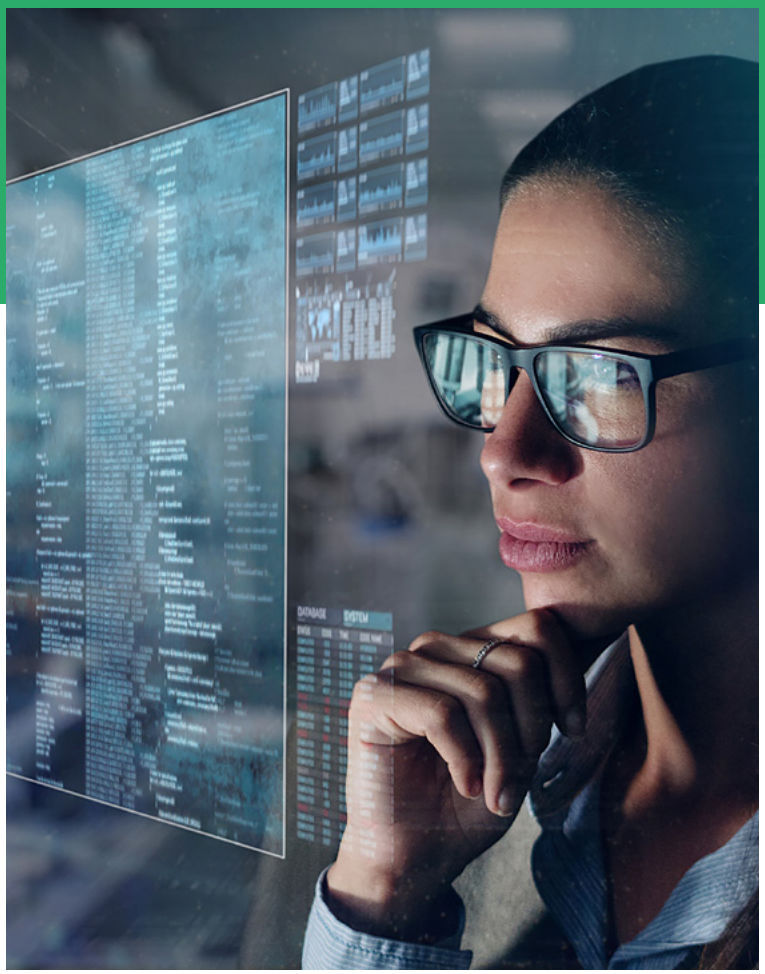

THE DATA PROBLEM WITH COMPLIANCE ARCHIVING

Evolving from EIA to DCGA highlights the need for a new approach to communications data infrastructure

THETALAKE

TABLE OF CONTENTS

INTRODUCTION

The FSI communications compliance landscape has expanded rapidly over the last two decades, with ever-increasing ways to communicate via eComms, Voice, Visual, IM, mobile, and Unified Communications tools like Microsoft Teams, Zoom and RingCentral. It's not just the communication "channel", though. Particularly with UC tools, there are multiple channels offered within a given platform. For example, Microsoft Teams enables communications like IM, Voice, SMS, Document Sharing and Video to name a few. It goes even deeper when looking at chat or SMS features within a UC platform channel that also include emojis, GIFs, and videos etc.

In summary, there is a large amount of data generated in almost any modern business communication. A single conversation can continue across multiple channels, multiple features within each channel, and all communications data are subject to regulatory scrutiny. In addition, regulators have made it clear that it is not the "channel" that they are concerned with, but the content of the business record. This has created a big challenge for IT and Compliance leaders who are charged with protecting their firms from regulatory and reputational harm as well as data security threats.

Enter DCGA (Digital Communications Governance and Archiving). The additional complexity and volume of data has driven the need for this new market category that expands well beyond email archiving. Text-based capture and supervision doesn't cut it anymore as more modalities are adopted. DCGA is described by Gartner as an approach to leveraging solutions that give companies methods to monitor and enforce regulatory compliance and corporate governance across a growing communications ecosystem.

In this paper, we will break down the data problem across current archive infrastructures, and how to fix the data problem to get your organization on the right path to a reliable, actionable and future-proof communications data infrastructure.

DATA SILOS

The problem with data silos stems from capturing different types of communications having multiple sets of infrastructure, creating gaps and lags in conversations. Take eComms (like email or text) and Voice for example, with vendors in the space offering or specializing in one or the other. Conversations happen across mediums, and have for a while. But increasingly, those conversations continue to be meshed across channels, with gaps being created because of two different sets of data from the same conversation. This also creates lags because they have timing differences in terms of when the data is reconstructed together downstream. Another critical issue, which is a challenge for procurement and budget-owners, is that the segmented approach doubles the costs to the firm.

Example: if there is a violation that occurs while using Voice, but that meeting also included IM, there is a risk that the Voice violation would be missed, or vice versa.

INCOMPLETE DATA

Solving data completeness challenges are at the heart of the new DCGA market segment. This impacts virtually any area of compliance, from regulatory capture challenges during audits and investigations, to downstream use cases like search, supervision, surveillance, e-discovery. If your firm is not confident in your surveillance tool, one primary reason today is working with incomplete data. If your firm is being scrutinized by regulators who aren't satisfied with meeting their requirements, in most cases it's due to incomplete data. As mentioned in the introduction, regulators are concerned with the content of the conversation, not the channel. The content needs to be complete for these reasons. Capturing all of the data with an infrastructure that was designed for email is problematic.

Example: Starting from a chat, then to a call, to a meeting, to a file, to an email, to a text, then back to a chat all in one conversation would be viewed as common in today's modern communications ecosystem. A conversation within a UC platform like Microsoft Teams can be loaded with images that have text inside of the images, a chat that might have a .wav file, drawings with comments and labels, shared documents with sensitive information, and of course, the video, voice, SMS features that run on the Teams infrastructure.

DATA STORAGE AND ACCESS INFLEXIBILITY

As the market has evolved from EIA (Electronic Information Archiving) to the new DCGA (Digital Communications Governance and Archiving), another critical problem to solve is data storage and access. As the modern communications landscape evolves, the data volume is ballooning with ever-increasing places that people are creating communications data. These communications are much larger now and will continue to grow. Think about images inside of chat or documents, or video - more volume, more detail. As an undesired, and in some cases unanticipated result, costs become increasingly unscalable in those proprietary storage environments from traditional vendors. The long-time vendors with archives have focused on the storage of your data, and notably, a cost associated with accessing it. Whether it's a fee per GB, or needing to pay a vendor's managed services team to do it because they are the only ones who can, the reality is that your data is not yours until you pay for it, and the vendor retains their customer by holding on to the data. In short, there is a data access problem in DCGA.

Example: A firm needs a segmented set of data to reside in Germany with the rest to be stored in the USA, then making sure the data that they want vs. don't want is stored or not retained respectively. This requirement can be very costly if even possible to accomplish with traditional vendors.

DATA ANALYTICS MYOPIA

The industry is well-aware of the many pain points in the communications surveillance segment of DCGA. Most supervision and surveillance tools can't meet expectations because they are restricted to the limited and narrow confines of available data, preventing the required criteria to meet desired outcomes from fragile analytics. The gaps, lags, fragments and lack of context and access to data leads to remarkable time and resources spent piecing together and reviewing conversations. As previously mentioned, limited access to data from traditional vendors further marginalizes the capabilities of surveillance and other analytics tools. The most common problem that surfaces after implementation is that firms take an overly-aggressive approach to try to detect violations, resulting in a high volume of false positives and a lot of wasted time. The alternative is that reviews become extremely narrow in an attempt to save time, which increases the risk of a true positive being missed. Neither are acceptable for DCGA.

Example: A firm is managing their UC platform, they need to capture a specific set of data, but there is no detection set or authentication needed when external users are involved, so there is a possibility that something is being missed based on the way the system is set up. If a downstream surveillance vendor is not involved in the capture integration or not considering the implications at the data layer, then the fundamentals are missed and accuracy is unattainable.

REACTING TO THE DATA PROBLEM VS. TAKING ACTION

With the myriad of challenges that the industry faces with the data problem, not addressing the problem at the fundamental data infrastructure hierarchy will doom firms to the same cycle of finding an issue and reacting to that issue, with a long list of recurring problems as new problems and costs continue to add up with no end in sight.

Example: Costs are increasing, data remains incomplete and fragmented causing search and supervision issues from a lexicon-based tool, a significant amount of time is being spent on reviewing false positives, violations are missed and a firm adds another communication channel from the same vendor because the off-channel policy isn't working.

SUMMARY OF THE DATA PROBLEM

The Data Problem from EIA to DCGA

Data Silos for eComms archives and voice recording create gaps, lags, and doubles costs

UC requires true voice & eComms together for complete "conversation" with in-sync delivery

Data incompleteness for meshed modalities

Images with text, chat with .wav, drawings with comments and labels etc.

Visual, voice and text support is a must for record keeping, search and analysis

Data storage and access inflexibility

Data hostage, unscalable costs, poor geo / data sovereignty support, inability to use data / AI

Data analytics myopia based on modality, limited context, and limited to communication content

Fragile analytics, FPs as well as significant missed TPs, missed UC setting controls equals increased risk surface area

Data reaction without actionability / response

No reduction in poor outcomes or governance improvement

FIXING THE DATA PROBLEM

While many of the issues that create the data problem are all too familiar, completely changing your systems, revamping your tech stack, starting from scratch all at once isn't an option for established firms, and Theta Lake would not recommend that approach. But there is one important aspect of your infrastructure that you can change today, and start realizing meaningful progress immediately.

Fixing the data problem starts with Capture. If you have a proper Capture infrastructure in place, not only will that solve the conversation missed/lag/delayed/incomplete issues, but all other downstream use cases will be more effective, reliable and produce better outcomes for your firm.

What to look for in a Capture solution

- For the channels that you capture, **get all of the data for any modality/communication feature.** This includes Meta-data, context, and enrichment
- There should be **no fragmentation or delays in timing**. No missing pieces delivered later from other sources
- **Monitoring and reporting on controls**, configuration drift and health matter
- Ability to easily **Reconcile records upstream and downstream** for source of truth and auditing is a must

HOW THETA LAKE CAN HELP

More Coverage of Unified Communications and communications tools, features, and modalities. Theta Lake's unique approach to ensuring that you can capture all of the data from conversations (not just eComms or voice) with in-sync delivery will solve for data completeness.

Better Reconciliation tools, reporting, and evidence. Reconciliation is absolutely critical to ensuring that your firm is capturing data as expected, and it's a priority for regulators. Theta Lake's easy-to-use and flexible reconciliation can be leveraged in multiple ways.

1. It can be viewed from the Home Tab in our Unified Capture UI.
2. A summary report can also be run from the same Home Tab.
3. You can leverage Theta Lake's API to create your own tools and compare data from source platforms against Theta Lake.

Health & Controls Monitoring of capture and tools. Theta Lake's Home Tab in the Unified Capture UI shows the captured data metrics, where the data is routed, and also provides easy-to-use health checks to ensure there are no issues with configuration or settings changes that will disrupt data capture.

Repeatability & Scale for any tools or communications. Theta Lake has dedicated teams for capture integrations, ensuring timely, complete and quality capture happens. What's more, Theta Lake can route your data to any preferred location, whether that's an existing archive to multiple storage locations, or to you your firm's location. This means the data is truly yours to store where you see fit.

Easier set up and maintenance. Unlike other firms, you can set up a channel integration in 5 consistent, easy steps.

...All of this fixes the data problem that fixes archiving, search, discovery, supervision, surveillance and improves compliance outcomes

Learn More:

How can Theta Lake can fix your data problem? Schedule a demo today!