we
1:33:01
wanted to share an additional uh way to deploy the fetus ecosystem so that
1:33:06
there's a more proactive element to preventing fraud and scams based on our conversations that we had
1:33:11
with mentors and regulators and other participants in the sandbox one of the things that we found is that

1:33:17
the current process of submitting suspicious activity reports uh to the ncaa is siloed and not very
1:33:23
conducive to proactively preventing fraud and scams so what i wanted to demo for you today
1:33:28
is how to utilize the ecosystem to submit and share the suspicious activity reports with the nca
1:33:35
and across institutions that have a relationship with that entity or customer so that being said
1:33:42
let's take a look at how that works as we're seeing here on the portal we can see the company details such as
1:33:49
the ubos the kyc status of this customer and any other relevant information
1:33:55
and so um once we go through this process we'll select what suspicious transaction
1:34:01
this particular customer has that is of concern and once we do this
1:34:06
we'll uh in a production environment we can submit documentation et cetera and report all of this into the nca so
1:34:15
that uh the nca can take the appropriate action when they're reviewing the
1:34:20
suspicious activity report all of this is customizable in the platform for the needs of the individual
1:34:26
institutions and also the ecosystem as a whole so once we submit the report
1:34:32
we'll go into our dashboard as the regulator so in this case the nca and we'll see that the suspicious
1:34:39
activity report has come through and again as mentioned in a production environment here we'll see
1:34:44
all of the documentation or data that is relevant to investigating whether this
1:34:49
is an actual valid transaction report or activity report and based on this
1:34:56
the nca can make a decision whether to confirm that this is indeed a suspicious transaction or or
1:35:04
kind of escalate or do whatever it needs to do so once this is confirmed
1:35:09
if we are anglia bank which is uh also part of this ecosystem and actually
1:35:15

shares a really shares a relationship with this customer um that the sar has been filed
1:35:20
against um we can go in and see that there's a report that comes through now all of the
1:35:26
information that's shown here is information that already has been gathered
1:35:31
on the ecosystem and is not shared so nothing proprietary no no information around the client
1:35:38
relationship or what bank reported the sar is shared with um banks on the ecosystems to protect
1:35:45
the privacy and the proprietary information but essentially here what we'll see
1:35:50
is that there's a remark that several linked cash transactions have been linked to this account and
1:35:59
or this entity and that essentially this allows anglia bank to make a decision on how to
1:36:04
secure this customer account and ensure that it's preventing any further fraud and scams from taking
1:36:10
place so all of these workflows again are totally customizable
1:36:16
and ensure the privacy of all parties involved and the benefit here is that the
1:36:22
blockchain-based ecosystem means that there's an immutable record um ensuring the accuracy of reporting
1:36:28
and enabling um auditing from regulators and parties that are um vested in in this ecosystem so that's
1:36:35
our demo for how to prevent fraud and scams with the fetus ecosystem and more than happy to answer any questions or
1:36:42
discuss anything further thank you very much indeed and
1:36:48
just because we had this slight technical glitch moving between slides we'll we'll give you that time back simon so
1:36:53
we won't we won't cut into any any q a is any q a coming back coming through from anybody
1:36:59
any questions burning questions uh for the team at norblock okay so i
1:37:06
mean i suppose a kind of a really practical one what do next steps look like for you simon

1:37:11
uh next steps um great question so i think for us the next steps um are to kind of uh get feedback around
1:37:18
the utility and the i guess what where we might see some gaps in the in
1:37:24
the needs from the various stakeholders here whether that's the regulators that would be involved or institutions
1:37:31
and really understand how we can um further build out functionality to support um
1:37:37
those needs lovely thank you and yeah as part of that i
1:37:42
mean you you really imagining that those conversations will start to happen with banks in terms of an implementation
1:37:48
pathway yeah i think for us um we're open to having conversations with banks and
1:37:54
regulators and you know based on our existing production ecosystem that's live in the
1:37:59
uae um we've we've worked with both parties to ensure that um the the solution that is deployed
1:38:06
is um deployed easily across all of those uh partners and done in an equitable way
1:38:12
so that um there's no one party doesn't have a more of a vested interest than another
1:38:18
perfect thank you very much indeed simon i can't see any other questions through so i'm going to wrap us up there with with our thanks
1:38:26
and move on to team futures ravi and andrew i think you are presenting
1:38:34
on behalf of team futures
1:38:40
uh good morning everyone my name's ravi uh andrew should be on the call as well
1:38:46
uh we're from team futures uh at bae uh we'll get straight into it uh because
1:38:52
we don't obviously have much time just quick intro to futures so we are the in-house innovation team with nba
1:38:58
systems uh creating new strategic capabilities for our customers and um it's kind of to that end that we
1:39:05
wanted to get involved with the sandbox so i quickly went through the first bit uh
1:39:12
we were dealing with use case 1.3 which was about looking for deployment of technology to detect

1:39:18
patterns or other indicators of consumer behavior our approach to doing this was to trial
1:39:24
a new ba systems develop technology to explore how risk could and should be flagged in
1:39:31
real time some of the key features that we wanted to test with our new technology on
1:39:38
sandbox were looking at those kind of real-time aspects so we were looking to test out
1:39:44
neil time near real-time uh incorporation of input data and analysis
1:39:49
uh and so near real-time incorporation of input data and the analysis on the impact on resolution and risking uh
1:39:57
crucially without the need for a batch rebuild we know that's one of the kind of uh the gold standard of analysis
1:40:04
is doing a big batch build to get some really significant complex analysis out we wanted to see if we could bring some
1:40:10
of those capabilities to uh real time we wanted to test out the ability to define groups of interest
1:40:16
defined by a flexible set of characteristics and features that we or our customers decide are important
1:40:23
and extract those results in near real time as well we wanted to look at whether or not we
1:40:28
could persist those groups and then receive proactive notifications so that operational users could actually
1:40:34
do something with that information and finally we wanted to test out whether those groups of interest
1:40:41
could themselves be grouped into networks to try and identify wider scale and organized attacks again
1:40:47
in real time so quick uh overview of the progress
1:40:53
that we made so these are kind of the things that we wanted to try out number one was deconstruct broad typologies into
1:41:00
identifiable behaviors we've done that we wanted to configure
1:41:05
our engine to identify these behaviors in real time we've done that
1:41:10

we wanted to group these instances of identified behaviors into networks in real time
1:41:16
we've done that and finally we wanted to close the loop by using our findings to trigger friction and explain our
1:41:22
findings to the end user and that's where we've started but we haven't quite finished
1:41:28
um so i haven't gone through all that uh very quickly i want to take you
1:41:33
through a quick demo video so i'll just talk over this as it goes through this is a
1:41:39
kind of mvp uh user interface that we built for the purposes of the sandbox i'll
1:41:45
just start talking you through it as it comes up what you can see here is the alert screen and in a second what you'll start
1:41:52
to see is alerts populating into here these alerts are actually being generated in real time so as data has
1:41:57
been fired in under those that are interesting get popped up on here
1:42:03
and you can start to see that this uh this alert window is filling up so this really is
1:42:08
happening in real time behind the scenes in just a moment let's push it forward
1:42:13
actually we select one of those to have a look at what's in there what we can see here is
1:42:18
an entity that's been selected along with the transactions around it that are interesting and we can see just down the left hand
1:42:25
side here that in this network graph view uh we've created what we
1:42:30
termed a group of interest and we have identified some group attributes so the total incoming the total outgoing
1:42:37
as well as per edge different attributes as well so actually all those attributes kind of
1:42:42
carried through into this visualization interface fine uh i'll just go back a second
1:42:49
clicking on a different attribute you can see actually different different properties come up one of which is that the cash the
1:42:55
channel has changed cash the amount has changed uh you can't quite make out on this uh on this video but these
1:43:02
arrows are directional so what you can see here in total is where did the money come in from a bbl
1:43:08

loan in this case and where did the money go out to lots of different transactions kind of capturing all of
1:43:14
that financial flow the next thing we wanted to do then was to
1:43:19
group that up into a network of associated entities and additional activities and
1:43:26
again we did that in real time so what you're seeing here is a network that's been constructed in
1:43:31
real time based on alerting code alerted characteristics so we defined some risk rules which
1:43:38
generated some alert which subsequently led to this network graph being built
1:43:43
this is the kind of capability that has historically been kind of restricted to
1:43:48
batch batch analysis and we're starting to pronounce much closer to
1:43:54
real time now and uh i will just run it through a little bit
1:43:59
because you'll see a couple of network graphs pop up here this like smaller one here as well which
1:44:05
is a little bit easier to follow but what we can see on this one is that we've got
1:44:11
a business here a business here and a business here and they're all connected by a couple of common individuals so
1:44:17
that's the kind of network typology that's pretty common that we expect to see what we do like i said our traditional analysis
1:44:25
that was a really quick run through everything you've just seen now was pretty much built for the sandbox so from our perspective
1:44:31
what's been really exciting is that from a technology perspective which is kind of how we've taken a focus on this
1:44:37
we've managed to do quite a lot of stuff during the period of the sandbox we've extended our data interest framework to accommodate
1:44:43
new data we've never seen before we've added a whole bunch of new features to our core analytics engine to generate
1:44:49
the insights that you've just seen we've validated that our flexible risking framework can actually identify
1:44:55
the things that uh that are required and that was all stuff that again we hadn't seen before
1:45:00

we didn't have to extend our framework too much actually to do that we developed a brand new user interface
1:45:05
uh an mvp one because uh actually we needed a we realized that we needed to see how we needed a
1:45:11
different way of interacting with the data to how we previously previously been doing so and finally
1:45:17
probably most importantly for us we demonstrated that alerts can be dynamically raised in real time
1:45:23
as new things come into the system so what next um
1:45:31
we've had really good fun doing working on the sandbox and it's really helped us kind of iterate our technology quite a lot um
1:45:38
we're now looking for partners to experiment in an operational context clearly synthesized data is brilliant
1:45:44
and it takes you up to a particular point but there is a point at which you want to get some real feedback from real users
1:45:49
um so that's kind of where we are we'd like to gather feedback about how well our approach of bringing stuff closer to
1:45:55
real time solves our partners problems interestingly the third the third aspect
1:46:00
of this we want to explore the impact of real-time interventions on business processes
1:46:06
if your alert screen is filling up literally second by second what does that mean for your
1:46:12
for your fraud intervention processes and practices to establish or to kind of flesh that
1:46:18
out a bit we've actually commissioned some internal research on this already because we think it's a pretty substantial question
1:46:23
and we'll have quite a lot of impact when you get these slides if you're interested just click on the box at the
1:46:29
bottom and you'll get an email pop-up which you can send over to us
1:46:35
and i will stop talking there thank you very much indeed that final
1:46:40
point you raised is a really interesting one isn't it it's around um you know behavior change and and actually how that interface will
1:46:46

work in practice with people uh and uh and that engagement so i think it's a really interesting piece of research
1:46:53
that you have commissioned and i'm sure there'll be lots of interest in it um a couple of uh
1:46:58
questions coming through could you expand on the benefits of real-time monitoring versus batch monitoring and
1:47:05
you might mentioned adding friction again could you give us some examples of what that might look like
1:47:12
i'm gonna ask andrew to step in on the first part of the question uh and actually the second question is
1:47:17
i'll show you that yet fine yeah so i mean i think uh for me
1:47:23
the the benefits of the real-time capability are about uh being able to take into account
1:47:29
what's just happened for them subsequent events so i guess if um in some of the traditional systems even
1:47:36
if uh say an application for whether it's a loan or for an insurance policy or something like that
1:47:41
can be uh can be scored against uh a batch bill system the data about
1:47:48
that thing often isn't incorporated until the next batch runs so um that means that if if someone is
1:47:55
testing the waters by putting in a number of different claims you often can't pull that picture together until later
1:48:00
whereas in this world we can do that we also have some uh it means that we
1:48:07
can also offer other use cases for things like when um and i guess this speaks a bit to the
1:48:13
intervention question and that data is immediately available for
1:48:18
people like uh call handlers so if someone's uh called up about something that they've just done we've already
1:48:25
assessed it against risk or we can at least see where it sits in the network and so they can perhaps change the
1:48:32
routing of that customer appropriately as to you know whether it's a simple thing that they can say yes to straight away uh or whether
1:48:39

it's something that requires further investigation because there's risk associated with it so um it really for me at least in
1:48:47
i guess in the in this sort of financial crime context um it it yeah it's all about being able
1:48:52
to have that up-to-date picture we've got some other use cases that we're working
1:48:58
on that are much more in the sort of law enforcement space and there obviously having that real-time incorporation of data
1:49:04
is uh you know important in terms of sort of interventions there and risk scoring
1:49:09
uh risk scoring events as they happen i want to add to that um i think it's
1:49:16
it's relatively well established to to assess transactions in isolation in
1:49:21
real time it's pretty novel to contextualize that as fully as we're proposing to do here
1:49:27
to get a really rounded view of the risk and i guess bringing that back to a real life situation
1:49:34
we're talking about vulnerable customers at the start and i'm going to hypothesize here an elderly vulnerable
1:49:40
customer will still go to a bank badge imagine having the capability to
1:49:46
process that elderly customers transactions and get it whipped around the entire technical system within three or four
1:49:52
seconds so that if something is of concern you can catch them before they've walked to
1:49:58
the front door and you can say actually do you mind if we have a chat about what you've what you've just done because actually
1:50:05
some something here doesn't look right and i know that's a particularly i know that's quite an emotive use case but i
1:50:10
also know that that's something that uk finance are interested in with the take5 campaign about trying to find people who had been coerced into
1:50:18
particular financial transactions so if you've got the whole system working behind you
1:50:23
so that you can catch them before they walk out of the branch that's pretty powerful
1:50:31
indeed thank you um i think that the point you raise and really bringing it back to kind of you
1:50:37

know who are we solving on behalf of and where where where do those where does the harm sit
1:50:42
i think is a really important uh reminder for us all thank you very much indeed uh
1:50:49
team futures uh just two more teams to go and so i would like to invite uh team
1:50:56
one span uh to step forward and i think sharon lee and professor stephen murdock
1:51:01
are taking the floor for team one span
1:51:09
okay thank you so um hello i'm sharon i'm a researcher um at one span um so our project is
1:51:16
about building up the adaptive learning algorithms for fraud detections
1:51:25
so um first of all i would like to talk about um our progress so the objective is to build and test
1:51:31
some additive learning algorithms using the fca digital sandbox in particular we are interested in the
1:51:37
uh device data and transactions banking data our data scientists including myself have
1:51:43
analyzed the data set we have implemented tested and compared several machine learning algorithms are some are static
1:51:49
and some are effective um we uh did improve the first phase in
1:51:54
the review and the reject categories we also have our internal floor consultants
1:52:00
are involved in the project he reviewed the dss and brought in some matter expertise to support our work
1:52:08
so um i would like to um use the uh device data um to to
1:52:14
explain the challenges that we have in the domain of fraud detection in digital banking so um in the data set
1:52:23
we can see there are 35 columns the number of transactions
1:52:28
is 5 million and within that 5 million data points there are only 2 997 quadrant transactions
1:52:35
the fraud rate is 0.06 as we can see it is a very extremely
1:52:41
imbalanced data set on the uh right hand side we can see the um the details of the um of the uh
1:52:49

fortran transactions uh scam is the most popular one and then we can see red and depending on the human expert
1:52:58
some people will put fraud in the in the labels um and we also see quite a lot of fraud
1:53:04
are the first party fraud so um the first question that uh came up is do we actually have
1:53:11
enough good features in the data set so that we can separate two crosses um
1:53:16
as i've also mentioned it's an extremely imbalanced data set so it is
1:53:21
quite challenging for the machine learning algorithm development another limitation about data set is um
1:53:28
many datasets they are not are interlinkable and and
1:53:33
it means that we can't actually uh leverage the alternative dataset so if we believe
1:53:40
the uh fca synthetic data is a good representation of the real world then it
1:53:46
will give us some idea on the performance of the fraud detection system nowadays
1:53:52
so um the frost detection system we are uh did classify the all the transactions
1:53:58
into three categories the path reveal and reject within the reject category it means that the system
1:54:05
will reject the transaction directly and there are only 15 quadrant
1:54:11
transactions out of 540. for the reveal category it
1:54:16
means that we require a human expert to view the data point one by one
1:54:22
um within the 33 000 data points there are only 318
1:54:27
quadrant transactions and in the past category actually it contained most of the
1:54:33
quadrant data points which is in total uh 2665.
1:54:38
um from this um statistic um we we learned that the fraud
1:54:44
detection system is doing something the first way in the reject and review
1:54:49
categories are high much higher than the average however most frauds are still in the
1:54:55
past category and it can pass through the system um here i would like to show the
1:55:01

normalized histogram of the quadrant transaction versus the general insight transactions
1:55:07
there are two columns in the dataset called the positive score which are divided by the human rules
1:55:13
and another one called digital truss id trust score which is um divided by some population
1:55:21
matching algorithms and it will tell you on how reliable is that um digital
1:55:26
ide so for the foreign transactions which is again is 0.06 of the population
1:55:34
you can see normalized um histogram distribution is like this and this is the gendering um data point
1:55:42
um normalized histogram and here is the overlapped um histogram and as you can see
1:55:50
the foreground transactions perform fairly well in the digital id trust score some of
1:55:57
them are very well very good um while the uh the policy score uh looks like um
1:56:04
more effective uh and uh many foreign transactions have lower process score however if we take into account
1:56:12
on the uh very small number of fortune transactions it is still very challenging to um like
1:56:19
separate the filtering transactions and degenerate transactions without having a very high false
1:56:25
acceptance rate so um before we look into data set we
1:56:30
hope that we can have some nice um engineered features to separate two
1:56:37
classes so that we can find a clear or nice decision boundary however the
1:56:43
reality is we found that our two classes are heavily overlapped
1:56:48
with some reasons first of all humans do change behaviors
1:56:53
and more importantly many frauds are conducted by trusted device for example the app fraud
1:57:01
so um for the next step what we would like to do is do more research and experiment to improve our existing
1:57:09
adductive algorithms we would also want to leverage the machine learning
1:57:14

algorithms to assist experts in the development of groups more importantly i personally believe
1:57:21
that we do need to design new features for fraud detection system just like
1:57:26
what corsair is doing but we need to do something much more it is also important for us to consider
1:57:35
the combination of different data sets which can help us to defend new type of thoughts so um
1:57:43
that's it and any question i welcome
1:57:50
thank you very much indeed sharon um any questions from the group coming through from our audience today
1:57:59
okay we've had one coming through does this type of solution require the customer to have specific devices such
1:58:04
as a smartphone or laptop and will it support customers segments who particularly use telephony so
1:58:11
i mean that's a that's very pertinent for the kind of older and more vulnerable segments i think
1:58:16
um we we do not have um the information in the uh data set
1:58:23
regarding the segment um or the type of the customer um we in the data set or we do see um the
1:58:30
transaction data from different devices so um the uh what we have done is try to
1:58:36
get uh the uh first 20 of the data to learn some global parameters and uh try
1:58:44
to uh use the parameters to set up the verso and run on the uh remaining data set uh we do find that uh uh
1:58:52
using this kind of adjusted learning algorithm can help us to um categorize more fraud into
1:58:59
the reject and review category but from what you can see from the uh data we do
1:59:05
have limited human power our bands doesn't like to have too many alerts
1:59:10
and they don't want to handle the alerts that they can handle so um there are really restrictions on
1:59:17
how many um data points we can put into the uh we jet category and the review
1:59:22

category and when we develop the um algorithm we need to check that into account so that it is realistic to
1:59:27
to be implemented by bands thank you and i think that touched upon a piece of research that ravi was mentioning about
1:59:33
earlier wasn't it about understanding what what what uh going to do with the with the proliferation of alerts coming
1:59:40
through thank you very much sharon i can't see any other questions um coming through from the team so
1:59:47
unless you had any kind of closing remarks um i will thank you and the team very
1:59:52
much in indeed and come to our final presentation of of this demo
2:00:00
uh team trust stamp and it2 fraud signals being led by
2:00:06
adam adam ridgeway adam are you ready to go
2:00:14
my name's adam ridgeway and this is trust dance it2 fraud signal sharing so we've actually
2:00:19
partnered alongside uh cfas uh lloyd's banking group and one banks for the delivery of this
2:00:24
and then on the line as well we've got yasek who is our technical project manager
2:00:33
okay so uh cases of identity fraud rose by 18 in 2019 with a 32 increase since 2015
2:00:41
and this is poured from the the cfas fraudscape report so 87 of this occurred uh via online
2:00:47
channels uh and my guess would be that uh poster pandemic this this number is going to have a huge increase
2:00:54
so we've got a unique solution to this problem this problem and that's based around um detecting
2:01:01
the the fraud for the biometric so um the one variable that the fraudster cannot change
2:01:07
is is their face or their biometric so what we can do is we convert the biometric template which is
2:01:12
typically captured during the customer onboarding or during enrollment and we convert this into our proprietary
2:01:18
it2 our irreversibly transformed identity token and what happens and by doing this what

2:01:25
happens is it enhances both the security and the privacy in that we can then discard the original
2:01:30
biometric that's been templated we can discard the original biometric template and this
2:01:36
then allows us to authenticate users without the risk of biometric fast um and additionally to this and
2:01:41
what we've done for this project is we're then able to probabilistically match or compare these tokens
2:01:46
as a means of identifying fraud so we can match verify and do that buzz and deduplicate
2:01:52
against these tokens so what we've done is we've created a
2:01:57
watch list of it2's and this essentially acts as a biometric safeguard
2:02:03
that um that denies access or acts as a flag if there's been a match or when there's
2:02:08
been a match and a way of doing this is you could have multiple watch lists made up of
2:02:15
known fraudsters or you could have watch lists of enrolled customers and where there's a match this would uh
2:02:22
be as a signal for identity fraud highlighting which could highlight velocity attacks over a very short
2:02:28
period of time so um as you can see here the the fraudster
2:02:34
uh the fraudsters data or their it2 can then be shared across
2:02:39
multiple organizations um without the risk of breaching gdpr or any data privacy regulations and
2:02:46
this is because once we've tokenized that data it's no longer deemed sensitive information and then we can do this in real time
2:02:55
which allows that ability to create a shared biometric fraud network additionally to this we can query these
2:03:01
tokens using zero knowledge proofs to extract sort binary yes or no answers
2:03:08
and uh what we originally intended to do was we were going to use the some of a
2:03:14
sample of the live cfas data um but we run into some infosec issues where we were unable to to do this so

2:03:21
instead we've we've replicated this and we've used images that we've collected internally
2:03:26
alongside sort of driving license and passport documents so what we've done is we've had 30
2:03:31
images of 15 real people uh and 15 of these were then used to make up that watch list
2:03:37
that you can see on the top right there and that's to to replicate the the cfast database of
2:03:44
uh images associated with fraud and then on the and then additionally to that we've got the the fraudsters in the
2:03:51
top left um but that is those 15 images essentially replicate that bank
2:03:57
enrollment process and additionally to that we've then got 13 images of 13 real people
2:04:03
40 us driver license images of 40 people and then 18 uk driver license images and
2:04:10
passports of 18 people so what we've done there is we've we've used the driving license and passport images to
2:04:17
replicate um images of real life a real uh of photos taken of real life ids
2:04:22
where there might be differences in the lighting or the image quality just to make sure that they're not
2:04:29
they're not perfect so we've got a total of 101 images of 86 people
2:04:34
and the expected results would have been that we would match the 15 images of the the
2:04:41
bank enrolled customers uh with the watch list and then we would have had 71 images passing
2:04:46
as genuine genuine or non-fraudulent users and that's exactly what we saw so this
2:04:53
this shows the results of our test here we've set the the score value there at 0.6
2:04:59
um and any any of the uh anything that that match below that 0.6
2:05:04
would indicate um a match so what we've seen here is we've got 15
2:05:10
unmated pairs and third of making up the 30 images of 15 people and then we've also seen
2:05:16

the unmated pairs and as you can see here we've got 36 images of unmated pairs which totals uh 72 people so what we do have is that
2:05:24
additional match but this is expected um as we've got an uneven nominated pair
2:05:32
so this is exactly this is consistent with the expected results highlighted before and really shows the power of this of
2:05:38
the it2 token so as a way of next steps uh
2:05:46
that we were limited with this test that it was a very small data set in the end so what
2:05:51
we would like to do is use a much larger data set and prove our scalability um
2:05:57
additionally to that what if we could revisit what we intended to do and use the live data from the the cfas database or a sample
2:06:03
of that then that be that would be ideal and really what we'd like to do is use
2:06:08
uh multiple watch lists for uh to highlight a velocity attack across organizations so
2:06:14
as you can see here we've got this this little image um the way we would like to do it is we'd have
2:06:20
three separate watch lists where we've got a 92 associate with fraud a temporary it2 database and the ic2
2:06:27
master database and that's that enrolling customer goes through he would then be added so he or she would
2:06:32
then be added to the temporary velocity database and if there was a match over a set period of
2:06:37
time this would highlight a velocity attack
2:06:48
and uh just to highlight a few use cases so uh identity fraud kyc aml so we've actually
2:06:54
been uh we've run a pilot where we've done used our deduplication technology against the
2:07:00
pepsi sanctions watch list just to clean up any inconsistencies or false positives in that data
2:07:05
and then as part of that frictionless or onboarding piece as well and then if anybody wants to uh reach
2:07:12
out or have any additional questions about this feel free to contact me on uh that email address in the bottom left
2:07:18

corner a ridgeway at trustar ai thank you very much thank you adam
2:07:23
um that was really interesting thank you very much indeed i've had a couple of questions through uh with uh biometric you'll be able to
2:07:30
see this as well in the in the chat uh with biometric what if fraudster opens the account uh the
2:07:36
biometric would be theirs um any insights on that
2:07:44
so if the if the fraudster was open to to open the account um so typically what we'd catch is any
2:07:50
any mismatch in in pii information so say for example
2:07:55
um the fraudster was using a fake id with his real image on there as well as his biometric when he would
2:08:02
enroll or he would enroll and regenerate that it2 token um if they've enrolled previously or
2:08:08
they're part of that database what we would then see is that there's a match in the biometric but a mismatch in that
2:08:14
pii information okay and that would flag that as a potential
2:08:20
foreign um and then um we've had a kind of a follow-up question around
2:08:26
kind of the known challenges with with with facial recognition and
2:08:31
particularly those with darker skin tones um uh and the kind of the
2:08:37
the the the challenges around structural inequality what specific criteria are you using to
2:08:43
determine the point at which you'll be legally and morally legitimate to put faith in the ability to use
2:08:49
uh watchlist pictures so uh we train our ai on
2:08:56
various data sets so we've got a rounded um and diverse collection of data so
2:09:02
from from that uh we're fairly happy that what what we're doing is doesn't
2:09:08
have any bias in it um and then sorry what was the the other part of that question so
2:09:15
um at what point would you you know which quest what criteria are you using to kind of assess at what
2:09:21

point it will be kind of uh morally legitimate or legally legitimate to put faith in the ability
2:09:27
to use watchlist pictures okay uh i'll um i'll pass you over to yasic to answer
2:09:32
that second part of that question um so i will actually uh add something to the first part first
2:09:37
um so it's just and we are very committed to um assessing the impact of
2:09:44
racial bias on biometrics and one of our research projects actually proves that
2:09:50
our technology does not have
2:09:56
a significant racial impact but that's something that we can pursue offline if that's something that
2:10:02
you're interested in on this note we do have um some other project that we are pursuing um with
2:10:09
other biometric modalities which includes uh
2:10:15
fingerprint palm uh and um voice biometrics um and we
2:10:22
also have another project where we are pursuing uh breaking vendor login so in case you are
2:10:27
not comfortable with using facial biometrics you could switch to using for example
2:10:34
fingerprinted data and you could potentially use it across vendors so you would have one fingerprint vendor and the second
2:10:40
fingerprint vendor where you could compare these two um so we do have
2:10:46
multiple projects that deal with this specific issue um on this note there are specific
2:10:52
things that we are doing inside the watch list which actually account for the fact that
2:10:58
there is a high chance that people not only with darker skin could potentially match with each other so we
2:11:04
do have assessment projects that are currently ongoing as well which are supposed to set the threshold
2:11:12
um at that specific level which will account for this um which is
2:11:18
something that we've been doing consistently since the beginning of the watch list just improving the biometric solution
2:11:23
behind it thank you um and a a follow-up question

2:11:29
and if the token token is still linked to identical data identifiable data sorry within the system
2:11:36
um doesn't the token then remain personal data under the uh under gdpr
2:11:45
do you want to answer this one yati oh sure so there are two things that we do here so
2:11:50
first of all the token itself so the id to token can contain
2:11:56
um pivot points to external databases um so for example in order to have just
2:12:03
unposted you do not amend the token with personality identifiable information um what we do
2:12:10
instead is we point to external databases so for example your database becomes the single source of knowledge about this
2:12:16
person since that person is your customer and that allows also for sharing between
2:12:21
organizations we do have other components that can be very useful in terms of gdpr compliance
2:12:28
which for example involve tokenizing pii basically we convert
2:12:34
pii to vectors which can be compared and instead of sharing pure pii you're basically uh performing
2:12:41
zero knowledge proofs across organizations so the answer to is the customer's name
2:12:47
yatsek is no longer is the customer's name yatsek it's um a comparison of two vectors and it
2:12:54
allows us to make judgments without the need of
2:12:59
sharing the data and an unencrypted format so i hope that answers your question
2:13:06
thank you very much um that brings us to time there are a few additional questions uh in the sidebar
2:13:12
so perhaps i can ask you guys to take a look um and pick those up
2:13:17
um that brings us uh to the end of uh this um demo uh session today and
2:13:24
it brings us to the end of the three demos the showcases that we've had across this week
2:13:29
marking the marking the end of the pilot and i'm sure um you will join me

2:13:35
in uh commending the teams for all the work that they have done um over over the past uh 10 weeks
2:13:43
and for the the time they have taken to really thoughtfully and articulately present to us
2:13:48
their their solutions and their and their and their progress um this morning it has been really rich
2:13:54
and and insightful and really in its in its entirety with with vulnerability and sme lending really
2:14:00
starts to show um the art possible um with the digital sandbox um and so my thanks to all the teams my
2:14:07
thanks as well to the fca and and uh city of london teams for all the work they have done throughout
2:14:14
supporting the teams managing these sessions and bringing it all to life and and showcasing the the range of
2:14:21
activity it's a huge amount of work that goes on behind the scenes so my uh deep thanks to them and to all the
2:14:27
mentors as well who have really engaged we've had the teams across today um provide their
2:14:33
shout outs to a few of their mentors who have really helped to kind of shape uh sense check critique
2:14:40
and challenge along the way and we've also seen some fantastic collaboration and participation across the different
2:14:45
teams which is something exactly that we were hoping to to see and start to start to develop as part of
2:14:53
this process so my thanks to all the mentors my thanks to our advisory panel as well
2:14:59
who have um been there from the start in terms of assessing applications all the way through to supporting the
2:15:05
teams and throughout their process all um all the videos from today and
2:15:10
indeed from all the sessions will be available on the team showcase pages of the digital um sandbox
2:15:17
um uh pilot web pages so please do go and check them
2:15:22

out and please do go and share them with colleagues who haven't necessarily been able to participate today or
2:15:27
or across the other sessions in the week and as i've mentioned we have been evaluating this as we go along and this
2:15:33
has been a really important part of the process to really inform the next steps that we
2:15:39
that we will wish to take uh with the digital sandbox so um uh watch this space um
2:15:45
thank you all very much indeed for your participation it has been a thoroughly uh
2:15:52
enjoyable and hugely insightful process and we are grateful for each and every
2:15:57
one of you who have participated and shaped and helped develop it along the way um so with my thanks from the fca team
2:16:04
and the city of london team my thanks to all the teams who participated today um thank you very much indeed um and
2:16:11
uh do keep engaged with the digital sandbox uh pilot web pages and continue to share
2:16:17
your feedback thank you all very much
2:16:25
indeed