



# Supporting innovation in financial services: the digital sandbox pilot

April 2021

# Contents





# 1 Executive summary

**1.1** The Financial Conduct Authority (FCA) has developed a range of tools to foster a regulatory environment to help promote innovation in financial services.

**1.2** Through our initiatives such as the Regulatory Sandbox and TechSprints, we have been able to accelerate responsible innovation, and have helped other regulators establish similar services around the world. Running our suite of market facing services has provided us with rich and valuable insights into the challenges and opportunities facing innovators. In an increasingly digital landscape, we are keen to ensure our services support the needs of an industry moving at pace, and to explore new ways we can support responsible financial services innovation.

**1.3** Increasingly, innovative new products and services are data centric. We identified an opportunity to support market innovators in the development of financial services 'Proofs of Concept' (PoCs) through the creation of the world's first regulatory digital testing environment.

**1.4** In partnership with the City of London Corporation (CoLC), we piloted the 'Digital Sandbox' to provide support to innovators. We also wanted to use this opportunity to measure the efficacy of a digital testing environment and rapidly gain user feedback to inform future iterations.

**1.5** Since the pilot concluded, the Kalifa Review of UK FinTech has recommended that a permanent digital sandbox be created to encourage further collaboration within UK financial services. The Review suggests that these efforts be led, in the short to medium term, by the FCA. We are committed to playing a leading role in the further development of the digital sandbox concept and have outlined below our next steps regarding both iterating and improving the pilot and exploring the options for a future and permanent operating model.

**1.6** In the spirit of collaboration and fostering an innovation-friendly ecosystem, we are sharing our insights and findings with stakeholders. In this report, we provide an overview of the digital sandbox pilot, evaluate how it was used by participants, and the key lessons learned for future iterations. We hope this will contribute to the current dialogue surrounding innovation, as well as be of interest to regulatory peers, as we look to progress future iterations of the digital sandbox.

## Key lessons

**1.7** Ninety four organisations applied for the digital sandbox pilot. Twenty eight were selected to take part in an 11-week pilot which ended February 2021.

**1.8** The pilot gave participants access to a range of development tools, such as synthetic data assets for testing and developing PoCs, an API marketplace, a coding environment, as well as access to expert mentors and observers.

**1.9** While it is too early to assess the long-term outcomes for participating firms, our on-going evaluation and feedback from participants has indicated a range of benefits. These include accelerated product development, validating and improving AI and machine learning models, refining business plans, and networking within the pilot ecosystem. This feedback was collected throughout the pilot, via entry and exit surveys, structured interviews, an independent evaluation, and usage analytics from the platform.

**1.10** The key lessons from the pilot:

- There is significant industry demand, particularly from early-stage firms and start-ups, for a digital testing environment. Analysis and feedback indicated that access to data and the benefits of an ecosystem convened by the regulator were the biggest underserviced market needs.
- Access to the digital sandbox was successful in accelerating early stage PoC development and improving product design, to varying degrees. However, we received valuable feedback on changes to the testing environment which would enable participants to make further progress.
- The synthetic data was the most valuable feature cited by participants while simultaneously the one with greatest potential for improvement. Notably, referentially linked data sets (see feedback summary) and more granular data would enable more effective testing, and for products to be developed further.
- Future cohorts should have a narrower focus. The pilot covered 3 broad areas exacerbated by the Covid-19 pandemic; scams and fraud, consumer vulnerability, and SME financing. A narrower focus in the future would result in more sustained engagement from a closer-knit ecosystem, and fewer data assets being required to meet the use cases, resulting in greater depth and granularity for those produced.
- The core set of tools/features made available in the pilot successfully met an underserviced market need. Participant's did not want additional tools but instead suggested improvements to the available tools. Moreover, certain features such as the API marketplace and integrated development environment (IDE) delivered limited added value for participants and could be de-prioritised for future versions
- The application and approval process should be brought forward, and successful applicants onboarded earlier in the process. Participants should then be involved in the data generation process so that data could be tailored to match their development needs
- Participants would benefit from a more structured journey through the cohort, rather than the digital sandbox being a 'self-service' platform. The expectations of expert mentors should also be clarified and formalised, to tackle inconsistent contributions.

**1.11** The vast majority of participants (84%) reported that access to the sandbox had accelerated their product development, with the greatest estimated gain being 18-24 months. However, there was a wide spectrum of utility, with a minority reporting more marginal benefits. These cases tended to be where participants ultimately found that the data sets provided were not sufficiently realistic or granular to meet their use case requirements and so did not accelerate their product development.

**1.12** However, participants often found that they derived other benefits particularly through collaboration with other teams and expert mentors. These included pivoting business models or better customer understanding and product design. In general, earlier stage firms found the features of the digital sandbox more useful than more mature participants.

## Next Steps

**1.13** Based on the feedback collected throughout the pilot and an independent evaluation, and to support the Kalifa Review recommendation of a permanent digital testing environment, the FCA and CoLC will:

- Run a second cohort of the digital sandbox in late 2021.
- Further iterate and improve the digital sandbox testing environment by incorporating the lessons learned from the initial pilot and making the suggested improvements to the platform.
- Expand on the use of the digital sandbox testing environment to highlight the opportunities and value it contributes to the financial services ecosystem.
- Focus the efforts of this second cohort around the theme of Sustainability/ climate change to support the UK's green finance ambitions ahead of hosting COP26 in November 2021.
- Explore, with industry and other stakeholders, viable sustainable operating models for a future, permanent version of the digital sandbox.

# 2 Introduction

**2.1** The digital sandbox pilot was launched by the FCA and CoLC to test the efficacy of a novel service to support innovation in financial services.

**2.2** Our engagement with industry and experience running innovation services like TechSprints and the Regulatory Sandbox indicated that a digital testing environment would be a valuable addition to the ecosystem. Research suggested it would enable rapid prototyping and proof of concept development, enabling technology solutions to be tested and improved, providing exposure for promising innovative developments, and fostering a collaborative ecosystem that promoted diversity of thinking and cross-pollination of ideas. This would benefit innovators by reducing time to market, improving the quality of solutions, and reducing early-stage development costs. Ultimately this could lead to positive outcomes for consumers and markets, in the UK and beyond.

**2.3** The pilot aimed to stimulate and support greater levels of innovation in financial services by providing access to a suite of tools and features, thereby:

- providing support to innovators in financial services
- validating the hypothesis that creating a digital testing environment would benefit financial services innovation
- collecting feedback to iterate the sandbox model both during the pilot, as well as to inform future versions

**2.4** This report details the process and operating model established for the pilot, the progress it made against identified success criteria, the key lessons learned, and recommendations for future iterations of the digital sandbox.

# 3 Digital sandbox overview

## Concept

**3.1** The FCA has an objective to promote effective competition in the interests of consumers. An essential component and key driver of effective competition is innovation. In addition to providing inventive solutions to meet consumers' needs, innovation enables agile start-ups to challenge incumbents, while driving incumbents to compete harder to retain customers. Innovation can also help to reduce standard operating costs and as a result reduce barriers to new entrants.

**3.2** Since its inception in 2014, the FCA's Innovation division (formerly 'project innovate') has a evolved a suite of market-facing services to support innovative firms and drive improved competition.

**3.3** While there is no 'one size fits all' journey for an innovative new product or business, the FCA's services are designed to support firms at multiple levels of maturity. This includes, from initial ideation and proof of concept stages, to obtaining authorisation, and scaling internationally through cross-border testing.

**3.4** Increasingly, data is playing an ever more prominent role in financial services and financial services innovation. As financial services become steadily more digital, through both new market entrants and large-scale digital transformation at incumbents, the FCA has seen an increasing demand for support services that focus on data and data access.

**3.5** Our experiences running TechSprints, the Regulatory Sandbox, Direct Support and the Advice Unit have indicated that there is a missing rung on the ladder in early stage development. TechSprints in particular have highlighted the importance of access to consumer and market data to potential new market entrants to enable them to develop, test, and demonstrate the value of new solutions. They have also shown the value of collaboration and cultivating focused networks to tackle long standing and complex issues in financial services.

**3.6** The FCA has frequently received requests from both TechSprint participants and from firms applying to the regulatory sandbox for access to a different type of innovation service, to assist at the early, 'proof of concept' stage of product development.

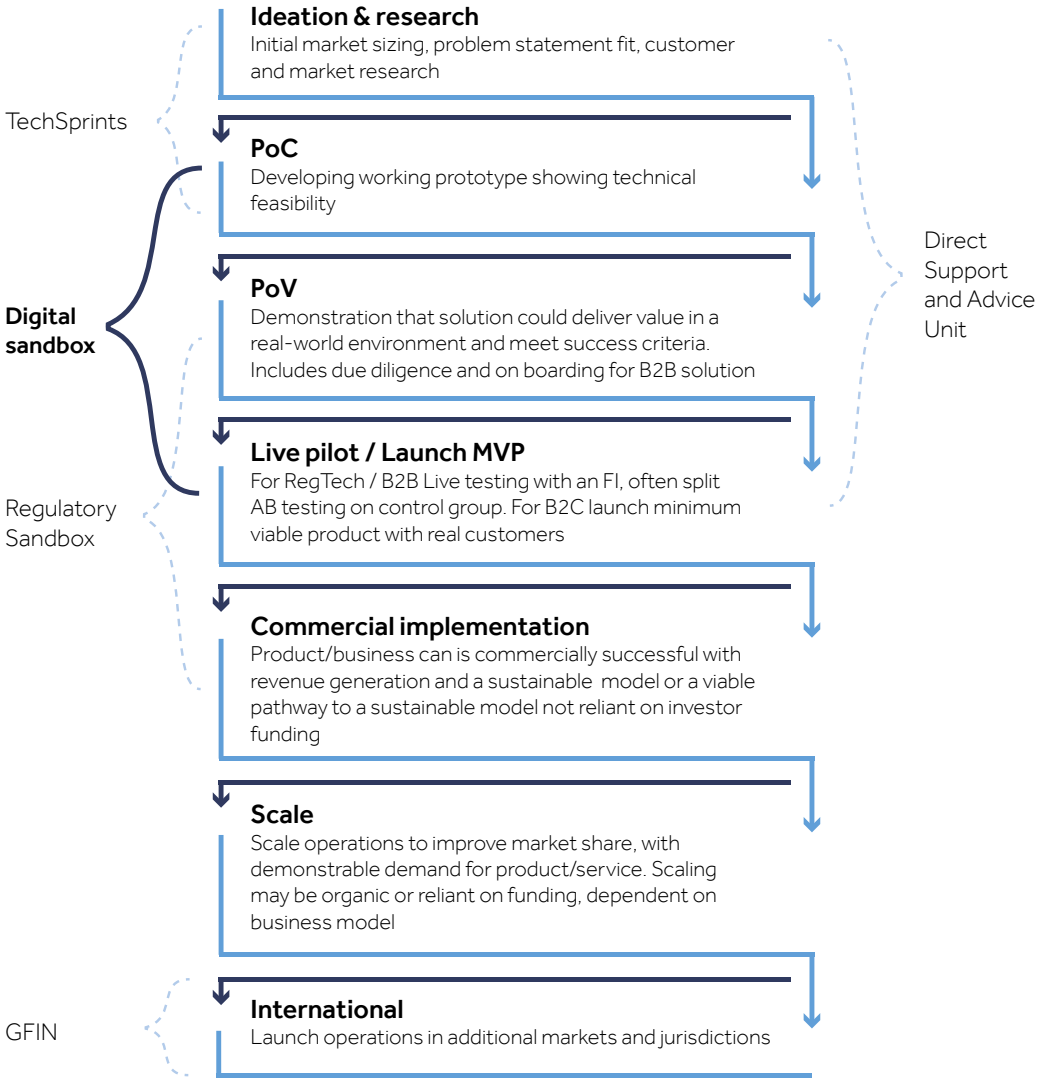
**3.7** The UK's 'pro-innovation' regulatory environment has been frequently cited as a key factor in the sustained growth of UK FinTech and RegTech. It is important for this regulatory approach to continually evolve and develop in order to meet market requirements.

**3.8** Similarly, CoLC is committed to supporting competitiveness of the UK's financial and professional services sectors. This includes enhancing the UK's position as a global leader in technology and innovation, which is made possible by the development of emerging technology products.

**3.9** CoLC's approach is to facilitate adoption of technology and to support the sector in developing, scaling and launching solutions to market more quickly. Frequent feedback from business engagement has suggested that technology solutions have difficulty demonstrating legitimacy to the market, and that access to data was required to do so. At the same time, financial and professional services continue to face challenges that are not addressed by the solutions currently available within the market. Put simply, there is a gap. CoLC were interested in finding ways and means to increase collaboration and communication to address this disconnect and to provide support to enable products to develop.

**3.10** These mutual areas of interest led to the FCA and CoLC ('the Working Group') entering into a collaborative agreement to explore options for a digital testing environment – the digital sandbox pilot.

***Figure 1 – Example of an innovative product/solution lifecycle and interaction with FCA innovation support services. The Kalifa Review of UK FinTech has proposed a 'ScaleBox' Concept with supporting recommendations to address the gap for firms at the scaling stage. The FCA, CoLC and others are exploring implementation of these recommendations.***



**3.11** In early 2020 the FCA ran a series of consultations and workshops with stakeholders to identify the tools and features that participants would find valuable in a digital testing environment aimed at helping early stage product development. The FCA had also previously asked Innovate Finance to undertake consultation on the feasibility of an industry sandbox, with a similar feature set, in 2017. The consultation received highly positive feedback, but ultimately the blueprint produced was not taken to production, given the start-up costs and diffused benefits across a range of stakeholders.

**3.12** Based on the findings of these workshops and recommendations made by Innovate Finance's Report, we established a proposed set of features for a digital sandbox pilot.

## Overview of features

**3.13** We know the value that ready access to data and a collaborative network brings for innovators to rapidly build and test PoCs.

**3.14** More established financial institutions have also indicated that a collaborative environment overseen by the regulator, as well a greater signalling of regulatory interest in specific areas of technology, would provide greater internal impetus to engage with innovative providers and/or develop new solutions in-house.

**3.15** The FCA also receives requests for support from firms which don't meet the eligibility criteria of the regulatory sandbox, but whose proposals may nonetheless deliver desirable innovations in UK financial services.

**3.16** Based on the insight gained from the engagement with its innovation services, the FCA determined that the pilot digital sandbox should provide participants with the following features:

### Readily accessible synthetic and other data assets

**3.17** Access to synthetic and other GDPR-compliant data assets to enable testing, training and validation of prototype technology solutions. For example, transactional banking data sets, SME lending data and customer accounts. Within the pilot environment, these data sets could be accessed either via an API or through an integrated Jupyter Notebook (which enabled participants to process and analyse large volumes of data).

**3.18** Additional publicly available data assets were also provided where these were expected to have relevance to the use cases. See annex 1 for a summary of the synthetic data that was made available to participants.

### An API marketplace

**3.19** An Application Programming Interface (API) marketplace where digital service providers could list and provide access to services via APIs. These test APIs enabled innovators to connect and test multiple integrations amongst several existing providers.

### Integrated development environment (IDE)

**3.20** A coding development environment in which applicants could write and test code collaboratively.

### A collaboration platform

**3.21** A platform including communications features such as direct messaging, access to a group of expert mentors, and community showcase spaces, to facilitate an ecosystem of key organisations to provide support and input to digital sandbox participants. Within the pilot this included experts from financial institutions, academia, government bodies, venture capital, and charities.

### An 'observation deck'

**3.22** A demonstration area to enable regulators and other interested parties to observe in-flight product testing at a technical level and provide channels to engage with the cohort.

### The use of synthetic data

For many firms access to financial data is key to developing new and innovative products. However, financial data relating to customers (such as banking or transactional data) contains personally identifiable information and is subject to strict obligations under GDPR. Legal analysis concluded there is not a sufficient legal basis for the sandbox to make large volumes of real personal data available for market competition and innovation purposes.

One way to address this could be the use of pseudonymised and anonymised data – data that have been stripped of attributes in order to make re-identification of real persons difficult. Nevertheless, such data still carries a privacy risk that it can be reverse engineered or combined with other data sets to identify real person(s). In order to provide useful data that is privacy compliant, we therefore chose to generate and make available synthetic data to participants.

Synthetic data is not 'real' data created naturally through real-world events, rather it is 'created' data generated to replicate the statistical components of real data. In the context of the pilot, the key advantages of providing synthetic data was its utility to train machine learning models and its ability to overcome the barrier of privacy preservation to enable the sharing of data sets, thereby increasing the potential for collaboration and testing.

### Synthetic data generation – the 'DataSprint'

In July 2020 the FCA hosted a 'DataSprint', before the launch of the pilot. The objective of the DataSprint was to develop, in collaboration with financial services industry participants, innovators, academics, technologists and data scientists, some of the synthetic data assets to be used by pilot participants. It also provided a unique opportunity to explore different methodologies of synthetic data generation with leading data scientists from industry and academia.

Over 100 industry participants voluntarily took part in the 3-week event and were divided into 2 main groups:

- data science experts to generate synthetic data
- industry experts who could define typologies and behaviours to be coded into data sets.

The primary method of synthetic data generation was a technique called 'agent-based modelling' where algorithms were used to generate behaviours and patterns that were reflected in the data sets. Generative Adversarial Networks (GANs) were also utilised at the DataSprint. These are a type of deep learning model, where two algorithms are trained together, with a generator model trying to produce synthetic results until they consistently 'fool' an identifier model that can no longer identify if the data are real or synthetic. The DataSprint was also supported by the Alan Turing Institute, who provided valuable input by repeatedly benchmarking the produced data against real data to improve its accuracy.

## Overview of the pilot

### Use cases

**3.23** Through consultation with a series of internal and external stakeholders the Working Group developed use cases around 3 topics. These topics were selected on the following criteria, that:

- they were areas that had been impacted or exacerbated by Covid19
- synthetic data generation was feasible for the anticipated data requirements
- the potential for technological improvement in current market solutions existed
- they were closely aligned to the Working Group's business plans and priorities

**3.24** The 3 topics selected, based on these consultations were:

- detecting and preventing fraud and scams,
- supporting the financial resilience of vulnerable consumers
- improving access to finance for small and medium enterprises (SMEs)

**3.25** During the consultation process we received input on a wide range of concepts, products and services that firms wanted to test in the sandbox, and there were calls for the sandbox to be use-case agnostic and open to solutions/PoCs from every sector.

**3.26** However, 2 cornerstones of the digital sandbox pilot were access to synthetic data assets and a focused ecosystem of collaborators. Synthetic data generation is a nascent field, and high-quality data sets are both challenging and costly to produce. It was therefore decided to limit the number of use cases so that a smaller body of data would be required and could be reasonably produced within the timeframe.

**3.27** Similarly, through its TechSprints programme, we have observed that collaboration benefits from having a tightly focused and active ecosystem which is highly engaged in the subject matter. We thought that a use case agnostic pilot would make building this focused ecosystem more difficult. Selecting a narrow set of use cases enabled the Working Group to target specific organisations and experts who were able to support the pilot.

## Application process

**3.28** The application window opened for 4 weeks in October 2020. During that period, 94 applications were received, of which 30 were selected by the Working Group, based on recommendations made by an external panel of experts from non-commercial entities. Applicants were required to complete a standardised application form and were assessed on the following criteria:

1. **Genuine innovation** – Is the solution significantly different from what already exists in the market?
2. **In scope** – Will the solution benefit UK consumers or financial services firms by solving one of the use cases?
3. **Need for a digital sandbox** – Does the solution require the features of the digital sandbox features to be tested, developed, or improved?
4. **Valid testing plan** – Does the solution have defined development objectives with a clear and viable pathway to production?

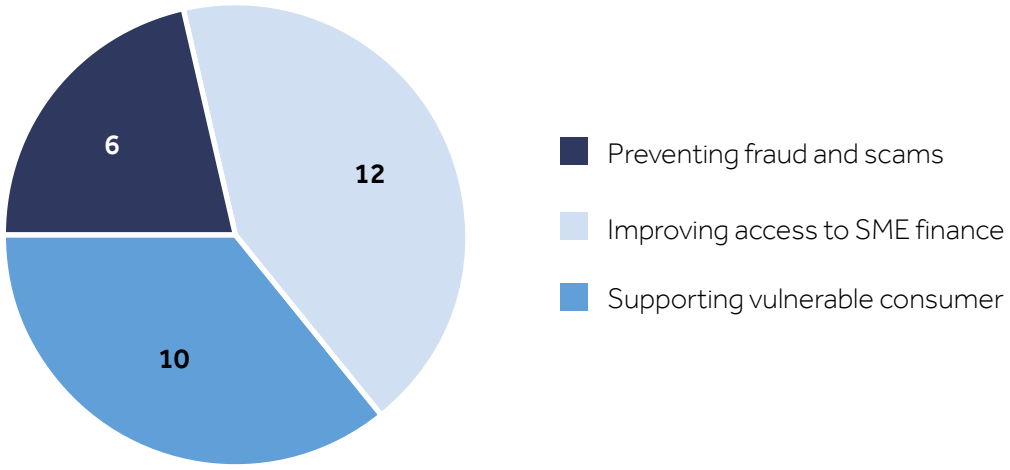
**3.29** Two successful applicants ultimately dropped out, one prior to launch and one in the first week of the pilot. Both cited shifting internal priorities resulting in a lack of resources to adequately test the proposals they had submitted, as reasons for exiting the pilot.

## Pilot participant overview

**3.30** The selected participants in the pilot represented a broad range of organisations across the 3 use cases.

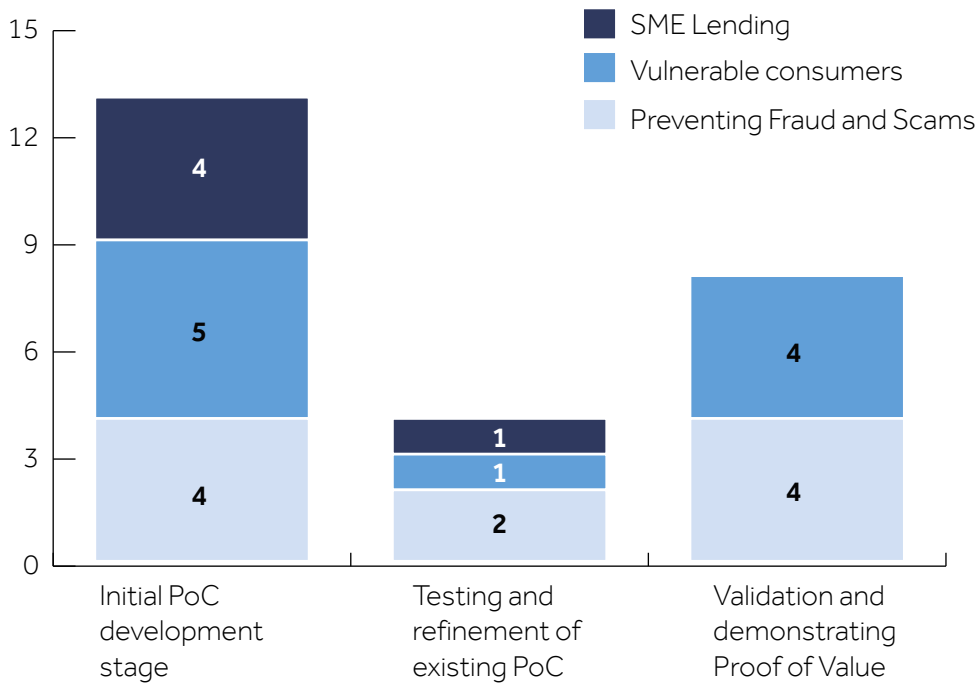
**3.31** The application success rate was proportional across use cases, with preventing fraud and scams receiving the highest number of applications and improving access to SME finance the lowest number of applications.

***Figure 2 – successful pilot participants by use case***



**3.32** Successful participants tended towards wanting to test and develop early stage proof of concepts, rather than latter stage proof of value. However, there was some variance between use cases:

***Figure 3 – stage of development at pilot outset***

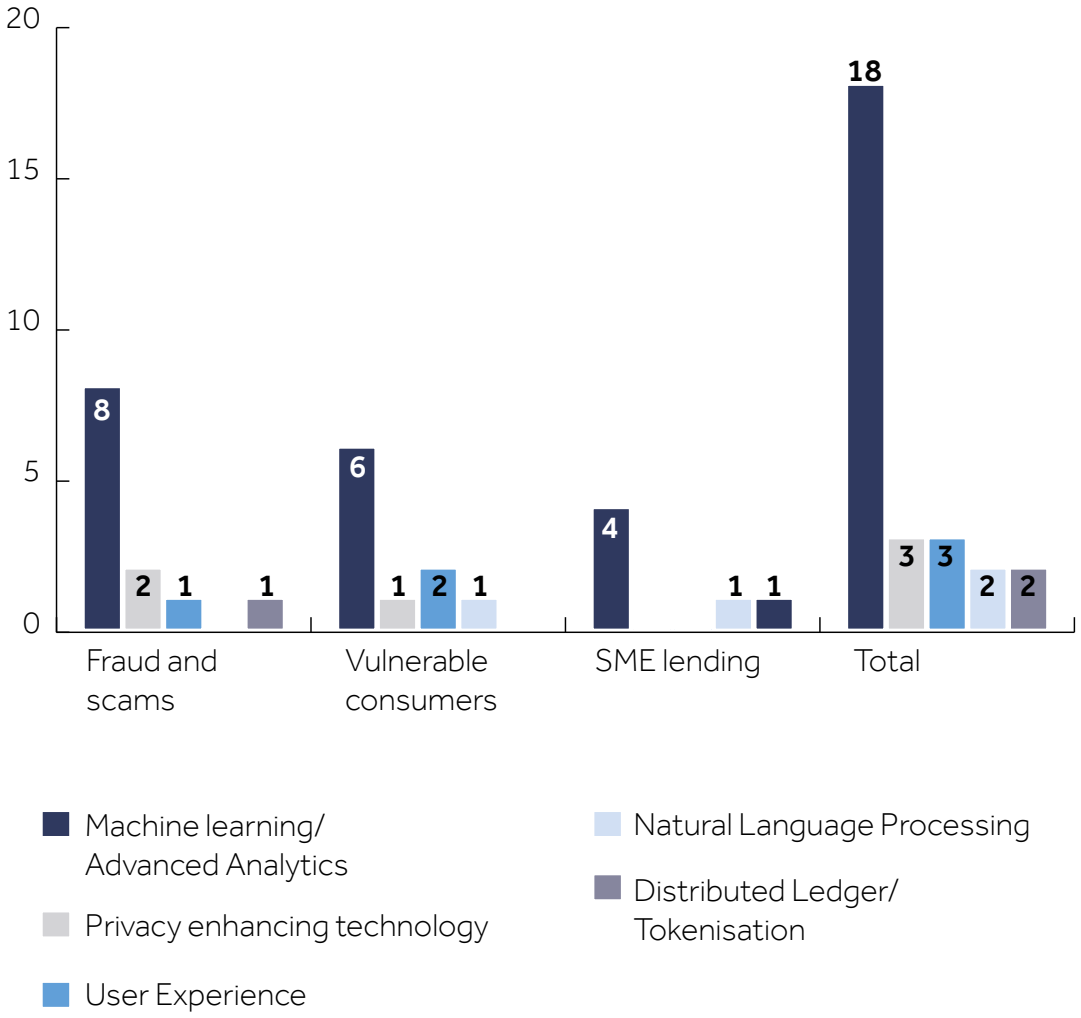



**Product/solution stage of development at pilot outset**

**3.33** Just over half (52%) of the participants entered the pilot looking to develop entirely novel ideas and solutions at the initial PoC stage. This was particularly true of the SME finance use case, whereas vulnerable consumers and fraud prevention also saw slightly more established organisations looking to validate and demonstrate the value of their existing models/solutions. This would suggest that disruption in these sectors is at a more advanced stage, with less focus on the introduction of advanced technologies and more about competing to show outperformance vs traditional methods.

**3.34** Across the pilot cohort, the use of advanced analytics such as machine learning, deep learning, or neural networks were the most common type of underlying innovative technology participants wanted to test. This is unsurprising given one of the core features of the pilot was access to synthetic data assets (which naturally lend themselves to these sorts of innovations), so we anticipated a degree of self-selection. Participating firms also developed solutions using distributed ledger technology and tokenisation, privacy enhancing technology, and natural language processing.

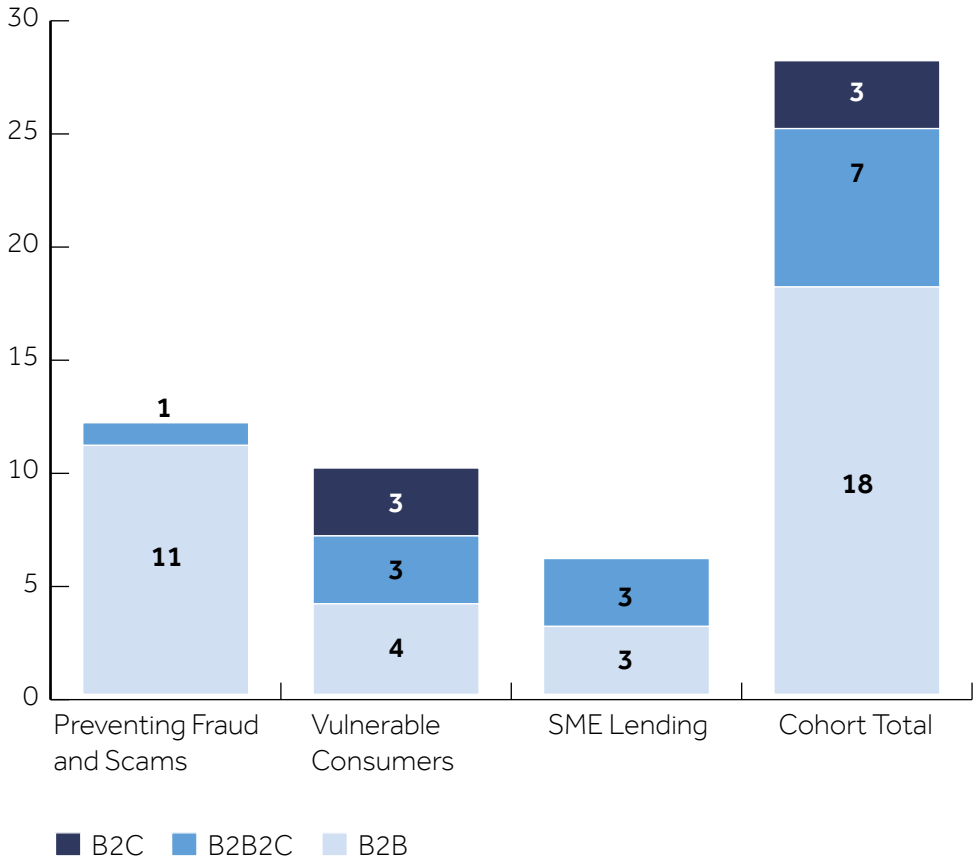
***Figure 4 – frequency of core innovative technology underpinning team projects***




**3.35** The majority (64%) of solutions tested in the pilot were business to business (B2B) models. All 3 of the direct business to customer (B2C) solutions were developing PoCs in the supporting vulnerable consumers category. Interestingly over the course of the pilot, 2 of these participants pivoted their business models to 'business to business to consumer' (B2B2C), as a result of collaboration with expert mentors and other participants. These participants concluded that a more viable path to launch would be by partnering with an existing financial institution and delivering their services through an established entity, rather than launching a standalone product directly aimed at consumers.

**3.36** The FCA's regulatory sandbox has a requirement for applicants to demonstrate direct consumer benefit. Innovative firms which focus on RegTech or B2B solutions have, at times, found it challenging to quantify the downstream benefits for the end customer. The digital sandbox pilot demonstrated the potential for this type of service to provide additional support to these types of innovative firms which have traditionally been underserved by our offerings.

***Figure 5 – business models of participants at outset of pilot***




**3.37** A full list of successful pilot participants, and a summary of the solution they tested during the pilot is available in annex 2.

# 4 Summary of pilot feedback

**4.1** Over the course of the pilot users were required to complete an entry, exit and data survey, as well as participate in a series of recurring touchpoints with the Working Group, who also analysed aggregated usage statistics from the digital sandbox platform. These and other feedback mechanisms were embedded to enable quick changes to the pilot to be made in-flight, and to ultimately inform future iterations of the digital sandbox. In addition, an independent evaluation of the pilot was undertaken by a third party.

## Usage statistics from platform

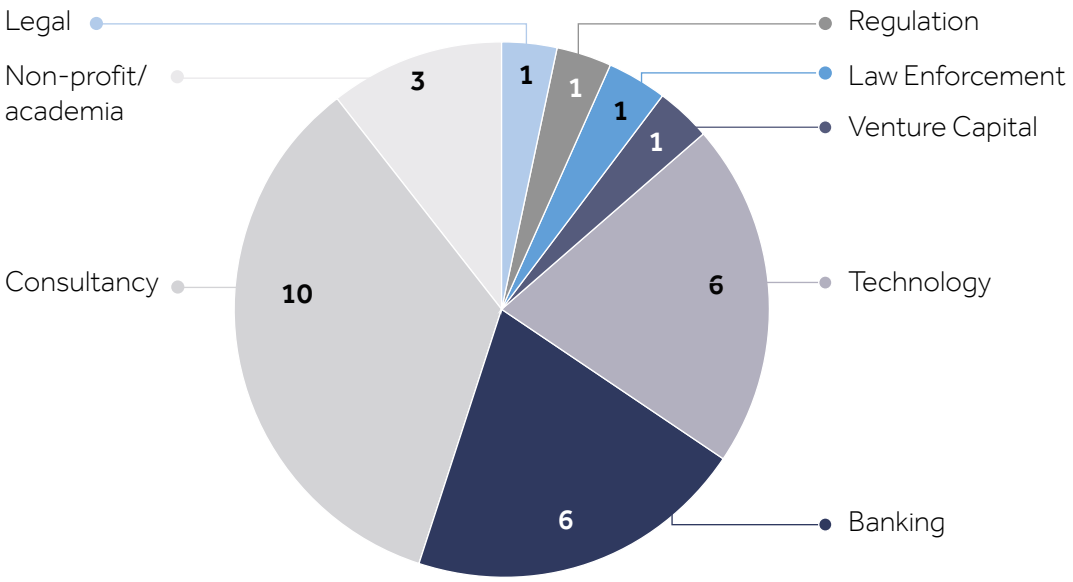
### Users

**4.2** In total, over 850 users created accounts on the digital sandbox platform over the course of the 11 week pilot. This total figure includes team participants, expert mentors, and observors. Mentors were individuals who had self-nominated in response to the FCA and CoLC's expression of interest. They were required to have skills and experience relevant to the pilot and be willing to contribute to supporting the teams pro-bono. Observers were individuals who had registered an account on the platform, and who had access to view the showcases and send messages to participating teams.

**4.3** The digital sandbox pilot landing page had over 5,000 unique hits, and over 600 viewers watched the live showcase sessions held in December and February. Recordings of each of the sessions are available on the digital sandbox platform. The final demonstration day videos can be viewed on YouTube.

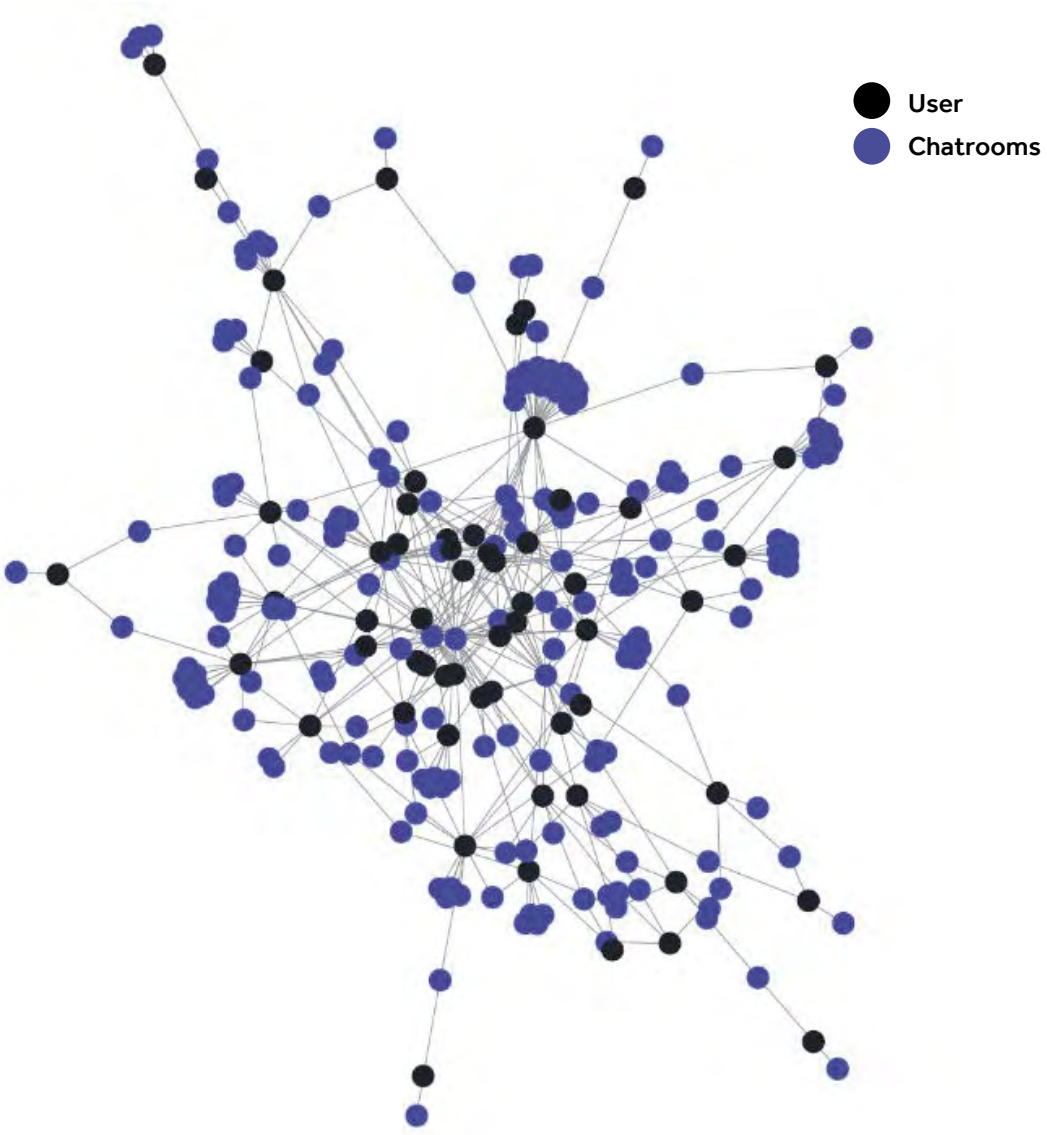
**4.4** There were a total of 29 expert mentors registered on the platform providing support to the teams. Several mentors engaged with multiple use cases.

***Figure 6 – breakdown of mentor expertise by sector***



**4.5** The digital sandbox pilot platform also allowed users to send private messages or begin group chat channels amongst any registered users. Over 115 unique communications channels were created between users over the course of the pilot.

***Figure 7 – diagram of chatroom and user clusters***




**4.6** There were 100 active participating team members across the 28 pilot teams, with an average of 3.6 individuals per team. Weekly active users (including observors) averaged 92.

## Data usage statistics

**4.7** The synthetic data assets were accessible in 2 ways; via API calls or by launching an instance of the integrated Jupyter notebook for large volume data processing. Aggregated usage statistics from the platform show that in total there were:

- 853,980 data API calls
- 652 launches of the Jupyter notebook and 1280 large data queries

**4.8** Interestingly, only a total of 20 API calls were made to third party services in the API marketplace. The services users tested integrations with were Stripe, the FCA Register, FinHub, Open Bank Project and TransferWise. In the entry survey users noted that they expected the API marketplace to be the least used feature, however the minimal level of usage was well below our initial expectations. Reasons for this are explored in the lessons learned section.

## Feedback surveys and questionnaires

**4.9** Participants were required to complete a total of 3 feedback surveys:

- an initial onboarding survey to capture user expectations going into the pilot
- a data survey specifically capturing users' feedback on different aspects of the data assets
- an exit survey during the final week of the pilot

### Initial survey summary

**4.10** Going into the pilot, 92% of participants stated that the feature they anticipated using the most was the synthetic data. The feature participants anticipated using the least was the API marketplace (52%) followed by the Jupyter Notebook (20%). When followed-up through interviews, users noted a preference for accessing the data in their own platforms and systems as opposed to the integrated Jupyter notebook.

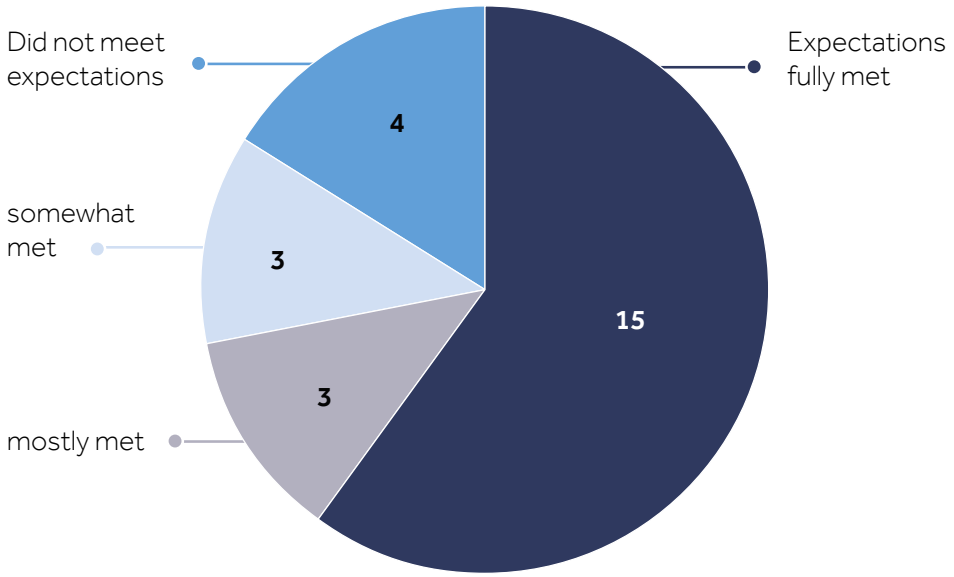
**4.11** There was also a significant desire for collaboration going into the pilot, with 56% of teams stating they intended to collaborate with at least one other team in the cohort. Participants stated that the adjacent nature of their solutions meant there would be valuable learning opportunities and the potential for partnerships.

### Exit survey summary

**4.12** In the final week of the pilot, users were required to complete an exit survey.

**4.13** Participants were asked whether the tools and services provided in the pilot had met their expectations overall. Of the firms who completed the survey, over half stated that their expectations had been fully met. Twenty five of the 28 participating organisations completed the surveys in full. Results in this report are based on completed survey findings only.

***Figure 8 – participant responses to whether pilot had met expectations***

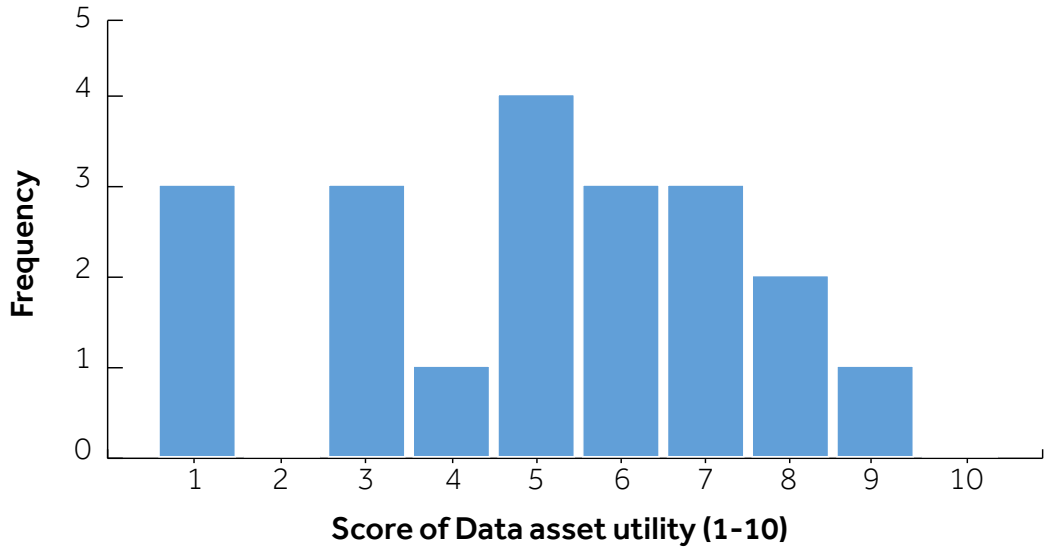



**4.14** The highest satisfaction rate was in the preventing fraud and scams area, whereas SME lending was polarised between being fully met or not met. Those in the vulnerable consumers segment were more evenly distributed amongst the categories.

**4.15** The consistent theme amongst respondents citing a lower satisfaction related to the synthetic data assets, and in particular that the data lacked the necessary attributes to meet the participant's testing requirements, such as granularity, referential linking, or lack of typologies within the data.

**4.16** Participants were asked to provide an overall score for the data assets, with 1 = no utility and 10 = extremely useful. Both the mean and median scores were 5, suggesting that for most participants the data assets provided some utility but did not fully meet their testing and development requirements. It is also important to note that in addition to the synthetic data assets provided, teams were also given access to other publicly available data sets. This included an anonymised sample of credit card data from a Czech lender in 1999, as well as other aggregated data sets from financial institutions, Google Location data, and others.

***Figure 9 – respondent's data asset utility score (n=20)***



**4.17** Interestingly, we observed several examples of teams using the synthetic data assets as a foundation for further tailoring and the creation of their own data sets bespoke to their needs. For example, some teams added additional typologies and behavioural patterns of fraud, which did not exist in the synthetic data provided. In the feedback responses, 2 respondents suggested that the data assets should be a 'living resource' which participants could incrementally add to, for benefit of the wider cohort.

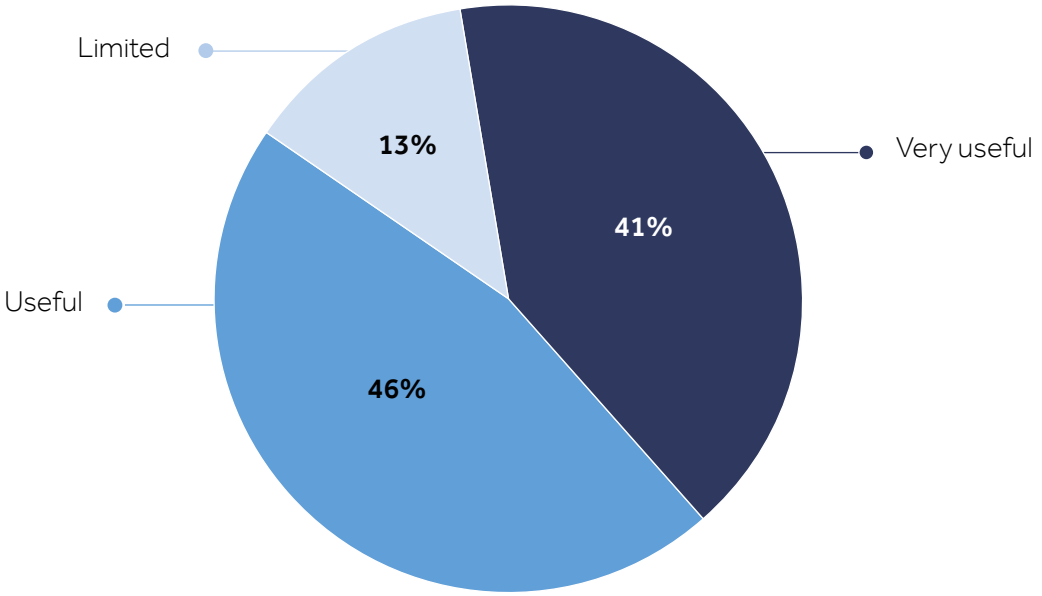
**4.18** This demonstrates the value that would be added by greater levels of specificity in the data, and we were able to get a measure of what might be required through some of the feedback collected in the data survey. It also demonstrates the utility of a network platform that can enable community development and iteration of assets, something we are keen to explore further.

## Mentor and collaboration feedback

**4.19** During the course of the pilot around two-thirds of participants collaborated with other teams. Given the similarity of the PoCs they were developing, this collaboration almost entirely took place within use case areas, rather than across the wider cohort. Teams found the showcase sessions particularly useful for learning about their fellow participants' projects and solutions, and several were keen to form an alumni network post-pilot. Feedback indicated that teams would have preferred more structured introductions at the very start of pilot as there was initial inertia around being first movers and a lack of clarity how they should be collaborating.

**4.20** The expert mentor system also received overall excellent but, in some cases, mixed feedback. Mentor support was requested 88 times, with a mean of 3.5 requests per team and a median of 2. Only 3 teams did not request support from any mentors, and 2 teams requested support from 10 separate mentors. As part of the exit survey, 87% of respondents stated they had found the mentor's support either 'very useful' or 'useful':

***Figure 10 – percentage of responses on the value of the support provided by mentors***



## Feedback on improving the sandbox

**4.21** Users' feedback in response to improving the digital sandbox focused heavily on improvements to the existing core tools/features as opposed to the introduction of additional new features. When explicitly asked about additional features, only 21% of respondents suggested additional features and all of these centred on new channels for advice, mentoring or guidance, for example regulatory feedback or industry panel sessions.

**4.22** During the soft consultation phase, tools such as a technical conformance testing suite or certification regime to aid procurement and due diligence were suggested as possible features the FCA should consider. However, there was no feedback from pilot participants that these would have been useful features for their stage of development.

**4.23** Overall, 84% of responders cited the pilot as having accelerated their product development. While it is difficult to ascertain or quantify this level of acceleration, analysis shows that the biggest factor was ready access to data in developing an early stage PoC. Several participants estimated they had accelerated their development by 4-6 months, with one going as high as 18-24 months, largely by negating the initial need to identify and work with an industry partner to get a PoC off the ground, or sourcing or generating data themselves.

**4.24** Those who found that the pilot tools had not accelerated or improved their development generally noted that this related to the data being insufficient in some way for meeting their use case. This was as a result of one or a combination of the following:

- Insufficient detail in the data, particularly a lack of relevant typologies or behaviours they required to model their solutions. For example, transactional spending patterns indicative of a customer experiencing fluctuating mental health (under the vulnerable consumers use cases) were not seeded into the synthetic data due to the complexity of doing so.
- Required data sets were not available. For example, large volumes of unstructured data such as consumer complaints text, to train and validate natural language processing techniques.
- Data sets not being referentially linked. For example, different data sets had been generated independently, so the behavioural patterns or characteristics of a synthetic individual 'John Smith', would not be consistent with 'John Smith' in a separate data set.

**4.25** However, even with these limitations, participants noted the utility of the data sets for 'bootstrapping' product design. Even where the data could not be used to refine an algorithmic model, there was value in providing the data models, data structures and formats that were representative of what they would be working with in real production environments.

**4.26** Overall, feedback received from the pilot suggests that the core features were sufficient to meet the identified and underserviced needs in the market. In the future, resources should be dedicated to improving the highest value features - the data and collaboration channels - as opposed to exploring or introducing new ones.

## Independent evaluation

**4.27** The pilot was assessed against 5 success criteria set at the start of the pilot. An independent assessment was undertaken using a combination of quantitative survey analysis and qualitative interview analysis. Structured interviews were conducted with 12 participating teams, 3 Advisory panel members, 3 expert mentors, and 10 observers. Further analysis was derived from the exit survey results.

**4.28** This evaluation concluded that the pilot had fully met 2 of the stated objectives and partially met the remaining 3. This corroborates our findings that the pilot has been successful in supporting innovation in financial services, but further improvements to the tools and services, in particular the data, would result in improved outcomes in future cohorts.

***Table 1 – Copy of high-level summary of independent evaluation findings***

| Success criteria | | Status | Summary |
|---|---|---|---|
| Innovation | Role played in encouraging innovation in financial services | Partially met | Pilot encouraged innovation within financial services, however this was mainly around longstanding issues rather than tackling issues that had been specifically created by the Covid pandemic |
| Speed of development | Role played in enabling quicker testing and development of proofs of concept | Met expectations | The vast majority of participants reported that the pilot had enabled them to develop their projects more quickly |
| Collaboration | Role played in fostering collaboration, facilitating diversity of thinking and creating an ecosystem of key stakeholders | Partially met | Teams were able to interact with other teams, mentors, and observers during SME and industry panels, but they expected more facilitated collaboration. |
| Testing environment features | Effectiveness of the feature set in stimulating and accelerating innovation | Partially met | Teams found the synthetic datasets useful notwithstanding some limitations. IDE and observation features were moderately useful, and the API marketplace was least useful |
| Sustainable future | Role played in informing and assisting the design and future operating model of a permanent digital sandbox environment | Met expectations | The pilot validated ongoing demand for this type of service and all stakeholders provided numerous suggestions for future improvements. |

**4.29** The independent assessment based the innovation success criteria on addressing issues 'created' by Covid, rather than 'exacerbated' by Covid. This difference in terminology accounts for this criteria being deemed 'partially met' by the independent assessment, while other feedback sources from participants described in the report suggest that the pilot did succeed in meeting the expectations of this criteria.

# 5 Lessons learned

**5.1** The following section has been broken down into lessons learned with corresponding actions. As previously mentioned, the FCA and CoLC intend to run a second cohort of the digital sandbox over the next 12 months and will look to implement these actions in the next stage of the initiative's development.

## Use cases

**5.2** Fraud, vulnerability and SME banking are significantly different domains with different requisite data sets and stakeholder networks. The quality of synthetic data assets suffered due to the sheer breadth of data sets required to service participant's needs across the 3 topics. Despite the breadth of data sets that were made available, at least 25% of survey responses wanted additional data sets, such as biometric data, unstructured data, and audio or voice data.

**5.3** Securing sustained engagement from the mentor and observer network would have benefitted from the members of these networks having similar areas of interest and expertise, rather than being spread across 3 diverse areas. It would have been easier to penetrate deeper into large organisations with a narrower focus of use cases to reach more senior decision makers.

**5.4** The use cases were established based on the criteria discussed above, CoLC and FCA priorities, and the strong desire to tackle issues that had been exacerbated by the pandemic. In the future, the use case selection process would benefit from a more formalised approach with consistent input and weightings from internal and external stakeholders. An external advisory panel could agree use cases for the cohort, which would sustain buy-in from those organisations throughout the cohort development period.

### Recommendations

- Select one 'focus area' per cohort, to ensure sufficient data quality and a more highly engaged community.
- Better identify and receive buy-in from financial services firms (banks, investment funds, FinTechs), and their senior stakeholders, that are essential to the initiative's success including professionals in innovation, banking, products and partnerships, IT, investment, and fintech - and engage with them at the start of development.
- Conduct a series of roundtable discussions or external workshops with senior industry stakeholders in order to refine the 'focus area' and identify use-cases that meet an industry need and clear success outcomes.

## Synthetic data generation

**5.5** At the DataSprint (described above), data requirements/behaviours were defined **in parallel** to data generation due to the short duration of the event. In hindsight, future events should clearly outline requirements and typologies beforehand, perhaps at an entirely separate Sprint, as very different skill sets and expertise are required for the two tasks.

**5.6** Industry participation at the DataSprint was much higher than expected but lacked the sustained participation that is key to producing high quality data assets. As a crude example, 10x data scientists for 1 hour does not equate to the output of 1x data scientist for 10 hours.

**5.7** The sheer volume of participants made it challenging to manage. The contributions of participants were varied, with much enthusiasm displayed to participate, however the participation levels were impacted by existing work commitments.

**5.8** Ultimately, it did not deliver the finalised data sets required for the sandbox but did lead to good engagement. Future versions of similar efforts should include participants to the sandbox being selected prior to a DataSprint, so they could define their requirements and work with data scientists to build bespoke synthetic data.

### Recommendations

- Separate the data definition/requirements/typologies exercise from the generation of the data sets, and include the eventual users of the sandbox, who have been accepted earlier in the process.
- Expertise for the first exercise (data definition/requirements/typologies) is best convened externally, from industry.
- For the second exercise (sustained data generation) it is difficult to rely on pro-bono external resource. It should be run internally under the digital sandbox banner, perhaps via procurement or dedicated FCA resources.

## Synthetic data

**5.9** Access to synthetic data was identified as the most important feature by far, by most participants (92% ranked data or Juptyer as the highest expected use). However, the average 'rating' of the utility of the data was 5/10 (this also includes additional data sets outside of synthetic data).

**5.10** For most teams, data did not fully meet their needs. As noted in the use cases recommendations, the sheer breadth of the required data across 3 different domains meant that there was a trade-off between volume and quality.

**5.11** It is important to recognise that synthetic data is a nascent and enormously complex field. Even with world-leading expertise contributed by the Turing Institute, the Working Group was unable to create the required richness across so many data sets.

**5.12** Feedback also indicated that better data documentation was needed at the outset of the pilot to avoid resources being required to answer participants' data questions. This documentation should:

- Provide clear descriptions of how the synthetic data was modelled, eg what typologies exist, in what ways is it representative of real data.
- Give an understanding of which data sets are linked to each other.
- Be published before teams apply. Teams had unclear data expectations going into the pilot while writing applications which contributed to the misalignment of expectations and teams having to pivot solutions.

**5.13** Many teams would have preferred full access to data outside of the platform, for example via S3 buckets. The IDE limited teams to a certain number of calls per day and they could not ingest data onto their own platform. Allowing this approach would also reduce the cost of computing ie the cloud credits required to be provided by the Working Group.

**5.14** However, ring fencing data assets behind the IDE enabled the monitoring of usage, prevented data leakage and ensured sustained engagement with the pilot. This contributed to the development of a supporting ecosystem for participants.

**5.15** Feedback from participants and observers (particularly financial institutions as potential customers of the solutions) indicated that synthetic data can only enable product validation if the synthetic data assets are commonly understood and recognised by industry - otherwise they are more limited to serving a role in PoC and product development. Industry participants can't evaluate the effectiveness of potential solutions if the input is not understood at a technical level.

### Recommendations:

- Explore the option of allowing participants full access to the data assets outside the platform environment, rather than ringfencing behind the IDE. This would give participants the option to remove data assets and work on local machines and systems but would require appropriate safeguards and agreements regarding access.
- Create narrower data sets with richer depth, to fully meet the needs of participants.
- Engage large financial institutions when developing data sets. A commonly understood industry data set may enable 'proof of value' assessments whereas a private ringfenced dataset cannot – modelling should be open source.
- Create better documentation for data sets. Modelling approach needs to be clearly articulated in advance of opening applications.

## Application process

**5.16** Having an external industry panel of experts provide their opinion on what was genuinely innovative and commercially viable was valuable in making accurate assessments of which proposals were supported.

**5.17** The platform features worked well for the online application portal and enabled efficiency gains compared with other previous application systems used for FCA innovation services.

**5.18** The application process was not subject to the same level of scrutiny and rigour as Regulatory Sandbox applications. This is appropriate on the basis that digital sandbox participants represent a significantly lower risk of harm, given that tests are conducted

on synthetic data, and no limited permissions or authorisations are granted to test with real consumers or live data.

### Recommendations:

- Keep the online platform for receiving and managing applications.
- Keep the external advisory panel format, with a changing format relevant to the cohort subject matter.

## API marketplace

**5.19** The API marketplace had minimal usage and was ranked as the least valuable feature in surveys (52%). This suggests there is minimal additional value to be added from collating existing API services into a one-stop-shop. There are a few reasons why this might be the case.

- Firstly, it is best practice for Tech firms to publicly expose their test APIs via developer portals, so many of the most well-known FinTech and RegTech services are readily accessible.
- Secondly, the limited cohort size meant creating a bespoke API purely for the pilot would be exposed to a maximum of 30 companies, likely not returning dividends on the resource cost to set-up.
- Finally, most participants were not at a mature enough stage of development to find value in accessing the services of a third-party provider.

**5.20** Almost no organisations external to the pilot (6 enquired, 2 joined) offered their APIs in the marketplace, despite open calls for interest from the Working Group on the option to do so.

### Recommendation:

- API marketplace should not be prioritised if it represents a significant cost to maintain.
- Enable industry self-onboarding with minimum standards, instead of investing resources into creating a curated marketplace.

## Collaboration – chat functions

**5.21** Going into the pilot 88% of teams expected collaboration to be at least moderately useful and 56% stated they planned on collaborating with another team during the pilot.

**5.22** There were over 115 unique chats created, of which approximately 35 were active on a weekly basis. However, overall analysis is that the chat feature was underutilised and could have delivered more value with a better implementation.

**5.23** Participants tended to make initial connections via the chat function and then take collaboration off the platform to more familiar services, for example to Slack, WhatsApp, private email, LinkedIn or Microsoft Teams. There was substantial

qualitative feedback that participants were experiencing 'platform fatigue' resulting in this reversion to the use of existing and familiar communications channels.

**5.24** There was very little platform usage by observers of the chat feature. Usage was mostly limited to teams and mentors even though observers also had access to this feature.

**5.25** At the beginning of the pilot, the platform did not have a notification feature to email observers, mentors or participants when they had a new message. In addition, participants did not have access to personal contact information for mentors or observers which made it difficult to reach out outside of the platform. This meant that many communications went unanswered in the first weeks of the pilot and this delayed collaboration, and some development. A notification was installed mid-way through the pilot in response to requests from participants and as part of our real-time improvements.

**5.26** The collaboration platform did not include a list of all observers and their roles or organisations. This made it difficult for participants to identify useful parties to contact.

### Recommendations:

- Explore additional features to improve the chat experience, such as email notifications for unread messages, user activity levels, default channels upon joining, user groups (ie investors, data scientists, use cases) and 'add all' options.
- Explore options such as integrating LinkedIn or e-mail information for back-up communication.
- For shorter, more intense collaboration events which may be hosted on the digital sandbox in the future (such as TechSprints) in addition to the embedded chat functionality, consider using alternative collaboration tools, such as Slack, that participants are familiar with and have a better feature set for rapid and frequent messaging.
- To engage observers more, reverse the onus of initiating contact onto pilot participants by creating a community directory of observers. Participants are more likely to be active and willing to contact observers that they are confident are interested in their particular use case.

## Collaboration – user journeys

**5.27** Feedback suggested that participants would have benefitted from a more structured journey throughout the pilot, for example organised theme-wide meetings throughout the early stages of the pilot. The creation of a community calendar for greater clarity of milestones and collaborative events would further support this objective.

**5.28** Observer engagement was relatively minimal outside of showcase sessions. The user experience of observers should be improved with the recognition that little changes on a day-to-day basis from an observation perspective so periodic defined milestones should incentivise observers to maintain engagement.

**5.29** It straightforward for observers to contact teams, but in practice this feature was very rarely used. On the other hand, it was difficult for teams to identify and contact observers, for example due to the absence of a community directory.

### Recommendations:

- Publish a 'calendar' before the application process showing what the key milestones are: touchpoints, participant 'team leader' meetings, demonstration sessions, feedback sessions to clarify user expectations etc.
- Enable a community directory where teams can contact observers. Tighter cohorts with narrower use cases should also contribute to a more engaged observer community.
- Rollout the 'live feed' feature to all users so that they can see activity happening within the sandbox. This is a continuous stream of public updates indicating what has been updated on the platform, for example changes to showcase spaces, new media, new observers, etc.

## Collaboration – mentors

**5.30** The 'mentor' function had a wide spectrum of outcomes. Some very valuable connections were made but some participants had a frustrating experience with delayed responses or limited feedback and input.

**5.31** Feedback suggests that there was broadly the right mix of mentors' fields of expertise. The only notable missing group was incumbent financial institutions. Overarching feedback from participants was the concept was correct, but that it was let down by execution. In the future it was suggested that LinkedIn integration would be beneficial and that mentors should be able to be contacted via email rather than the chat function on the platform.

**5.32** Unresponsive mentors were a recurring problem. Clearer expectations of the role that mentors are signing up to should be made to ensure they have skin in the game and remain active.

### Recommendations:

- Allow mentors to be contacted via email and/or LinkedIn, rather than, or in addition to, the chat feature.
- Create a Terms of Reference for mentors - clarifying expectations of them.
- Spend more resources on finding appropriate mentors before launch day.

## Showcase sessions

**5.33** Participants were required to present their solutions to a broad audience of pilot observers at the mid and end points of the pilot. These showcase sessions - designed to create a 'bridge' between innovators and the broader financial services ecosystem - generally worked well. There were, however, differing degrees of preparedness and quality of presentations among participants, with several requesting a more standardised format and preparation material. Overall, however, 92% of participants agreed the sessions were valuable.

### Recommendations:

- Create firmer expectations and more structured feedback on demonstrations before final showcase sessions.
- Keep format and midpoint and end of cohort demonstrations sessions. The midpoint was helpful for enhancing cross-cohort collaboration and focusing minds.

# 6 Next steps

**6.1** Since we have concluded the pilot, the Kalifa Review of UK FinTech has recommended that a permanent digital sandbox be created to encourage further collaboration within UK financial services. The Review suggests that these efforts be led, in the short to medium term, by the FCA.

**6.2** Based on the feedback collected throughout the pilot and an independent evaluation, and to support the Kalifa Review recommendation of a permanent digital testing environment, we will:

- Run a second cohort of the digital sandbox in late 2021.
- Further iterate and improve the digital sandbox testing environment by incorporating the lessons learned from the initial pilot and making the suggested improvements to the platform.
- Expand on the use of the digital sandbox testing environment to highlight the opportunities and value it contributes to the financial services ecosystem.
- Focus the efforts of this second cohort around the theme of Sustainability/ climate change to support the UK's green finance ambitions ahead of hosting COP26 in November 2021.
- Explore, with industry and other stakeholders, viable sustainable operating models for a future, permanent version of the digital sandbox.

**6.3** We will work closely with industry and other stakeholders throughout the Summer to identify the specific sustainability use cases that the second cohort of the digital sandbox will look to address and to identify a permanent future operating model.

# Annex 1
# Data assets available to participants in the digital sandbox pilot

| Grouping | Data | Description | No. of Records |
|---|---|---|---|
| Reference | Synthetic Entities | Synthetically created entities that are used as reference data to link all of the datasets together within the ecosystem. Statistically representative of UK Companies House. | 600K |
| Reference | Synthetic Individuals | Synthetically created individuals that are representative of the UK population, based on Office of National Statistics data. | 7 million |
| Banking | Transactions | Retail and Wholesale banking data representing transactions for consumers across the UK through the creation of 5 fictional banks. We have also looked to represent some of the behaviours that we would see as a result of the Covid-19 crisis. | 400 million |
| Banking | Device Data | Device data that represents the devices used for faster payments across the UK. Can be used to detect fraudulent behaviour. | 5 million |
| SME Lending | Loan history | Issued loans to SMEs. | 65k |
| SME Lending | Credit Card History | Business credit card statistics | 190k |
| SME Lending | Current Account History | Summary statistics about entity current accounts | 750k |
| SME Lending | SME Directors | Directors and officers for entities in the ecosystem | 2.5 million |
| SME Lending | Covid SME Lending | Covid-19 Business Impact based on applications made for Covid-19-specific government relief or loans | 500k |
| SME Lending | Factoring | Factoring information for SMEs who have factored their accounts receivable | 500k |
| SME Lending | Profit and Loss | Profit and loss statements for SME entities | 500k |
| SME Lending | Accounts Receivable | A list of details around invoices which inform credit decisioning | 38 million |
| SME Lending | Lending Providers | Institutions providing lending in the market, both traditional and alternative | 350 |

# Annex 2
# Overview of participants' digital sandbox solutions

| Detecting and Preventing Fraud and Scams | |
|---|---|
| **Faculty AI** | A tool which combines cutting edge outlier detection methods with robust and accurate explanations, enabling a user to uncover and investigate anomalous transactions quickly and easily. |
| **Sedicii** | An AML solution that allows financial institutions to securely share knowledge about clients or transactions without disclosing any underlying data or information. |
| **EalaX Ltd** | A solution that creates digital synthetic twins of real financial data which can then be used to detect fraudulent patterns and complex problems that are being experienced during Covid-19. |
| **Callsign & HSBC** | A solution to protect customers' digital identities through a unique positive identification approach, tackling the rising issue of online banking fraud, without compromising on customer experience. |
| **Financial Network Analytics** | A solution that uses neural networks to establish the usual patterns of behaviour between organisations and individuals to highlight anomalies in order to detect fraudulent payments. |
| **Synectics Solutions** | A solution that uses both supervised and unsupervised machine learning techniques to analyse transactional data in close to real time, to determine a risk score to aid in the identification of account push payments fraud. |
| **MPC4AML** | A solution that uses a Privacy Enhancing Technology (PET) known as Secure Multi-Party Computation (MPC) to run risk scores on a transaction network of data from multiple banks. |
| **Futures** | A solution that uses AI-based approaches to provide a real-time risk assessment across a network to detect consumer behaviour indicative of fraud and scams. |
| **Norbloc** | A solution that uses blockchain to allow for a secure and GDPR-compliant sharing of verified KYC files across multiple institutions in real-time to create a single profile per customer. |
| **LexTego** | An open source solution to detect fraudulent and money laundering activity within a financial ecosystem – in real time for selected typologies. |
| **IT2 Fraud Signals - Trust Stamp, Cifas, Lloyds & OneBanks** | A solution that uses biometric data to create an identity token that can be used to match, de-duplicate and verify across institutions while protecting the users personal identity information. |
| **Onespan** | A solution that will combine machine learning with the development of rules by experts to identify types of fraud. |

| Improving the Financial Resilience of Vulnerable Consumers | |
|---|---|
| **PrinSIX Technologies** | PrinSIX is focusing on detecting vulnerability within credit applications, testing and deploying dynamic onboarding journeys that identify applicant vulnerability flags and trigger highly personalised digital assessments to improve customer outcomes. |
| **Qpal** | A digital assistant focused on financial wellbeing, powered by Open Banking, that automates financial decision for consumers. |
| **Elifinity** | A solution to create a two-app ecosystem to allow users to share their financial data with debt advisors and for debt advisors to directly send their solutions to the consumer via an app. |
| **Automated Regulatory Monitoring** | An AI platform aimed at SME's that identifies the drivers and characteristics of vulnerability and provides a series of preventative measures to avoid negative outcomes for consumers. |
| **PORTABL** | A solution to create a 'Freelancer Risk Score' that fairly represents independent workers who might miss out on financial inclusion because they are not well served in typical markets. |
| **FLANK** | A solution to assist vulnerable young people who have borrowed money from friends and family in managing and repaying these loans. |
| **Applied Blockchain** | A solution to allow a range of lenders to assess the credit risk of a borrower without requiring direct access to private and sensitive financial data by using privacy-preserving technologies. |
| **DirectID** | A solution that uses transactional data to predict the probability of default based on an individual's historical and predicted cash flow. |
| **Kalgera** | A solution using financial transaction data to identify indicators of impaired financial decision-making and to help firms more efficiently and quickly allocate resources to support vulnerable consumers. |
| **Amplified Global** | A solution that uses a 'cognitive risk engine' to assess the level of comprehension of information and apply a 'cognitive risk score' to a consumer, and natural language processing to appropriately simplify presented information. |

| Improving Access to finance for Small and Medium Enterprises | |
|---|---|
| **Fluence** | A solution that uses natural language processing AI to interpret and analyse financial applications and claims handling processes in order to automatically interpret new applications and claims. |
| **Open Banking Reporting** | OBR creates analytical assessments, metrics, forecasts, network and threat models for SMEs and lenders to build successful businesses. |
| **Sota Signal Analytics** | A solution to signal risk in a business's financial condition to reduce uncertainty and create greater confidence in a lending counterparty. |
| **Company Watch Ltd** | A solution to enable finance funding providers to predict, analyse and risk assess the ability of a small business to repay credit within certain time periods. |
| **Fractal Labs** | A solution that uses Open Banking to create a cashflow forecast in order to help assess the eligibility of SME's for small working capital loans. |
| **Untangled Finance** | A solution using tokenised assets on a blockchain to enable simple, cost-effective and transparent ways to securitise SME loans and invoices. |

Pub ref: D495



12 Endeavour Square London E20 1JN
Telephone: +44 (0)20 7066 1000
Website: www.fca.org.uk