

Improving SME access to finance – Digital Sandbox Pilot demonstrations Transcript

0:04

all right good morning everyone my name is simon horner i'm the innovation director at the city of london corporation and welcome to this

0:11

uh digital sandbox pilot showcase day um i think the most important thing i want to say

0:17

is if any of you do have any questions about the pilot about the sandbox itself the architecture the infrastructure the

0:23

participants the team the process um just get in touch we'll cover

0:29

as much as we can today in particular we're obviously focusing on on the sme user case but if you would

0:35

like more information about the broader sandbox pilot and also what we're going to do next um

0:40

after this um yeah do reach out so i'm from the city of allen corporation partnering with the fca

0:46

on this initiative um sca colleagues will also be on hand and to provide more information um next

0:54

slide please teresa thank you yeah so just the background um we

1:02

accepted 28 teams across the whole of the three use cases um they are obviously

1:08

sma finance which we're talking about today um vulnerable consumers and then also

1:14

fraud and scams and there'll be there'll be other demo days at the showcase days this week

1:19

um i guess the sme use case day was particularly interesting for us um because of the background

1:27

focusing on covid as this pilot was the huge challenges that smes have faced in

1:32

accessing finance uh the huge issues with getting lending where it's needed fast

1:38

using digital channels which you know either lender or participant were less familiar with pre-crisis and

1:45

seeing what we can all collectively do to try and develop those channels improve solutions raise awareness of

1:53

what lending solutions are out there and that's really what what much of this pilot has
has been about so we have 28 teams we're

2:00

going to talk about six these tip six solutions uh in the sme finance

2:06

user case today um pilot officially closed on the 5th of feb so the teams have had 10
weeks to

2:13

work on their solutions and today is really all about showcasing how far they've come

2:20

on on the sme finance solutions next slide please teresa

2:26

so the purpose of the pilot you've been thinking about digital fire uh digital sandbox
concept for a while

2:33

um i mean really it's to try and create the opportunity for firms to test solutions

2:38

in particular to try and create new synthetic data sets to test and scale solutions

2:45

but also i guess beyond that building a sophisticated development environment uh
with digital tools but also

2:51

opportunities for collaboration with peers and also mentors

2:56

from across the financial services system so that not only do they have access to
data to develop products but

3:03

you know a whole suite of support um to help them you know build their business
with of

3:09

course the sva on hand providing that regulatory angle as well

3:14

next slide please theresa um so for those of you don't know these are the key
features um

3:21

of the sandbox i think data is obviously one we've we've talked about and we'll talk
about more today

3:27

um you can see the other the other features uh on the slide i guess in particular for
us at the

3:33

corporation we're we're interested in the observation deck because really it's about

3:38

trying to raise the profile raise awareness of some of these solutions give them as
much kind of air time as

3:45

possible you know whether the audience is the broader government whether the
audience is the business community

3:52

they get more understanding more visibility of these solutions uh mentoring from maybe incumbent

3:57

financial institutions who you know stand to benefit or stand to partner on some of these solutions and

4:03

use cases and then even maybe beyond that scaling thinking about access to investment

4:08

opportunities to bring some of these ideas to market next slide please

4:19

so where we are now so we've had over 800 registered users so that's you know including the 28 teams but

4:26

beyond that huge huge and growing ecosystem of other interested participants

4:32

observers partners mentors investors agencies etc um all the social media

4:39

chats are on their uh social media stats are on there that's the one i'd highlight um is the the collaborative element of

4:46

the sandbox it gives the opportunity for teams to work um outside of their own individual businesses so 100 plus chat

4:53

channels created to foster that collaboration and it's that sense of digital community that we've

4:58

been trying to build beyond the actual digital tools the features of the sandbox um

5:06

next slide please teresa thank you um so turning to the content today so most

5:11

importantly um hearing from the businesses that have been developing solutions it's an interactive session so each team

5:18

is going to have six minutes and then there'll be a short sharp q a those of you want to know more about the product the process

5:25

the next steps etc uh so do post your questions in the chat box and we'll try and get to

5:30

them all if we can uh and we'll we'll read them out once the teams are finished presenting

5:36

um so without further ado i'll ask uh david from fluent to kick things off

5:47

okay so yes i'm david i'm the ceo and co-founder of a company called fluence

5:52

and we are ai implementation specialists and we have a particular focus on applied

5:58

linguistics and the objective in any projects that we work in is to deliver either time savings or

6:05

combination of time savings and performance improvements by embedding document understanding ai into complex business processes that

6:12

involve documentation and there are two particular components of the ai that are quite exciting so it

6:19

can process billions of words of any content and extrapolate patterns indicative of

6:24

quality opportunity inconsistency or risk relative to outcomes that are

6:29

interesting to the company and the other benefit is it can be trained to actually predict outcomes on

6:35

new documentation based on for in this context historic applications and patterns of reporting so in this context

6:42

we're trying to risk profile companies so the objective is to understand

6:48

lending data uh we had a couple of false starts because the primary focus of the um digital sandbox is on synthetic data

6:56

which doesn't necessarily include commentary or documentation but with the help of ollie at digital sandbox we were

7:03

able to end up analyzing uh consumer complaints data instead and what we're trying to do is to

7:08

understand the macro level behavior of the industry what are the lending priorities and red flags within the

7:14

industry what are the bottlenecks in decision making and are there any industry or time related trends in

7:20

the documentation so the data that we actually ended up analyzing

7:25

for the digital sound box was the consumer complaints data which came from the consumer financial

7:31

protection bureau and this is a map of a 108 000 complaints that they have handled

7:40

with that contained documentation that contained actual text within within the complaint and the color coding is us putting in

7:47

for example this is harassment versus investigation and we can start to extrapolate patterns in behavior

7:54

within the large data sets so the way that we can actually study this data once we've mapped it

8:00

it's actually very similar in respect to how the um covert variants are being detected so

8:07

well effective using genetic sequencing techniques to profile the type of complaints that are being received

8:13

and what's very interesting about that is you can then start to overlay interesting metadata over the top of

8:18

that so what's happening in different uh corners of the map uh how does it influence results and how

8:23

does it actually does it allow us to actually risk profile whether a company is uh

8:28

performing well or morally or if there's any weaknesses in their process that make them a risky company so within

8:36

the context of um this map we've i've just zoomed in on one particular uh corner of the map

8:42

which is uh abusive language which made up about 70 of all lending uh complaints uh with a

8:49

particular you know there was debt collection data in there but also vehicle lending data as well and we can actually start

8:56

to profile so now i've zoomed a little bit further and there's a couple of dots in red and that for example is general motors

9:03

and what we're able to do with this kind of data is to actually or the next steps for this data we don't have this data

9:09

within the digital sandbox is we could start to risk profile the companies themselves so for example

9:15

how many documents uh landing into this corner of the map are indicative of bad future performance

9:21

you know in two years time for example uh so with that in mind effectively what

9:27

we've been able to do with the data with the digital sandbox is to demonstrate our ability to demonstrate strategic

9:32

oversight of what's going on within an industry or within a particular quarter of the industry we can also show

9:39

how it's possible to conduct ai assisted company profiling so we can actually show that you know the profile of complaints

9:46

of a company can be linked to a particular outcome we can also demonstrate so the model is actually

9:53

visual so we can actually explore the map understand why the ai is making the decisions that it is

9:58

and we can actually create company company efficiency profiles for example and then finally the tech

10:05

stack held up even though we were dealing in literally uh billions of words of content so we've got a scalable aws

10:12

microservice architecture which is easy to adapt to any new problem space and we've actually got a hyper computing

10:17

environment that allowed us to to build the model so we're able to roughly process about a billion words per hour so with that in mind um

10:26

happy to take any questions okay so i'm saying about are there any language limitations

10:33

um so yeah there's no language limitations all languages in any language not just

10:40

english all work to a a very similar zif distribution so like a kind of frequency of use

10:46

and we leverage that to sequence the documents that we're dealing with so no there aren't any particular limitations

10:51

not even for other languages um so i've just seen another question here

10:58

give an example of a complaint coming through the system so we actually weren't able to do our original use case which is to actually process a claim

11:04

because we didn't have that kind of data so we were using um lending data so the best we could do is profile or risk

11:10

profile um the lenders themselves uh using a kind of patterns of behavior

11:17

in the claims that they would be uh that were being received on uh on their behalf

11:22

uh so that's all we were able to do within this particular study um but effectively what happens next

11:29

with this kind of modeling is that when we receive 50 or 100 complaints from a potential an sme we'd be able to say uh this

11:36

this is a potentially risky one because we've seen this pattern of behavior in previous companies that have gone downhill or have gone up uh

11:42

or are doing certain things that we wouldn't like them just to do um

11:49

so i'm just looking so there's one sorry you go for it shall i talk about

11:57

about resilient ai um the yeah the system is completely is is literally baked there's about

12:04

seven trillion inputs into the model so it's very very difficult for any individual

12:10

complaint to be um to be kind of gained in any particular way and the other

12:15

thing that you're just dealing in such vast quantities of data that would then we tend to have seen everything before

12:20

but there is an advantage in the fact that we can visualize the model is if we don't have if something comes in and it looks rare to the system then we

12:28

can actually say you don't don't process this yet we don't have enough data so not only can we make predictions we can also switch on

12:34

and switch off the predictive capabilities if we don't feel it's relevant so that actually hand answers the next

12:41

question which is can it handle missing data so i feel like i've answered two for one there um

12:46

and yes to be absolutely clear we are profiling um the lenders at the moment we didn't

12:52

have enough data within the data sets uh on this on the digital uh

12:58

sandbox to actually do the original thing we were going to set out to do which is to profile the claims and the applications

13:04

themselves so david one from the chat box and and

13:10

just to restrict treats this comment do try and use the q a function on zoom

13:17

so how do you think this will be used in practice to help increase access to finance smes and that's from

13:22

matthew patrick okay so that's a very good question the

13:27

use case we've ended up running with is more of a regulatory one so for example the organization that

13:34

supplied the data in the first place could be using this to profile the quality of the lenders themselves

13:40

so in this context we can actually start to based on triangulating maybe you know they might

13:45

receive 100 complaints from a particular company we can profile those complaints and actually say this is a high-risk lender

13:51

for example or this this lender is not performing to our standards or they are

13:56

if we've got metadata linked to the lenders themselves we can actually trace that back to patterns of behavior with the types of lending

14:02

types of lender that they're actually supporting

14:08

any other questions i should answer i don't know how much time we've got so

14:14

let me just check the chat again yeah no i think we're okay um i've just

14:20

seen one here about the economic business model um i'll just quickly answer that one as well

14:25

so the business model is adaptive so we're an implementation company so we work alongside

14:30

tends to be multinationals who are looking to implement ai assisted business process and we hold

14:37

their hand hand through a three-step process so the first stage is the data audit which is all we did within the context of this

14:43

environment then when we start to embed outcome data we can start to build models stress test it and then we leave them

14:50

with a a license to the software that then allows them to actually deploy that service live on an ongoing basis

14:55

and that's a kind of a sas revenue model

15:01

excellent thanks david um final call for questions for david

15:08

otherwise we will move down the agenda could we analyze uh is that i've just

15:15

seen another one coming with from marcus uh could you analyze documents that are not structured

15:20

um yes yes we can so we we specialize in unstructured data so from all the way

15:27

through from [Music] university essays for example predicting the grade of an essay all the way

15:33

through to handwritten doctor's notes as pro risk stratifying patient there's

15:41

all sorts of different applications um and the idea is that we can then

15:46

predict whether or not the form is as it should be if there's types of complaints that

15:51

have been made or if there's types of forms that have been rejected for similar reasons before or whether indeed the model actually

15:57

picked up so we can have even with handwriting conversion we can even detect whether the system's accurately converted the handwriting so

16:04

there's all sorts of steps that you can take to to improve the accuracy and the worst thing you can do is make a false prediction so there's a lot of

16:10

steps you can take to just either pass it back to a human or to kind of risk profile the quality

16:16

of the of the document itself so for example in an nhs context half of it is do we have enough

16:21

information from the patient to make a prediction so in in different contexts we do have to deal with a lot of unstructured

16:27

data thanks david and thanks fluence um right time to hand over to

16:35

ian and simon from open banking reporting

16:43

good morning everyone my name is ian douglish i am the commercial director for open banking reporting

16:49

i'm joined the call by my colleagues ken robinson uh who's the chief data officer and simon cockel

16:55

who's the chief technology officer um so to start with uh an introduction

17:02

to where we were when we kicked off the process with the sandbox

17:08

and we set ourselves four key uh objectives um and

17:15

we identified quite quickly that one was going to be a challenge but three were most definitely doable and we

17:22

were successful in creating a data and creditor analysis the peer group analysis

17:28

and covered business model stress and that's down to um a lot of support

17:35

from the fca um clc uh alan turing institute and authentic sandpit so a big thank you
17:42

for for those helping in particular can work his way through the objectives

17:49

so that was what we set ourselves that's what we achieved and what i'll do now is
we'll run on to

17:56

a demo of our solution and i'll call out the areas within the demo that we've

18:01

been able to take forward so our open rep solution

18:07

utilizes data from traditional credit risk sources macroeconomics open banking data
and the

18:13

key area of our solution is access to real-time data from the cloud-based accounting
packages

18:19

now traditionally data sources from the credit reference agencies were historic

18:24

or are historic and therefore by definition of date whereas utilizing real-time data from

18:30

the accounting packages together with advanced analytics and robotic automation
which is our ip generates proactive

18:38

alerts actual insights and next best actions enabling lenders to focus on their

18:43

customer relationships while open rep is relevant across a number of sectors namely
investment

18:50

groups business angels procurement teams insurance and smes our key focus at
present is on the sme

18:56

lending market across both banks and non-bank lenders and then saying that but we
identify um

19:04

right up front that the solution we're creating here will really have to be offered to the
smes

19:10

very quickly to ensure there's a appropriate value exchange but the devil i'm running
through to snipe

19:16

excuse me was designed for relationship managers within lenders and on the screen
you

19:22

could see what they would see when they went into the screen they've got 258 smes
the current lending

19:29

interviews and up to date and there's five risk categories now we've got more than
five but for the

19:35

purposes of the demonstration we're is articulating or demonstrating five-year money laundering through to fordress

19:42

and and the relation manager would then kick on on this one we're going to conduct risk

19:49

and that brings up a list of companies that are being flagged under the conduct

19:55

risk cad risk category and it's important to call out here that it's not the open banking

20:00

reporting risk category the risk categories will be deployed on our bank by bank bases and

20:07

it was certainly their risk categories that would be deployed now we can see

20:13

what the current lending is we've got big speeding credit score coming through as well and if we click on global trends

20:19

clothing clothing it takes us through to what is the key

20:25

area and the key benefit of our solution we're bringing through management insight and actions

20:32

and it's designed so that a layperson can quickly come to terms with what the issues and

20:38

challenges are of a particular sme and what we're listing here is conduct liquidity money laundering

20:46

and training i'll come back to training in just a second but if you go to business under business

20:52

snapshot you'll see that we've got age debtors and suppliers and that

20:57

aligns back to the first objective that we set ourselves with debt or credit analysis and the

21:04

data that we've been able to take from the sandbox has helped us shape these particular widgets where we're

21:10

able to give an sme and a lender our view of the debtors suppliers debtors and creditors

21:16

and to be able to identify upstream and downstream where there may be issues and challenges

21:22

which they want to address earlier in the cycle and the obvious one to call it here would be

21:28

john madison the sheet metal that would be one that the lender of this and we would wish to take action

21:34

upon immediately we can then go further down uh we have the

21:41

balance sheet profit loss and the cash flow we have additional trends and forecasts

21:47

which will be useful to the relationship manager of the bank and if we go into the profit and loss

21:55

you can see that this information is to hand as well so you've got the p l um again underlying data

22:02

and various insights and information cash position here and that just continues so well this is

22:10

not designed the tool is not designed just to give access to this information it's the management

22:15

insights and actions it does give the relationship manager the option to see the information very

22:24

quickly now come back to trading

22:32

if we click click on that one this takes us through to it's coming up

22:39

this takes us through to the second one which is the peer group analysis and again using the data within the sandbox we were able to do

22:46

but compare a company versus the peer group and that enabled them to be able to at a

22:53

glance see where they were not performing as they should be or maybe over performing

23:00

in any case be able to take get an insight and then be able to set the strategy to investigate what the issue

23:06

or challenge was where was that an issue was causing them a problem or was there a missed opportunity

23:11

but this information would give them that sort of insight and it gives them a ranking within the peer group which

23:16

we're then able to articulate in various graphs

23:23

this is just to let you know that we're at times if you could wrap up and then coming to the third one which

23:29

was the the covert piece which was the covered business stress uh we were able to use the sandbox data to be

23:36

able to help identify uh customers during covert that were at high risk and

23:42

that the banks wished again to discuss earlier in the business cycle and we'll stop there

23:52

that's great thank you ian uh questions for ian and the team at open banking

23:58

reporting please okay we've got one there uh here

24:06

should i read it here it's how would the how would the app have access to the financials of all

24:11

those companies well the we are using utilizing a third

24:17

party called kodak um who are this aggregator bringing in the accounting data

24:23

from the accounting platforms and they pull that data in but they pull it in in a raw format what

24:29

we do is then take that data and put analytics on top of it which creates

24:35

management insight and actions

24:41

great and chris asks regarding peer group comparisons how do you deal with

24:46

the possibility of different entities having different accounting policies eg for measuring values such as

24:52

depreciation stock valuations etc kane can i pass that one to you yeah

24:59

sure so in terms of identifying a peer group it's been identified on some objective i guess less subjective

25:05

measures such as the the turnover in the business and the um the number of employees and the sector

25:12

so it's quite an objective criteria in the example we've presented in terms of going into particular values

25:18

you're absolutely right there are a few things that we need to be very careful of um but we're taking from the p l and

25:24

we'll be able to look at each in individually to make sure that it is is valid but the measures that have been

25:30

chosen i think we've tried to keep away from some of these uh complexities at this point

25:35

um it is a poc we're trying to judge kind of what is the market appetite but we believe that we're able to

25:42

to work with a lender on this product would be able to look at their portfolio and increase the the understanding of

25:48

that portfolio

25:55

great we've got another question did you use any data from the sandbox in your presentation

26:00

so we did so we used the data for all of the work streams that uh ian mentioned so first of all

26:07

for the contagion which is looking at the upstream downstream from the account receivables file we

26:15

were able to look at the particular invoices we recognize there's not a counterparty on the invoices so we generated that

26:22

synthetic ourselves in terms of the um looking at

26:27

the peer group analysis that was completely uh originated from the sandbox and for the stress modeling

26:35

um that was using the ons business interruption survey to understand what was the responders

26:43

we know that's been fused onto the sme data set and we use that for modeling purposes so we have used it exclusively and

26:52

finally from emily how do you determine the money laundering risk based on data from accounting software

27:00

so you're right so understanding what might be an indicator and we would work with our lenders to

27:06

understand what it is but we would think about whether there's certain assets that might sit with on the balance sheet

27:11

um in terms of what might be some of the techs that might come through equally we could see if there's any uh

27:18

secular transactions that's working through the invoices in it and um that's in through the invoice stack

27:24

as well um i think these things are things that we will uh illuminate but it's certainly if there

27:30

is an insight that is um that we can find we'll certainly be looking for it

27:35

and there are adults sorry i think i'd also add that um to ensure that the data coming

27:42

through the cloud-based accounting packages is as accurate as it is whether that be for uh for money laundering

27:48

um then our partners have got capability to vet that against open banking data

27:55

and be able to vet that the data coming through the cloud-based accounting packages is accurate

28:03

perfect thank you i'm going to move us on now but thanks to the team open banking reporting and and the questions can be addressed uh

28:11

in the chat box uh as we work through the other presentations as needed i'm so going to hand over to

28:16

adam stones now from company watch limited

28:24

so good morning everybody i'm adam stones i'm a data scientist at company watch and today i'm going to give you an

28:30

overview of our participation in the pilot so just some context on us we're a commercial credit reference agency based

28:37

in the city of london we serve a broad base of clients particularly in financial services so bank can get insurance

28:42

but really anybody with trade credit or supply chain risk and our main activity is to assess the

28:48

financial health of companies and we do that mainly based on financial data in the public domain so our

28:54

principal product is the h score which captures the probability that a firm will experience financial

28:59

distress in the coming year and historically that's been very accurate predicting around 90 of corporate insolvencies so to give you

29:06

a flavor of how that works here i've got a single company tesco plc and you can see that in the last few

29:12

years its financial health has been steadily improving according to our h score and users are

29:18

able to dig into the seven factors that make up the score and see how factors that

29:24

how those factors changing uh has an impact on the main score and it's really that explainability that

29:30

transparency that's a key strength of our platform and it's something that really runs

29:36

through the the entirety of the platform so moving towards our participation in the sandbox

29:41

and our motivation there are challenges to working with publicly available data

29:46

in particular there's a lag between the financial year end of a company and when it files its accounts and mainly for smes

29:53

which is the subject of today they often file abridged or even micro entity accounts and typically

29:59

they're unaudited now those challenges have been exacerbated by covid for instance the

30:04

government have extended their filing deadline for financial statements but that comes at a time where really

30:09

access to this information is really important because of the increased risks of doing business and also the need for smes to

30:16

access short-term credit so we acted really quickly as a company in incorporating macroeconomic data from

30:23

the obr in order to forecast our age score and forecast scenarios on a sector by sector

30:30

basis so if i just go back to tesco you can see here that we've been actually able to provide a forecast

30:37

h score for the end of the last calendar year even though their financial year until the end of

30:42

february 2021 is still underway and of course the nature of risk

30:49

is granular and it's actually better to look on a company by company based than on a sector by sector basis and so

30:56

we actually provide our users with the ability to just play around with those assumptions and that's really a second key theme of

31:03

our platform is that users can use our platform as a starting point and then incorporate their own knowledge

31:09

and expectations in order to stress test the metrics that we provide now our ambition as a data science team

31:17

is to develop improved models to better forecast company outcomes and company distress and our

31:23

big hope for the sandbox that was that we would be able to use real-time data available on the sandbox

31:28

to really engage with the fast-moving covert situation and provide for a more timely estimate

31:34

of credit risk now working with the data in the sandbox was challenging at times because

31:39

even though the transit transaction or banking data had the kind of time resolution we needed for our use case it

31:46

was clearly synthetic and speaking to the teams that made up the data it's clear that any insights that we

31:52

glean are unlikely to apply in the real world but just to show you what we did do i've

31:57

got some cash flow projections here so this is five companies in the sandbox the solid lines represent their cash

32:04

balances relative to the balance at the beginning of the year and this widget enables me to choose a

32:09

projection start date let's say the beginning of march and then the dashed line is the mean

32:14

cash flow projection based on the activity prior to that date and the shaded region is represents uh

32:22

sort of a reasonable best case and worst case scenario and you can see as i move the projections further into the future

32:29

that uncertainty grows so finally i just want to give some

32:34

areas where we have fulfilled our ambition and some tools that are already on the platform so to begin with the tech score so that

32:41

complements the h score in that while the h score looks at the numbers in financial statements the tech

32:47

score looks at the text and actually that's a very good predictor as good as the age score

32:52

of company outcomes and company distress out of that comes searched so searched

32:59

if you've ever used company house pdfs then you'll know that they're just images and actually difficult to read in

33:04

any automatic way so with searched we provide um searchable pdfs

33:10

and you know that means you don't have to read all of the accounts manually and the possibilities for modeling and

33:16

research that come out of that are enormous and finally aphrodite so it's actually possible for a single

33:23

director to have multiple multiple registrations at companies house and aphrodite is a way of tying

33:30

those back together back together so you can properly see you can properly see the history of the

33:36

individuals associated with a certain company so to sum up then i just want to say

33:41

that we've had a really positive experience in the sandbox we've really enjoyed interacting with other teams

33:46

mentors the people who created the data and in that spirit you know please do reach out to us if you've got any

33:52

feedback or suggestions or data or anything it's all welcome and i just want to emphasize that one of

33:57

the strengths of our platform is that it's additive and we do work with clients to combine

34:02

their own data that might not be in the public domain with the data that's on our platform and we've got the expertise to create

34:09

bespoke solutions in order to meet the particular business needs of a client so with that i'd like to thank you for

34:14

your attention and i'm happy to take any questions thank you adam i've got a couple of

34:20

questions coming in from the q a uh i can read them out if

34:25

that's helpful uh yes please yeah sure so how exactly did you or are you planning to tackle

34:31

the specific sme challenge you outlined in your slides yeah so as i as i've said in the uh as i

34:38

said in the presentation um unfortunately when we when we came to look at the data

34:43

we found that the data just wasn't uh wasn't good enough to to really address the challenge we have

34:49

so we sort of made the most of what of the data that we did have and we've been able to provide cash flow

34:55

projections for those companies but we just don't think that those companies that those projects sorry we think the technology is okay

35:01

but we don't think that the the insights that the technology gleans from the data are actually um would actually apply in

35:09

the real world so what we'd be looking for in future iterations of the of the sandbox i do

35:14

want to say that we're very keen to engage with the fca and the city of london corporation with future

35:19

iterations perhaps even having some input on the kind of data that would be useful for this use case

35:25

what we'd be looking for would be data that was more realistic perhaps even anonymized data rather than

35:31

synthetic data and actually really data that we have prospects to get hold of in the real

35:38

world perhaps with open banking um you know so that we can we can take the use case that we developed within

35:44

the sandbox and then take that out um and then and then take that out into the real world for

35:49

our clients thanks adam i think that's a really important point on the dataset i think

35:55

that's all we want to do in terms of this was a pilot we want to learn you know the lessons

36:00

on you know what worked and what didn't and in particular on the data set i think pursuing some of those opportunities

36:06

with anonymized real data alongside synthetic is definitely something that we want to consider

36:11

that's music to my ears um so final question from nicholas who will be your main

36:17

targeted client your main type of client excuse me so as i said already we have got a lot

36:22

of clients and they come from all sorts of sectors both both in the uh both in the private

36:29

sector and in the public sector so we're really looking for solutions that can that can serve those clients

36:35

but it's it's a very broad mix and it's really anybody who works who does business with other companies you know we look at

36:41

all the companies across the uk and ireland and we also look at global

36:46

uh internationally at quoted companies so it's anybody who's doing business with those companies is a potential

36:52

client for us and you know our our ambition with the sandbox it hasn't quite come to fruition

36:58

but i'm sure we'll get there is to be able to get more real-time estimates especially nowadays you know

37:04

with covid um and you know it's not likely that this is going to be the last such situation that we face we're

37:10

probably going to face similar fast-moving situations in the future where we need to have

37:15

access to data that's more current than what we would get from the financial statements so it's basically

37:20

anybody anybody who works with companies or a target client for us

37:26

thanks adam um with that i'm going to hand over to the team at fractal labs if nick and

37:32

silaji are online

37:40

okay hello everyone uh i'm nicholas i'm ceo fractal and i'm joined by

37:45

senhaji our head of data engineering so first fractal's a platform powering

37:51

sme payments and financial services by providing a smarter and more affordable

37:57

and faster payments engine built on open banking and a business insights platform with

38:03

cash flow forecasting at its core the objective is ultimately delivering the right financial product to the right

38:10

sme at the right time in the sandbox we focused on using the data to enhance our

38:16

cash flow forecasting engine to improve access to sme finance here are the key features of the fractal

38:22

platform payments are initiated using account to account bank transfers behind the scenes the fractal

38:28

forecasting engine predicts future cash flow this analysis is

38:34

combined with the lender's credit risk model to determine loan eligibility and financing opportunities are returned

38:40

to the sme along with a revenue-based repayment plan our focus is on facilitating better business

38:46

management for smes and in turn less credit risk for lenders

38:52

let's take a closer look

38:58

here's an example journey of the fractal platform in action so fractal makes onboarding

39:05

smes seamless connecting directly to companies house or facilitating account creation by sole

39:11

traders we connect to 99 of all bank accounts in the uk through open banking and are able to aggregate permission-based data

39:18

from any number of transactional bank accounts we also connect to major cloud accounting platforms such as xero and

39:25

quickbooks pulling in invoices bills and management accounts

39:30

once the data is connected our platform immediately drives financial insights for the sme

39:35

for instance through deep linking with our partners that are untied the uk's personal tax app we are able to

39:42

surface future tax liabilities last month alongside untied fractal was the first

39:47

business to ever initiate an open banking payment directly to hmrc but rather than paying

39:53

the tax immediately the business might also set up a virtual tax savings pot

39:58

our research indicates unforeseen tax bills are one of the top reasons smes need

40:04

financing our analytics engine also picks up recurring transactions

40:10

in this example fractal initiates a standing order thereby saving the business valuable admin time during

40:15

manual payments outstanding invoices can also be paid directly from the app

40:21

the forecasting engine predicts the impact the payment will have on future cash flow

40:26

providing the business with greater visibility into their upcoming needs when forecasting cash flow it's

40:32

important to determine low points in the coming weeks and months behind the scenes are lending partners use the

40:38

connected data to determine credit eligibility and automatically underwrite a loan in this example a credit card with a 25

40:46

000 spending limit shores up any shortfall in the coming months digital processes make lending more

40:52

efficient and embedded finance reduces the risk for both smes and lenders

41:02

over to you senhaji hello everyone i'm sanhaji head in data

41:08

engineer at fractal labs thank you nick for presenting our product using payment automates banking services

41:14

helping smes with the cash flow forecast and enabling the loan applications our approach was data-driven we

41:19

extracted valuable insight that drive automated and partial decision making by financial institutions for small

41:25

businesses and individual traders using sandbox data set we trained multiple models to

41:30

build a solid forecast building a complete view of what when and how much cash flow would be

41:36

available for the smes we use two different approaches the first one is clustering transactional

41:42

data using frequency trees to identify our current patterns the second one is an

additive model for non-linear strength fitting with seasonality to predict trend and seasonality within the time series

41:54

in both cases we managed to use the data sandbox the sandbox data to measure accuracy and fine-tune the

42:00

models we present now some analysis of the

42:06

result we managed to extract from the data on the left figure we have a business account for which we cluster

42:13

the bank transactions based on their currency amount category and contraparty each dot is a transaction each caller is

42:20

a cluster we then analyze each cluster and projected a prediction into the future

42:25

our previous model were was less accurate model has less than 30 percent mp error it also

42:32

allows validate predictions on categories like salaries and loan repayment on the right

42:37

figure we have the same business account this time we focused on accumulated cash balance

42:43

for which we identified the trend in the weekly monthly quarterly seasonality and we use the regression to project it

42:49

and get and get the forecast and this is an outsample testing we trained nine months

42:55

and we trained nine months and we predicted one three months the orange graph is actual the board is historical and

43:01

the blue is the fractal forecast as you can see the projection is really accurate

43:06

and we have also a cloud error that increase over time

43:13

although the timing of these pilots the quality and type of data were an issue at first or we managed to really extract value

43:20

from from these pilots we want to thank the fcs sandbox team that enabled and enabled this project for fraction

43:26

and provide the data the framework and the support as next steps we want to reach out to

43:31

any businesses looking to offer sme clients value-added solutions including faster payments and smarter

43:37

cash flow forecasting and lenders who want to drive profitable revenue from this underserved segment of

43:43

the economy it's 2021 embedded finance is here

43:48

let's make cash flow thanks for your attention brilliant thanks nick and thanks

43:54

shanahachi um any questions for the team at fractal labs we've got a couple of minutes for

44:00

those i think we have a question from stewart

44:05

adam um probably this was for adam we have a question from ian

44:11

a fracture watch improvement were made to the sandbox data thanks um we we basically used the data

44:18

the data as raw initially uh but we we extract some so we pre-process it

44:25

basically we processed the categories and merchant uh and we also processed the amount in a format we can actually

44:33

use it for the projection i hope this one help ian major

44:42

we've got one from timothy are you targeting lenders internationally as well as in the uk

44:50

we're primarily uh uk focused um but our payments license under psd2

44:57

is valid across europe we also are working with other players in the

45:03

u.s and australia but that is not based on our license and

45:08

more based on the fact that cash flow forecasting so long as we are able to pick up

45:14

transactional bank data or accounting information is valid so the answer the short answer

45:20

is yes but our primary focus is the uk and europe

45:26

and i've got one from the chat box and once again please do use the q a tool

45:31

can you please explain your data aggregation tool can you aggregate multiple transactions and data sources

45:39

um yes we yes we do we we basically use the banking

45:45

transactions at first to uh to do a cash flow forecast

45:50

we use uh user inputs to improve these these forecasts so those are data that's

45:56

provided by the users we also use accounting data so we use um bills and invoices

46:02

to strengthen the transaction the forecast um accuracy so the baseline for the

46:07

forecast is around like 30 20 percent error and we can decrease that error to less than 10 percent with

46:13

with the other with other data sets that we that we get and data sources

46:18

and through open banking connections uh we have 99 bank coverage in the uk so we're able to take from

46:24

basically any bank in the uk perfect and then finally uh if you can

46:32

add and add anything to this that you've not already covered but hi fractal how did you use the sandbox data

46:43

so uh synagi maybe you want to briefly summarize that again uh well to

46:48

summarize so we focus we really focus on the on the banking transactions um and we

46:55

we basically uh split those um by entity um so by uh by sme company and then we

47:00

eventually did by uh by accounts uh so the all the analysis we did are account based

47:06

and we we focused on on a couple of features from the data we focused on the

47:12

merchant which which control party basically is on the transaction we focused on the category and so

47:18

basically from the transaction descriptions extract some category and and use that and to categorize the transaction and

47:24

categorize the cash flows amounts and dates are very important so we use those to uh

47:29

to aggregate and projects into the future and so those are the main features and data we got

47:36

brilliant thank you i think there are a couple of questions in the q a if the fractal team want to have a look at

47:42

those um otherwise thank you and i'm going to move us on now to untangled finance

47:47

thank you very much

47:54

hi everyone um and i are co-founders of entangle finance we have been working in fintech

48:00

and blockchain space since 2017 and

48:05

our product is being used by banks in africa so we started um antango finance during

48:11

the covid lockdown smes took the biggest hit as a result of covid and many of them will be

48:17

struggling to operate and repay their balance back loans these loans asset originators need

48:22

cheap and diversified funding and investors on the other hand are looking for

48:29

are looking to invest in higher yielding assets however at the moment the capital market

48:35

money is not flowing back to the real economy very efficiently because the securitization process is very

48:42

complex with very high transaction costs for the smaller portfolios and there is lack of transparency one of

48:49

the things that caused 2008 financial crisis was that there was a loss of nexus between the assets and also the node

49:00

holders so antango finance is a um as the digital securitization platform

49:07

by incorporating blockchain tokenization automated legal documentation

49:13

and live reporting advanced analytics we make securitization process

49:19

of sme loans invoices and green access cheaper simpler and more transparent

49:26

with fca sandbox data we were able to train our platform and have completed the third

49:34

third part of our solution which is the portfolio analytics part therefore today i will demonstrate an

49:40

end-to-end securitization profile for a portfolio of sme loans downloaded

49:46

from the fca sandbox and i will demonstrate including uploading the local assets

49:52

structuring the asset pool sell the pool to an spv or another originator structuring the

49:58

security securities issues and also redemption

50:07

so the process starts off with an asset originator first is to create a pool by filling in

50:13

some information of the assets for sale to transfer assets into the pool with a

50:21

set of electric eligibility criteria there needs to be an asset transfer agreement

50:27

which is generated by answering a set of questionnaires on our platform so these

are smart legal clauses that

50:33

contain computation within them once executed we will be able to track performance of

50:39

triggers of these items the next step is to upload loan assets in a formal

50:44

spreadsheet we follow the european central bank's standards so motion of the field

50:49

can be identified automatically if not the user can select from the drop down box

50:55

or self-define after validation these 201 assets that were worth 12.8

51:02

million pounds would be tokenized each loan is represented by unique token with a unique obligate payment terms etc

51:10

you can click into these tokens to see underlying information and also the record will be stored

51:16

on the public blockchain which you can have access to

51:21

now after filling some pool sale information this loan portfolio offer will appear on our private

51:28

marketplace for institutional investors here example of an investor trying to

51:34

place a bid when the sale expires the seller is able to review the bids

51:39

and select the preferred bidder successful bidder will pay for their bids and then the

51:44

seller will then transfer the asset pool to the new owner

51:53

so now it has been transferred to issue security from the pool the issuer or the new owner of this asset will go through

52:01

questionnaires so that subscription agreement and also investor prospectus can be

52:07

generated so this is to approve we run a dutch

52:13

option for securities issuance it works on a reducing prices or

52:18

increasing yield basis auction will be closed one order is fulfilled

52:23

or when the time is expired so at any time investors can access

52:33

portfolio and it's on the latest performance of these assets they can also define their own performance

52:38

metrics and our platform will be able to provide report based on that so these are the data

52:44

the underlying loan repayment is tracked as you can see from the transaction history here

52:49

once the pool has enough to make a redemption as per the legal agreement the payment will be made to the node

52:56

holder in accordance with the with a pool payment waterfall and of of course all these are reflected on the

53:04

public blockchain so you can see here that um now the note

53:10

has been redemption has completed now let's go back to my presentation

53:19

so here are some of the use cases that our platform is capable of facilitating right at this moment and because of the

53:26

setup of our platform um to adapt for a new use case is very easy for us because we

53:32

we are done by building blocks there's a public blockchain um there is a token layer there is a

53:37

smart contract layer containing business logic and then the ux ui in the front so with the um

53:46

and with the help of fca not only we have data we had we have been able to access to a

53:52

lot of useful data we also had some very interesting conversations with many of the mentors and

54:02

these are potential business company investors now we have arranged a few follow-up conversations already

54:08

um as part of this final demo we're looking forward um to have more business partners like

54:13

originators institutional investors and of course company investors who want to join us in making funding

54:21

to smes cheaper and more efficient thank you excellent thank you very much for that

54:28

so have we got any questions about untangled finance

54:36

i've got a question from chris do you guys want me to read it out or we got it in front of you

54:42

yes i got it okay so this is a question that whether our solution is being used in africa so um we have another

54:48

venture we've been working in blockchain space since 2017 our first solution was

54:54

to um it's supply chain financing so we work on the origination part we we help banks to land based on invoices

55:02

collateral other and anatomical clusters like complementary purchase orders etc is that solution

55:09

that is now being used by by some banks in in in africa

55:14

so i'm trying to finance these um targeted at more yeah

55:21

is there a special reason why has been launched adopted there okay so that's that i have answered that which country

55:27

are using it so the the supply chain origins

55:34

so second question is apologize if i miss is challenging is available within these

55:40

two uh operating agency methodology based modeling that's right so basically with uh without who you can

55:48

issue two type of uh node tokens we call them um obligation um payment application

55:55

token so one is the senior obligation token and the uh sort and the other one is job which is

56:01

junior obligation token so charging is uh is possible and for

56:07

a certain asset class like sme loans and invoices we do follow the rating

56:13

agencies criterias in terms of exclusion eligibilities and reserves

56:18

calculations now the next question how do users carry out long due diligence okay so when the assets are

56:24

uploaded there needs to be a valid data and it's typically our accountants which will perform due diligence on these assets

56:32

um there wasn't enough time for us to demo that but it's a it is a process relating to that that happens on our

56:38

platform as well and the validator will have to approve it so our personal data stored on public

56:43

blockchain or outside the blockchain that's right so the personal data is stored outside of the blockchain

56:50

so we only saw on the blockchain some [Music] pertinent information about the the

56:55

assets but no personal data is on that the personal data is stored in

57:01

centralized relational database how do you collect non-level data from

57:08

bangko yes so so that is possible either through uh daily secure file uploads

57:15

also apis so um as memory demonstrated you can see there's a mapping with the data field

57:21

between the originators data set and our platform according to

57:27

the european central bank um data set so we're trying to make ourselves as flexible as adaptive as

57:33

possible so starting the lowest integration is just a daily upload of spreadsheet to a next level which is the

57:39

api connection and we have a lot of experience on that working with banks in the last three four years and then the ultimate which

57:46

we are very future proof is also that um that we're working towards a a scenario

57:51

where the the assets are originated in a token form there is no longer need of

57:56

paper documentation but our technology and solution is capable of all three so we can work

58:02

with the existing and the future can asset be sorry are we running out of

58:07

time yeah we're just out but if you guys wouldn't mind answering that in the chat that would be great and then we can um

58:14

we can we can make sure everyone gets their questions answered um so thanks guys thank you so thank you so

58:20

much for presenting i'm going to move us on to our final team so it's alexis bell from

58:25

so to signal analytics thank you

58:32

i'm alexis bell i'm the founder of fraud doctor and the visionary for the shortest signal analytics team

58:38

i'm excited to be showing you our results addressing advanced analytics to improve modeling and risk assessment

58:44

shoda is an early stage first to market solution that signals potential risk and uncertainty

58:50

in company financials whether you're a large or small entity confidence in company financials is at

58:57

the core of our collective economy and society as a whole weaknesses in corporate financial

59:03

reporting negatively impacts every part of society if you fall into

59:08

one of these categories you are impacted by what occurs in the market people can lose their investments

59:14

others can lose their access to credit confidence degrades in an attempt to solve this

59:20

problem existing solutions have tried to scrape websites looking for patterns others have tried to analyze the news

59:27

looking for early indicators here's the deal we know that it typically takes over two

59:32

years before anyone realizes there's a problem and we all find out about it when it hits the news

59:38

by the time major misstatements hit the news it's too late by then the damage has already been done

59:45

there is a more effective objective and timely way to solve this problem

59:52

as a forensic accountant i've spent nearly 20 years across various industries using science to look for patterns of

59:59

human behavior if i was investigating a case for a particular allegation of fraud it was always in the back of my mind

1:00:06

what if i missed something really huge behind me this analysis helped me ensure i was

1:00:11

looking in the right place over time i was able to identify risk with greater accuracy

1:00:17

and even more importantly to understand what it looks like in the data before that decision was made we

1:00:23

call this the signal of pressure so why do i think the best way to understand a company's risk

1:00:29

is to read their financials i have an interesting case study to share with you i don't have time to go into it in

1:00:35

detail right now we do have a 15 minute video that does that but for right now i want to call out

1:00:40

just a few of the highlights cases like tango highlight deficiencies

1:00:46

in existing approaches especially those with ai that are not grounded in domain expertise

1:00:52

many existing risk solutions rely too heavily on one data point which is a model risk concern as you can

1:00:59

see here over reliance on the increasing stock price the blue arrow allow tango to misrepresent the

1:01:05

company's financial health in this case inflated financials persisted for years

1:01:10

the first sign of trouble was a year after the ipo when short sellers sounded the alarm we

1:01:17

wanted to know how much sooner would show to have seen any risks for material misstatements

1:01:23

well let's take a look not only did we detect the specific fraud schemes alleged by the united states sec

1:01:29

and later admitted to by django's executives the ones in the yellow but we identified risk of additional

1:01:36

schemes the ones in red that could have been pursued had they not remained hidden

1:01:41

shoda identified the susceptibility of material misstatements while it was still private before the

1:01:46

ipo in 2011 and after the ipo although shoda's

1:01:51

algorithms were originally designed for publicly traded companies they can equally be applied to private companies

1:01:57

when the data is available the sandbox helped us to prove our concept and accelerated our commercial

1:02:04

plans the fca and the co llc's digital sandbox pilot gave us access to mentors

1:02:11

who helped us to build it the right way from the start in less than two months we've been able

1:02:17

to automate and scale the previously manual analysis process into the groundbreaking solution we are

1:02:23

bringing to market we have ingested all of the current and historical fundamentals

1:02:28

for every publicly listed company traded on the us and uk stock exchanges we've embedded

1:02:35

the algorithms into an existing ai enabled analytics engine and finally using live

1:02:41

data we have tested the performance of soda against over 33 000 restatement impact years

1:02:49

for the companies traded on exchanges in the united states we tested over 15 000 restatement impact

1:02:55

years and had an efficacy rate of 87 percent a unique aspect of shoda is that is industry

1:03:01

agnostic it works equally well across all industries another really important

1:03:07

aspect is that it is currency independent which means that it does not matter what accounting system is used by a

1:03:12

particular country to illustrate this point let's take a look at testing we conducted for the uk

1:03:19

for companies traded on exchanges in the united kingdom we tested against 18 000 restatement

1:03:25

impact years and had an efficacy rate of 85 percent this is really really exciting for us

1:03:31

because these results are before implementing any machine learning this is solely based on domain

1:03:37

expertise we anticipate that these efficacy rates will increase as we continue to tune the show to

1:03:43

scores we also attempted to validate shotify smes however the data was available was

1:03:50

insufficient for responsible testing the pandemic has created an

1:03:55

urgent need to risk score companies and monitor across markets in financial distress we

1:04:02

know that lending is often being undertaken with little knowledge of counterparties you need the ability to triage a company

1:04:08

during an unprecedented time and understand the true financial health of every company you touch

1:04:16

board of directors are under increased scrutiny to demonstrate adequate duty of care asset managers or

1:04:22

insurers with large portfolios trying to highlight emerging risks can ingest show to scores

1:04:28

into customizable dashboards and our platform or theirs while showed it is not a lending

1:04:34

decision tool it does augment your understanding about a company it does provide you with the direction

1:04:40

of inquiry it tells you precisely what areas need your attention some of you will be interested in your

1:04:46

own company or the portfolio you are managing while regulators will be interested to uncover potential risks across the whole

1:04:52

system imagine what you could do when you could see if an entire market

1:04:58

or specific sector was experiencing pressure maybe you would offer help or make

1:05:03

monetary policy adaptations shoda is flexible enough to answer all these questions

1:05:10

shoda is scalable customizable and reliable if you have a need for any

1:05:16

of the areas we've covered here today or if you want to help us grow we want to hear from you

1:05:23

so thank you i'm ready to take your questions

1:05:29

brilliant thank you alexis let's see what we've got

1:05:35

ah topical question uh do you want me to read them out for you alexis is that helpful

1:05:41

so i've got from anonymous did you pick up patisserie valerie

1:05:46

we did um we we did pick them up we scored them um i don't remember the exact

1:05:52

exact scores at the top of my head um paul do you remember paul dunlop is here he's a partner for a

1:05:58

doctor yeah the ticker cake is very familiar to us as many of the other cases in the news obviously here while there's ongoing

1:06:04

litigation with what we'd say but we absolutely scored it and we'd be on happy to share those scores offline

1:06:10

if the person asked in the question anybody else wants to send us a question

1:06:16

directly great any final questions

1:06:22

uh for shelter

1:06:30

great okay thanks guys um so that's we've reached the end of

1:06:36

our six teams um now it's just for me to wrap up so thanks everyone for joining

1:06:43

i really appreciate you taking the time thanks for the teams for their presentations and for participating in the sandbox um

1:06:49

and thanks especially to all of you who have joined us today and ask your asked your questions

1:06:55

um next steps this was a pilot the object obviously as well as developing the use cases um

1:07:02

and trying to provide a useful service to all of the participant teams was to trial the approach draw the technology

1:07:08

trial the data um in the pilot and see what we can learn and see if there's appetite and

1:07:14

enthusiasm to create a more permanent fixture a more permanent version of the digital sandbox

1:07:20

architecture so that we can continue to support different use cases different industry-wide challenges

1:07:26

so do keep in touch we'll obviously let you guys the market participants know what our

1:07:31

plans for the future are but suffice to say we've learned a lot during this project um and

1:07:36

we're going to want to put those learnings into future iterations of this process and

1:07:42

product so the contact details are there do keep in touch so [digital.sandbox fca.org](https://digital.sandbox.fca.org) uk or do reach out to any of the

1:07:49

team here across the sca or indeed the city of london corporation to to get in touch to find that next

1:07:56

steps to continue to ask questions of the team or any of the above so thanks again for joining and hope to hear from you all

1:08:03

soon thank you

English (auto-generated)

AllCoursesRelatedFrom The Global City