Lessons Learned from the 2021 DFS Techsprint on Digital Regulatory Reporting in Cryptocurrency

On March 1-12, 2021, the New York State Department of Financial Services (DFS), in conjunction with the Alliance for Innovative Regulation (AIR) and the Conference of State Bank Supervisors (CSBS), hosted the first techsprint ever organized by a U.S. state financial regulator.[1] Techsprints—hackaton-style events in which teams compete to help solve problems identified by regulators—have become increasingly popular since the UK Financial Conduct Authority pioneered the format in 2016. DFS saw the March Techsprint as a way of jump-starting its efforts to transition to digital-first supervision by gaining more frequent access to the data and documents regulated cryptocurrency firms regularly produce to DFS. DFS also hoped the solutions put forward by the Techsprint teams would help DFS make more use of IT in analyzing the data it receives, enabling the Department to better leverage its resources.

DFS stands to benefit from increased digitization of data reporting and analysis across all of its supervisory activities. But for purposes of the recent Techsprint, DFS decided to focus on the regulation of cryptocurrency companies in particular, because DFS-regulated cryptocurrency firms are small in number and relatively technologically advanced, making the industry "low-hanging fruit" for trying out some of the needed changes. The sponsors of the 2021 Techsprint hoped that any improvements in the supervision of cryptocurrency companies that resulted from the event could also be applied, in some form, to other industries DFS oversees, such as insurance and banking.

Nine teams comprising 83 participants competed in the event, counting among their members technologists, regulators, consultants, lawyers, and managers from across the cryptocurrency community. Almost 500 people joined the virtual "Demo Day" held on March 12, at which teams presented their proposed solutions to the problem statements the Techsprint's

[1] More information about the Techsprint, including videos from demo day, can be found at https://www.dfs.ny.gov/techsprint. The team that organized the Techsprint included Matt Homer, Matt Siegel, Olivia Bumgardner, Seema Shah, Justin Friedman, Evelyn Castillo, and Cici Matheny from DFS; David Ehrich and Randy Repka from AIR; and Dan Berkland and Brennan Zubrick from CSBS.

organizers had posed, and prizes were awarded. The proposals were varied and imaginative, and the organizers are now working together to advance some of the proposals to the next stage of development.

This document identifies the salient lessons organizers learned from the Techsprint: both the competition itself and the months of planning leading up to the event, during which organizers held five workshops with U.S. and international regulators, cryptocurrency professionals, academics, and others. The organizers hope that the information they gained from the Techsprint process will be useful to anyone interested in conducting similar events in the future, and to the community of financial regulators more generally.

Overarching Themes

Reflecting on what we learned from the 2021 Techsprint, we can identify a few overarching themes:

- **Now is the time.** The Techsprint fortified our belief that now is the right moment to be working on Digital Regulatory Reporting (DRR) approaches for the U.S. cryptocurrency industry. Many U.S.-based cryptocurrency firms provide a steady stream of data and documents to multiple federal and state regulators, and keeping up with these reporting obligations is a significant cost of doing business. The data reporting function cries out for more, and better automation. Also, because this industry is new, and based on assets whose valuations and trading patterns are, as of now, still quite volatile, we can expect to see continued boom-and-bust cycles; service outages as companies struggle to meet spikes in demand; and, very likely, more cybersecurity incidents. In this environment, companies' risk profiles can change substantially from month to month; so regulators need a window into their supervised entities that is more up-to-date, and more amenable to quick, automated analysis, than is now common, if they hope to achieve the basic regulatory goals of safety-and-soundness, market integrity, and consumer protection.

- **Change is needed.** The Techsprint process also made it clear that regulators cannot simply burnish the supervisory tools they have employed for decades, but will need to develop entirely new approaches to ongoing supervision. For example, we need to think about getting direct, real-time access to company data, in some instances, to replace a periodic "push" model of data production. Regulators will have to weigh the risks against the benefits of using real-time or other quick-turnaround data, because such data may be less fully vetted, and thus more susceptible to errors, than periodic snapshots are.

- **Collaboration is essential.** Both the pre-Techsprint workshops and the event itself clearly demonstrated that neither regulators nor industry participants will likely be able to design the needed DRR methods on their own. Multilateral collaborations will be required. Regulators will have to speak for consumer protection, safety-and-soundness, and other regulatory goals, and for the practical needs and capacities of their agencies. Cryptocurrency executives will be needed to explain the cost/benefit tradeoffs of making data available on a certain schedule or in a certain form, and to help ensure that DRR

approaches are, to the extent possible, valuable to the companies that participate in them, rather than being seen as simply an exogenous burden they need to comply with. Technologists will also be needed, to explain the objective constraints on DRR, including the costs of adopting proposed new data-production regimes or new types of data analysis.

- **We must think about the market holistically.** Any approach DFS takes to DRR should also take account of the cryptocurrency companies that are not regulated by DFS. For example, as the pre-Techsprint workshops confirmed, observers see a clear regulatory need for better tools to identify and arrest attempts at market manipulation on cryptocurrency exchanges. But, because most large cryptocurrency exchanges are not regulated by DFS—and because market manipulation can often operate across platforms—a solution to this problem will likely need to consider non-DFS-regulated exchanges in addition to those DFS oversees.

- **Incrementalism is fine, and likely preferable.** In the cryptocurrency industry, the current state of regulatory reporting is so far from being optimal, and the technical and institutional barriers to the optimal use of data are so formidable, that it is hard to imagine any regulator adequately addressing these issues with a single effort. It is more realistic, and probably more desirable, for regulators and supervised entities to implement a few improvements that can be accomplished in the short run with relatively low cost/benefit ratios; learn from those implementations; and iterate. This sort of incrementalism may not come naturally to regulators who are used to a more "rifle shot" approach.

* * * * * *

Problem Statement Learnings

To help craft the problem statements that were the basis for the competition, DFS held its five external workshops in November-December 2020, which were aimed at identifying the most salient risks posed by the cryptocurrency industry, and the areas in which DRR approaches are most acutely needed if regulators are to adequately address those risks. The following takeaways emerged, based on the broad categories of risks identified by participants in the workshops:

- **Illicit financing.** Regulators and cryptocurrency firms would like to be able to better use company data—including, to the extent possible, detailed transaction data—to identify and stop instances of illicit financing, including attempts to launder money; attempts to finance terrorism; and unlawful transactions with sanctioned nations or persons.

- **Market manipulation.** Many regulators and industry participants believe market manipulation is a particularly serious problem in the cryptocurrency industry, due to factors including the price volatility and thin trading volumes of many cryptocurrencies; the fact that cryptocurrencies are traded on numerous exchanges with varying degrees of regulatory oversight; and the paucity of information-sharing among those exchanges.

- **Consumer harms.** Another issue on the minds of cryptocurrency industry observers is consumer protection. Cryptocurrency prices tend to be very volatile, yet there are no clear

standards for determining whether a cryptocurrency investment is suitable for a given individual. Payments with cryptocurrencies are often irreversible, and cryptocurrency senders or recipients looking to avoid being identified can often do so; so industry players often observe "scams" against their customers, e.g., fake emails aiming to induce them to transfer cryptocurrency or reveal private keys to bad actors. The irreversible quality of cryptocurrency payments means errors can cause permanent consumer losses. DRR may be able to address some of these concerns.

- **Insolvency and other financial risks.** DRR may be able to provide more timely information on the financial health of regulated cryptocurrency companies. Particularly in times of market stress, regulators would benefit from knowing more about the day-to-day financial condition of their regulated companies, rather than having to rely on traditional quarterly and annual financial statements, which are not well-matched to the volatility of the cryptocurrency markets.

- **Engaging in cryptocurrency business without a license.** Technology could also help regulators identify companies that are providing cryptocurrency services without the requisite licenses or other permissions.

The sponsors ultimately chose four problem statements for the Techsprint. These statements did not touch on every risk or every promising DRR target we had identified, for reasons including the practical limitations of the Techsprint format and the need to winnow our list of problem statements to a tractable length. The competing teams were invited to address a specific problem statement of their choice, or to combine more than one of the problem statements as they saw fit. The final problem statements were as follows:

TECHSPRINT PROBLEM STATEMENTS

I. How can DFS achieve real-time or more frequent access to company financial data from virtual currency licensees and receive early warning signs of financial risks to the companies or their customers?

II. How can DFS obtain real-time transaction data from its licensees and automatically analyze the data to safeguard against illicit financing risks?

III. How can DFS use tools such as natural language processing, machine learning, and artificial intelligence to identify risks by processing and analyzing supervisory reports that are submitted by licensees in a wide range of formats?

IV. How can DFS use technology to facilitate information-sharing among licensees to help them more quickly identify and stop scams, ransomware strikes, and other criminal enterprises that put licensees and their customers at risk?

Solution Development Learnings

Participants in the March 2021 Techsprint were organized into nine teams, which developed potential solutions over a twelve-day period. The resulting proposals embodied several

innovative approaches to DRR, including some that had the potential to lead to implementable solutions. Although the teams addressed an array of problems, and took very different approaches to them, several common takeaways emerged:

- **Identifying DRR's benefits to regulated companies.** The benefits of any DRR project to DFS must be balanced against its benefits to the regulated companies. DFS has substantial authority to require that regulated entities submit additional data as needed. But, as a practical matter, DRR tools will be better-received and more enthusiastically implemented if they are clearly useful to the companies that are being asked to provide the data.

- **Determining the location of the data and the analytic engine.** In implementing DRR tools, DFS must weigh the merits of housing the data, analytics, or both at DFS against the merits of locating one or both functions off-site. The latter option would include, e.g., leaving the data with the regulated entities and enabling DFS to "pull" data as needed. It might also mean making the data available (with appropriate confidentiality protections) to all DFS-supervised cryptocurrency firms—and possibly even a broader group—through a decentralized network. Alternatively, the data, analytics, or both could be housed at a third-party organization, such as a trade association, joint venture, or self-regulating organization ("SRO"). The decision of where to locate the data and analytics will be driven by considerations such as confidentiality; DFS's IT capabilities; and the tradeoff between a DRR solution's utility to DFS and its benefits to the regulated firms. As one team noted, if the data and analytics were kept at an SRO, the arrangement could shift the maintenance and development costs from DFS to industry, and could allocate those costs to the individual SRO members in way that avoids disproportionately burdening smaller firms.

- **Thinking hard about data sources and formats.** DFS needs to consider several important questions regarding the data that should be included in any DRR approach to cryptocurrency. For example, will the regulated entities be asked to provide their data in a standardized format? Standardization facilitates cross cross-company analysis, but can increase compliance costs, particularly for companies that need to produce similar data to several different regulators. Also, to what extent can the analysis of company data be augmented with data and tools that are available from public and commercial sources, including analyses provided by blockchain analytics firms and payment data that is freely available on public blockchains?

- **Not making the perfect the enemy of the good.** A DRR regime must strike the right balance between working with the data DFS already receives and acquiring the data it would most like to have for analytics purposes. Production of "optimal" data sets may engender costs for firms that could outweigh the benefits to be gained from the additional data in some cases, and truly optimal collection of data is likely to be a moving target in such a fast-changing industry. Solutions proposed to DFS by the Techsprint teams demonstrated how much value can be gained by simply by gathering all the existing data into a single system; running the data against existing risk flags, such as blockchain addresses associated with suspicious activity by blockchain analytics firms; and presenting the information in a single place, so DFS examiners can get the full benefit of the data they already collect.

- **Making better use of transaction data.** Many of the Techsprint teams' proposals aimed to make better use of the transaction data from DFS-regulated companies—seemingly an underused resource. One proposal demonstrated how DRR tools can enable a regulated cryptocurrency company to query and share with other regulated companies red-flag information on particular individuals and their transactions, by using cryptographic techniques such as zero-knowledge proofs, which can allow for the sharing of information about customers (in this case, a "proof" that they have also been flagged by another party) without revealing any identifying information about the customers or entities themselves.

Process Learnings

The March 2021 Techsprint also taught DFS much about the process of running such an event, including lessons that DFS and other regulators can tap into if they decide to host techsprints in the future:

- **Focus on the problems.** Techsprint planners should focus their efforts on identifying the problems they are looking to solve, resisting the temptation to try to solve the problems themselves. Regulators who are deeply familiar with the problems being studied may assume they already know at least the basic categories of possible solutions, and may unconsciously design the techsprint around those categories. That can preclude the discovery of surprising solutions, which is the techsprint's core purpose.

- **Engage stakeholders early.** One of the most valuable steps organizers took in planning the 2021 Techsprint was to hold the pre-Techsprint workshops. The workshops served their primary purpose of identifying possible problem statements for the Techsprint, but also created a diverse network of people in the cryptocurrency community who knew about the planned Techsprint and supported its goals. These "friends of the Techsprint" helped publicize the event, and made themselves available for informal consultations as planning of the event went forward. Some participated in the Techsprint themselves.

- **Determine resource dependencies.** Organizers of a techsprint should try to be as specific as possible about what they need from others for the event to be a success. If you need data sets as fodder for your techsprint teams, consider exactly what sorts of data you will need, and in what form. If you need participants with specific skills and professional backgrounds, make those needs clear to yourself and potential participants.

- **Identify best approach to team formation.** Anyone organizing a techsprint should consider whether the teams will be assigned or can be formed by team members. Assigning teams helps ensure that each team has members with the requisite skills, and may achieve benefits by combining people with more diverse perspectives. On the other hand, some participants may feel more comfortable entering a techsprint if they can work with friends or colleagues, and some may work better with those they know and have worked with well in the past.

- **Consult with others about running a techsprint.** DFS had not run a techsprint before the 2021 event and benefited tremendously from obtaining input and advice from public and non-public organizations that had substantial experience in planning and implementing techsprints.

The first time out, seriously consider consulting with entities that have conducted techsprints in the past.

- **Evaluate data needs.** If the teams participating in a techsprint will be working with data during the event, organizers should think carefully about where the data will come from, and how it will be produced and organized. That means leaving enough time to gather the data needed for the techsprint, recognizing that data-gathering can be time-consuming, particularly if some of the data has to be anonymized and/or subject to legal waivers, confidentiality agreements, and the like. Organizers should also consider whether teams will be free to use "outside data" not provided by the event's organizers—a practice that could provide increased scope for creativity but could also be perceived as unfair.

- **Plan around the tech platform.** Techsprint organizers will likely want to use any one of a number of software platforms for coding aspects of their techsprint. Two lessons emerged from the 2021 Techsprint relevant to these platforms. First, organizers should consider making their platforms available to participants at least a few days before the event, so coders who are going to participate in the techsprint can get up-to-speed on the platform's use before the event begins. Second, organizers should remember that the work product of the techsprint teams will likely reside on the platform; so there should probably be a transition period during which a copy of the platform remains open after the techsprint, so the teams' work can be organized and archived.

- **Plan for continued action after techsprint.** A techsprint's organizers also need to think about the next steps after the event is over. For example, how will organizers further develop and, ideally, implement some of the solutions that the techsprint developed?

A propos of this last bullet, as of the publication of this Lessons Learned document, DFS had begun the process of building consensus among its executive team about how to transform the proposals submitted by the Techsprint teams into a plan of action. The path to expanded implementation of DRR may be daunting, but our regulated industries and the threats and risks facing them continue to change, day by day. DFS's only real option is to change along with them—even if change is sometimes hard, often iterative, and always imperfect.

A final thought for staff at other regulators who are considering holding a techsprint themselves: DFS found that these events can have cultural benefits for the regulators that undertake them, apart from any direct benefits from the competing proposals. At regulatory agencies, a techsprint can not only help solve technical and practical problems, but can also send an important internal message that the organization values experimentation and innovation.