Increasing Spectrum Data Sharing and Transparency with a Secure, Extensible Data Fabric

Dan McCloskey
US Government Division
Palantir Technologies Inc.
Denver, CO
dmccloskey@palantir.com

***Abstract*—Regulators are faced with challenges to increasing spectrum data sharing and transparency, such as engendering trust in industry to share data and establishing more robust and standardized methods for making use of data when it is shared. Palantir Technologies discusses how today's existing commercial software can accelerate the future of spectrum data sharing. Proven, secure commercial software is rapidly deployable to reduce the technical and procedural barriers for sharing competition sensitive spectrum data, offering regulators a more complete picture and greater situational awareness of spectrum usage.**

Keywords—data sharing, interagency data sharing, data transparency, spectrum management, spectrum usage, spectrum demand analysis, cloud, commercial software

I. Introduction

Regulators tasked with managing spectrum repurposing and sharing rules face continued challenges with accessing the data they need to make informed decisions. Today, wireless industry participants with information critical to upcoming regulatory decisions are not sharing data of sufficient granularity, quality, and frequency to be of use in feasibility studies and related data-driven assessments of current and potential demand and usage of spectrum. Regulators are faced with two primary challenges to increasing data sharing and transparency: (1) engendering trust in industry participants and incentivizing their participation in data sharing efforts; and (2) establishing a more robust and standardized method of receiving, synthesizing, and making use of data when it is shared.

Conflicting interests and trust asymmetries among industry participants cause hesitancy in data sharing and transparency. To get an accurate picture of spectrum usage and demand, regulators require data from multiple industry competitors, which is often competition sensitive and highly proprietary. Regulators need methods of assuring industry participants that their data will be secure and inaccessible by competitors before they are willing to share it.

Regulators also need more standardized and streamlined methods of integrating and operationalizing data once they are shared. Currently, industry does not view sharing data with regulators as worth the perceived effort and risk. As it stands, data sharing often involves manual translation of information from Excel spreadsheets into fillable forms, and with little consistency in place regarding the scope, format, and frequency of data. Reducing the effort required to share, analyze, and secure proprietary and sensitive data lowers the barrier to participating in market research or regulatory analysis.

Palantir Technologies ("Palantir") presents this abstract to detail how existing, rapidly deployable commercial software can both:

- Reduce the technical and procedural barriers impeding the ability to share high-frequency, granular, and competition sensitive spectrum usage data with the Government
- Streamline and automate the processes necessary to integrate disparate data from industry into a single source of truth for informing feasibility studies and related decisions, without compromising security

We build software that lowers barriers to access and use of data, particularly for complex inter-organizational collaborative efforts. This paper will explore the specifications and outcomes of a **secure cloud platform for sharing data among the interagency, and between industry and Federal spectrum regulators**. We also present selected case studies where we have deployed our software to deliver similar outcomes to those required for spectrum regulatory analysis and data calls.

II. Secure Cloud Platform for Spectrum Data Sharing

A secure cloud platform for spectrum data sharing would make it easy and safe for industry participants to share useful data. In turn, regulators gain a more transparent and complete picture of current spectrum use and can make more accurate, data-driven predictions of future spectrum demand. This enhanced awareness of the current state and potential future state will ultimately improve decision-making in the generation of sharing rules in an increasingly congested spectrum environment that includes complex and dynamic spectrum coordination between Federal and commercial users.

Such a platform must be able to ingest, store, secure, organize, analyze, and operationalize data to enable effective regulatory analysis supported by an accurate picture of Federal and commercial spectrum usage. Based on our experience deploying software to meet complex data sharing challenges — including on and between classified networks — we recommend the following features for a spectrum data sharing platform:

An accredited, granular security architecture trusted for interagency data sharing workflows in highly competitive and classified environments. To incentivize industry to share their data in a useful fashion, regulators must be capable of ensuring that data will be protected regardless of the context in which it is used**.** This can be accomplished with a data sharing platform that secures data with: end-to-end encryption; user provisioning and authentication/authorization framework; role-based, purposed-based, and/or classification-based access controls; data anonymization and obfuscation; and governance structures such as data access request checkpoints. Granular permission structures can be defined down to the cell or row level of a spreadsheet.

Ability for industry participants to retain full control over their data. When ingesting and sharing data within a cloud platform, organizations should be able to maintain control over their data. With granular access controls, data stewards can expand or revoke data access in highly flexible ways. Data stewards can audit and track data permissions in the platform to ensure that not only the right individuals have access to their data, but also for the right reasons. Further, each individual data contributor can host their raw data in a private environment, strictly limited to users from the given institution. In this environment, users can transform their raw data — including applying any necessary privacy and access controls — and perform data quality checks to ensure data is ready for broader use.

A sustainable and secure framework for ingesting data from multiple sources into a single, secure environment. Manual processes for data sharing, such as by e-mailing spreadsheets, present information security risks, and it is often labor-intensive to make data usable when sent in such piecemeal fashion. Instead of relying on inconsistent, insecure systems and processes, regulators can leverage a secure data sharing platform that streamlines and automates complex processes associated with integrating information from across industry. An interoperable cloud platform with out-of-the-box data connectors, data validation and health checks, data synchronization, data modeling, and data catalog features could transform the spectrum data sharing process. It would reduce the risk and effort associated with manual data sharing and would open up the opportunity for more automated, ongoing data sharing efforts over time.

Analytic and operational tools for situational awareness, predictions, hypothesis testing, and decision support. To truly make use of the spectrum data foundation envisioned in this paper, regulators need that data to be accessible and usable by analytic and operational tools. An effective cloud platform for spectrum data sharing will have integrated modeling and simulation tools that enable regulators to: (a) visualize and understand the current state of spectrum use across industry; (b) model and predict future use; and (c) simulate the impact of potential regulatory changes.

III. Selected Case Studies

We offer the following sample case studies to demonstrate how industries facing similar challenges have leveraged data sharing platforms powered by our software to bring industry participants - including competitors - together for improved outcomes. This software is in use today bringing together members of industry in healthcare, defense, aviation, life sciences, and more.

A. Data Sharing & Anonymization at National COVID Cohorting Collaborative ("N3C")

Palantir software—configured as the National COVID Cohorting Collaborative ("N3C") Data Enclave—is supporting a collaborative COVID-19 clinical research effort across 70 National Institute of Health ("NIH") Clinical and Translational Science Awardees (CTSA)—serving as the largest centralized patient-level data asset of COVID-19 clinical data in the world. The Data Enclave integrates 6.5M+ patients' data from 50+ institutions (over 7.3B rows of data) and provides secure access to 1,000+ authorized researchers across dozens of institutions and the public. The Data Enclave has supported over 50 scientific publications (published or in progress) and conference presentations.

To protect data and ensure only users are receiving proportional access with their intended use, N3C researchers must request access to relevant data sets (at a specified level of sensitivity) in the Foundry platform. Once approved, researchers can use secure, collaborative workspaces and scalable analytical applications to uncover insights. Beyond initial data access approval, Foundry also supports NIH governance workflows that provide oversight for data use. For example, the N3C Data Access Committee reviews and approves or declines submitted requests for download or proposed publications—all within the platform.

B. Skywise – A Connected Platform for the Flight Industry

Skywise is Palantir's partnership with Airbus to deliver an open data platform for the aviation industry. Skywise integrates in-flight, engineering, and operations data from Airbus and participating customer airlines and suppliers into a single, comprehensive data foundation and platform. Today, over 140 airlines worldwide contribute data and leverage Skywise capabilities. With granular security controls they can trust and the openness and interoperability needed to make data sharing easy, Airbus' suppliers and global airlines integrate and share data, which used to be locked into silos, into the Skywise platform; e.g., maintenance, flight management, aircraft monitoring and safety, and other critical workflows. In this way, Skywise not only powers the operations of Airbus' partners, but also allows Airbus to collaborate with them directly in the same environment to ensure the timely delivery and safe operations of its aircrafts.

C. Syntropy – Collaborative Data Ecosystem for the Healthcare Industry

In order to better understand each patient, leading Health Care Organizations (HCOs) must accelerate their ability to

integrate and utilize patient data while simultaneously safeguarding their privacy. Syntropy assists HCOs in this goal by licensing them access to Palantir Foundry, a fully managed cloud-based platform-as-a-service for governing, structuring, and harmonizing HCO data. The Palantir Foundry platform empowers HCOs and their collaborators to derive insights that help them better understand and care for their patients. Those insights are contextually reincorporated into a centralized data asset that grows and improves in quality over time.

End-to-End Governance: The instant data enters Foundry, it is stored and accessed under zero trust principles. Data access and use is comprehensively audit tracked. The platform's accredited security architecture and granular controls ensure that sensitive, proprietary data (including PII/PHI) are protected at all times.

Data and systems integration using innovative data management technology to bootstrap provider initiatives with high-trust data: Syntropy establishes deep partnerships with providers focused on bootstrapping internal initiatives (research, operations), in order to: (a) get direct access to structured and unstructured (e.g. clinical notes) data coming from the source with full context and provenance; then (b) structure and harmonize data into a common, institutional model, with direct engagement from clinical SMEs.