Modern Chat Compliance and Security

Addressing the Compliance and Security Gaps in Email Archives and CASBs

THETALAKE

Modern Chat Creates Archiving, Supervision, and Data Exposure Blind Spots. Theta Lake Can Help Your Email Archive and CASB Address Them

Contents

Introduction

An integral part of the modern workplace, chat platforms like Cisco Webex Teams, Microsoft Teams, RingCentral Glip, Slack, and Zoom Chat have become primary information-sharing channels. Their skyrocketing use, due in particular to the adoption catalyst that COVID-19 and resulting work-from-anywhere (WFA) introduced, has exposed new compliance and risk scenarios organizations must overcome.

As these chat platforms are used to facilitate internal communications as well as interface with external customers and partners, organizations have regulatory, privacy, and security obligations for content shared within these channels. Sectors like financial services are no stranger to the capture, retention, and supervision requirements for chat such as those required by the SEC, FINRA, FCA, and CFTC. Most industries must also comply with privacy and security mandates like GDPR, CCPA, HIPAA, and a growing number of additional emerging state laws, since sensitive personal identifiable information (PII) is often shared, and could be exposed, within these chat channels. Failure to archive, supervise, and act on content within these channels can result in compliance and privacy fines, litigation exposure, and broadscale reputational risk. These robust, media-rich communication channels present complex risk and compliance challenges that neither legacy email archives nor cloud access security brokers (CASBs) can handle in-depth, creating risk while driving up the cost of compliance for organizations.

Unique Challenges: Persistent, Ongoing, and Exposed

As a modern communication and collaboration tool, chat platforms function differently than many existing workplace communication tools. Unlike email, which typically has content type or size limitations, chat supports many types of content with no size restrictions. Text, audio files, video files, links, images, gifs, emojis, and even reactions to existing content are all supported and can be used in any combination. Yet challenges go beyond content type; the ways employees use chat platforms are also fundamentally different.

Content in chat is persistent. This means that any content shared within a chat conversation stays in the channel unless it is purposely removed. A conversation can be commented on after days, weeks, months, or even years. A file or link is still accessible even after a conversation has concluded. Therefore, risks and sensitive information continue to stay exposed, or can be shared, unless it is removed.

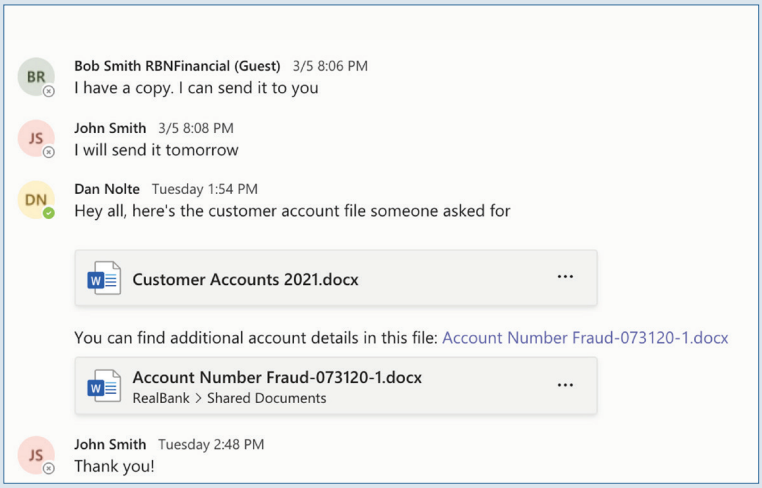
Chat conversations are also ongoing and fluid. As previously mentioned, participants can still reply or react to a chat conversation weeks later. That ongoing conversation behaves differently than email, often with many conversation "threads" or topics spanning days with multiple messages and replies to messages from different participants. Trying to reconstruct it using email-based threading and artificial insertion of "To" and "From" fragments context and creates misalignment in messages between participants that breaks supervision processes. At the same time, the historical, ongoing, and unstructured nature of a conversation easily bypasses blocking systems looking for specific, structured data, which are typically not present in collusion and insider risk behavior.

Examples of Risk Exposure and Repercussions

Recent cases of risk exposure highlights the importance of full visibility and awareness of risky behavior as well as the ease in which modern chat and communication can result in the leak of sensitive information.

Desjardins Data Breach

Desjardins Group, Canada's largest cooperative financial group, was found at fault in a data breach due to their lack of attention in protecting sensitive PII. To easily access and share resources for business purposes, Desjardins employees moved data from a secure location to a shared drive. This allowed a malicious former employee to exfiltrate sensitive data undetected for 26 months.



Unmonitored sensitive file sharing via Microsoft Teams can be the next public data loss example

It's easy to imagine that those files could have been moved to a SharePoint site and then the files or links to them easily shared in private chats, group chats, or even team chat channels. That is the most likely way they would be shared in today's collaboration-first workplace.

Finally, chat content is easily exposed by participants. Not only do organizations use chat to communicate broadly with employees and contractors, chat can also be used with external participants like partners and customers. Its shareable nature makes it easy for any participant to copy/paste, take a screenshot, or even provide a direct link to a specific part of a conversation to others. Chat's support for various content types also obscures transparency when sensitive material is exposed. For example, a link to a SharePoint file of an earnings report can be shared on chat, but it may not be easily apparent, even to participants, that the link exposes sensitive material. Further, what is shared and downloaded from the SharePoint link can be changed later in the SharePoint site to something benign without anyone knowing what was shared at the point in time in the chat. Potentially worse, once exposed, anything shared – a sensitive back-and-forth, a link to a risky website, a video recording, or sensitive file shared in a one-on-one, private, group, or channel chat – stays exposed in the chat, available to any participant to download, screen shot, or share unless deliberately removed. There is also no indication if anyone shared it outside the channel or for how long.

Eight Gaps Legacy Approaches Cannot Solve

Email archives have been the primary tool used for retention and regulatory requirements for electronic communication, while CASBs are increasingly used for aspects of data loss prevention (DLP) and enforcement of acceptable use policies (AUPs) as organization app usage continues to grow. However, these tools were not built to cover the complex risk scenarios of robust chat platforms presented by their myriad of content types. At the same time, previous supervision requirements were fairly simple – regulated organizations merely had to capture content and make it available for search to prove the existence of an archive of record – while today's obligations for proactive supervision have expanded.

If an organization cannot adequately provide compliance and risk coverage, it means they have to disable valuable features that drive productivity and top-line activity while satisfying customer and end-user demand. The alternative is assuming an untenable non-compliant and risky stance.

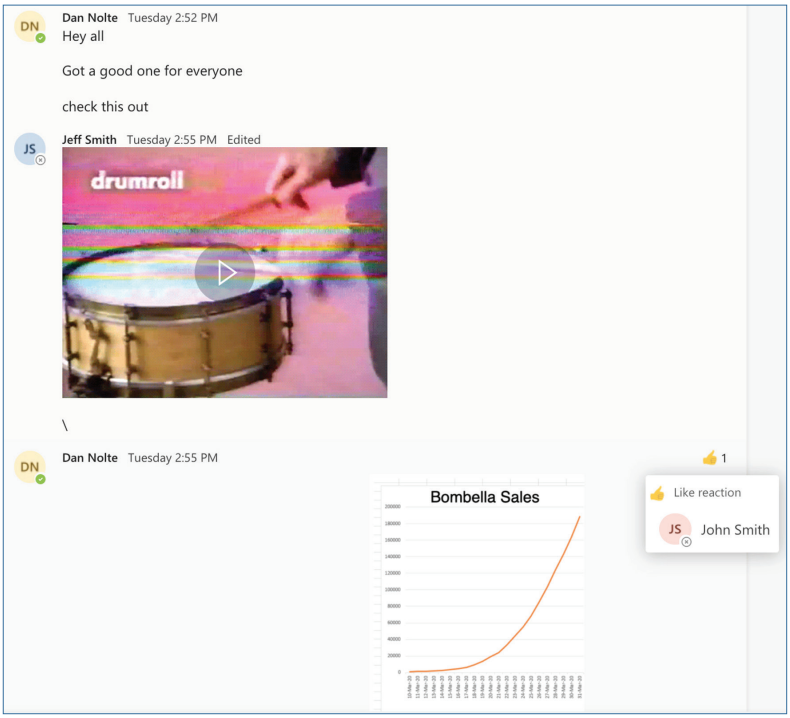
Here are the key ways that email archives and CASBs fail to support modern chat platforms and force regulated organizations to make costly trade-offs.

1. Not all chat components are captured and archived

Legacy email archives are often not fully capable of the API integrations required for capturing modern chat application data, and they may also miss the capture of the newer content types present in modern chat. Built for simpler journaling and SMTP-based content interception (and not the rigors and nuances of working with complex platform APIs), traditional tools often leave blind spots.

For example: only capturing teams and channel chats but missing one-to-one private chat, missing whiteboard and note content in a chat, or, in Microsoft Teams, not capturing the exact file version shared via a SharePoint link at the time the link was shared. Further, built for text-focused content, email archives often miss capturing and archiving emojis, reactions, gifs, images, video files, audio files, and related content in context of the conversation.

Not only is this an archiving failure at the most standard level, it also means adequate supervision is next-to-impossible due to the gaps in content and context in chat conversations.

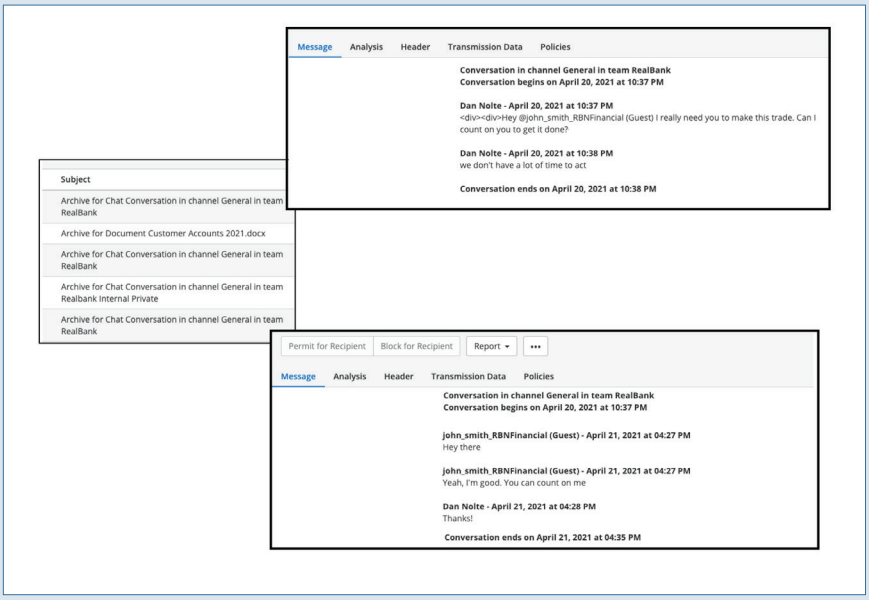


Commonly used Emojis, gifs, animations, and reactions change the meaning and risk of chat conversations and must be captured in context.

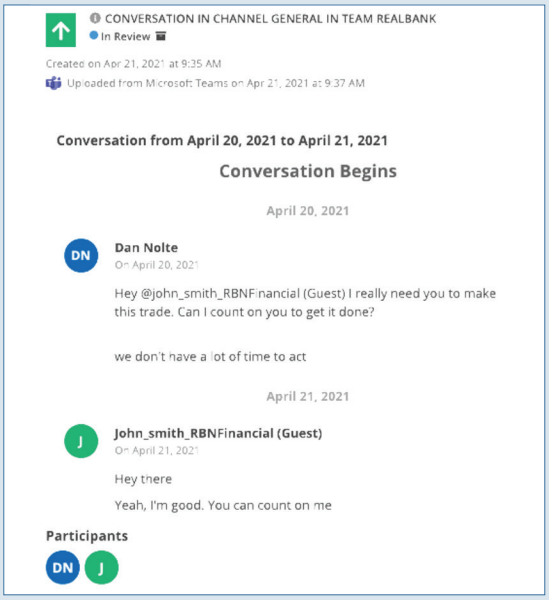
2. Message threading for email archives misses context and breaks search

Even content that is captured may not be put into the proper order and context, creating analysis flaws. Taking the example of email archives again, email threading is a common approach for grouping an email message with all its subsequent replies and forwards. Email archives apply this approach to ingest chat conversations so that they make sense for an email-centric system.

However, this results in inaccurate analysis of chat content and missed context, such as in a chat scenario between colleagues from different time zones or spanning different days. Assume colleague 1 located in the U.S. Pacific time zone sends a chat at the end of the day to colleague 2 located in the U.S. Eastern time zone. For colleague 2, the message is received past the end of the day (i.e. the next day). Traditional threading may not pick up the right time intervals for this conversation, therefore losing context by treating it as two separate conversations where the "To"



Threading views for email archives often break search around improper 'To:' and 'From:' or missed reactions to historical conversations while also making navigation difficult.



In contrast to threading chat as email messages, a native view is easy to search, review, and gives the full, contextual picture to better identify risks in chat conversations.

and "From" are captured incorrectly. The work burden for this inaccurate analysis is compounded in subsequent search and review processes, as reviewers must put in additional work to manually reconstruct the conversation.

An even simpler example is threading around reactions. As noted, an image, file, or even a message may not be risky in itself, but emojis and reactions to the image or chat message may indicate inappropriate behavior. Imagine a prospect list shared in a private chat channel. A "like" reaction on an older chat message may indicate collusion between participants. In many cases such as this one, it is the conversation over time or the context happening around a document that turns it into a risky scenario. Only a native, contextual view that surpasses threading can unlock the ability to effectively supervise these interactions.

In addition to generally being hard to navigate and understand as a chat, the above example shows two sets of messages that look like different conversations. This is the case even when presented in a threaded view in an email archive and in contrast to that same single conversation as it appears in an actual native view to the right.

3. Content is not analyzed or is analyzed incorrectly

Aside from legacy email archives not fully capturing all content and struggling to put it into native context, they also struggle to provide analysis of modern content types (e.g. emojis, reactions, images, gifs, videos, and others), much less analyze them in the chat conversation context they were used in. If the reaction is a "like" on a video file covering an IPO filing where the only chat text is "check this out," systems not built for these content types and unstructured communications are going to miss the potential collusion and endorsement risks in that chat.

As another example, outside of missing capture and analysis of files shared via a SharePoint link in a Microsoft Teams chat, an email archive that might capture a file shared directly in a chat is still unlikely to analyze the file content itself for risk. It is even less likely to mark the chat itself as risky if the file shared in a conversation is a risky file.

4. Search-dependent lexicons and "bolt-on" analytics for stored data introduce search and detection latency while producing unreasonable implementation and ongoing service costs

Legacy archives were originally implemented to capture and store data without the same imperative for more complete and proactive supervision. The need to search that repository for targeted legal or audit purposes was built around basic searches using key words. That lexicon approach evolved into complex and unwieldy burdens for organizations that typically yield high false positives and significant misses for legitimate detections. Furthermore, the

analysis on the repository of stored data has been so slow that it created the concept marketed SLAs on search time performance.

The next wave of attempted solutions was to leverage a new set of vendors that implement compliance analytics using bespoke machine learning. These tools have typically taken years to implement with high service, time investment, and personnel costs at low scale. Often, the detection and workflow needs have changed by the time they are implemented, creating a never-ending cycle of service and re-implementation costs. The resulting ongoing service and subscription fees tend to create prohibitive costs for compliance and introduce friction when adapting to changes in communication types and regulatory requirements. What is worst of all, none of these techniques and tools are designed for the modern content and unstructured sharing models that chat applications use today.

5. Archives have no ability to remediate exposed risks

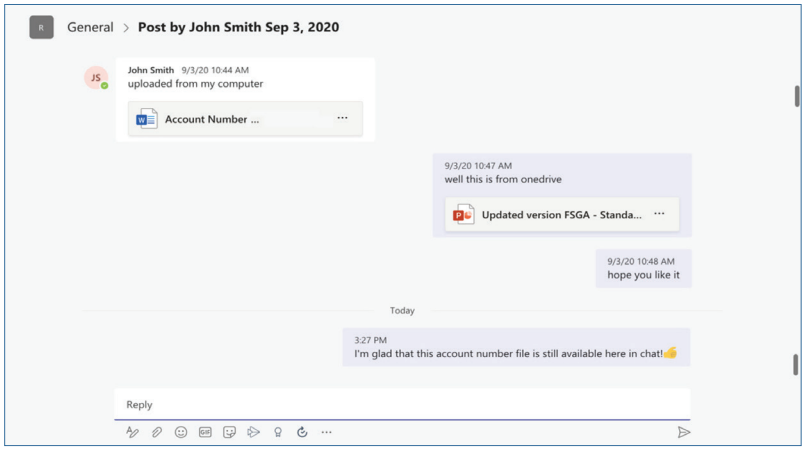
Even when a legacy system detects risk in chat, it lacks the ability to effectively remediate the risk. However, with the persistent and ongoing nature of chat, organizations need to take action to prevent the risk from continuing to happen or exacerbating it through additional shares and exposure.

Consider this scenario: a participant shares a SharePoint document within a Microsoft Teams chat channel that contains sensitive information. Most archive and search-based systems will only trigger when a search is run and can do nothing to directly address that risk. That can create an exposure window of days, weeks, months, or more.

An email archive's inability to remediate or remove content inside persistent chat channels highlights why the old store-and-search-later model is not suited for modern collaboration compliance needs.

6. Legacy archives have prohibitive storage costs

The rapid increase in volume of communications and the vast amounts of files and media included in modern chat exponentially increases the amount of storage required for compliant archiving. Legacy on-premises archives and hosted, private cloud archives have expensive storage models and pricing that are challenged by this increased volume of content. Along with this base challenge, the common integration model of email archives in converting chat content to multiple emails and email attachments increases inefficiencies, triggering more costs for the organization.



Chat is persistent and so is any risk or sensitive information exposed in them. Typical archives and eDiscovery lack needed automated, and instant remediation capabilities.

7. Archives lack a privacy-friendly model for redacting stored sensitive data

One of the challenges in the era of data protection regulations is that the need to both archive and supervise communications creates a unique set of data storage and handling issues that legacy archives are not designed to handle. This includes the need to redact captured PCI and PII so that it is not subject to potential breach and exfiltration. That also includes protecting such data from numerous and growing numbers of compliance and information security staff working in the organization. The majority of tools used for capture, archive, eDiscovery, and supervision provide no mechanism to redact data, properly protect it from exposure, and meet regulatory requirements around data privacy while also meeting communication retention and supervision requirements.

8. Lack of a modern review workspace drives inefficiency and costs

All of the limitations mentioned above culminate in an abysmal review workspace for compliance and information security reviewers. Since legacy archive tools are not designed to properly capture, analyze, and represent the fluid conversation flow and rich media content in modern chat platforms, they naturally also lack a workspace that reviewers need to quickly navigate chat conversations and risks detected within them. The burden is on reviewers to manually piece together chat conversations, likely without the full set of content around that conversation and with misalignments in timeline and chat participants. At the same time, they probably have incomplete detections of risks. All this increases the number of manual resources required to review while simultaneously increasing the likelihood of missing key risks without the ability to directly act on them even if they are found.

API-based CASBs and API-based cloud DLP systems can assist in some data protection use cases but are not suitable for full chat archive, workflow, supervision, and unstructured data loss and risk remediation.

A new area of security capability utilizes Cloud Access Security Brokers (CASBs) and related Data Loss Protection (DLP) tools as part of the stopgap plan for addressing risk challenges for cloud-based chat applications. Only API-based systems should be considered as proxies or agents. Browser-plugins create complexity, management overhead, and unnecessary end-user friction. API-based tools can help in some areas of brokering feature access in chat applications, scanning for sensitive data, and even remediating structured PII, PCI, and PHI data. However, this addition does not solve compliant archiving and deep supervision needs for modern chat tools. Furthermore, they are not at all designed for contextual detections of risks in chat conversations and information exchange that happen over time.

Concepts like collusion; triangulating risk based on the participants, the chat over time, and data shared directly or referenced via links; or references to side conversations are not in the scope of these tools. These tools also lack the reporting, workflow, and integrations to fit more seamlessly with legacy archive and eDiscovery systems, and often exacerbate the problem. For example, sending more chat converted to email files (.eml) to an email archive without fixing the native view, just adds more confused content to the email archive. This can increase missing content, context, and manual work while increasing compliance violations without reducing data loss.

Addressing the Compliance and Security Gaps in Email Archiving and CASBs

Fully addressing gaps in email archives and CASBs require a thorough identification of each gap as well as understanding the systems and processes your organization uses to achieve compliance and security goals. Here is a quick breakdown of what is needed to close existing compliance and security gaps.

1. Understand the features available in the chat tools your organization uses and assess which of the eight gaps your email archive or CASB has in coverage.
2. Identify the types of information security risks and conduct risks you want to detect in your organization's chat tools.
3. Identify the review workflows required for your compliance and information risk teams. Ensure the workflows align across user groups, geographies, and risk types.
4. Identify the systems (e.g. email archive, eDiscovery, alerting and logging systems) that need to integrate with and ingest information captured from modern chat platforms.
5. Use modern technology such as Theta Lake to address the gaps in compliance and security coverage for your modern chat platforms to reduce risk, improve compliance processes, increase the utility of your pre-existing compliance stack, and enable more effective use of modern chat tools.

Key Capabilities and Benefits of Using Theta Lake for Modern Chat Compliance and Security

Theta Lake's security and risk management solution is purpose-built to help organizations effectively solve compliance and security challenges with today's modern chat platforms. In quick summary, Theta Lake provides:

1. Non-disruptive and flexible deployment modes including archive coverage for Cisco Webex Teams, Microsoft Teams, RingCentral Glip, Slack, Symphony, Zoom Chat and more. We also have built-in integrations for pre-existing archive infrastructures such as Microsoft O365 Archive, Proofpoint, Mimecast, Smarsh, Veritas, ZL Technologies, and more.
2. Full coverage for capture and analysis of team, channel, group, and private chats with full capture of emojis, reactions, gifs, images, memes, documents, whiteboards, audio files, and video files in chat. This includes the capture of files shared via SharePoint links in Microsoft Teams.
3. Detections beyond lexicons, using machine learning to find compliance, conduct, data loss, data privacy, and security risks.
4. Robust, instant search with advanced review workspace with native and contextual chat view to quickly search all chat records as well as effectively review chat conversations.
5. Support for actual remediation and removal for problematic content in chat conversations in the chat application itself, as well as the ability to redact PII, PHI, and PCI stored in Theta Lake.

Learn More

To receive a live demo of Theta Lake's archiving, supervision, and data loss protection for modern chat platform, click here.

ABOUT THETA LAKE. Theta Lake, Inc., dual winner of UC Today's 2020 Best Compliance Product and Best Security Product, was founded in 2017 by proven entrepreneurs and enterprise technologists with decades of leadership experience and recognition from Global 100 customers and top industry analysts. With a mission to provide modern collaboration security and compliance, Theta Lake's patented and multi-patent pending AI helps security and compliance teams more effectively and quickly scale their risk detection and the workflows for communication security, data loss protection, and supervision of modern video, voice, and unified collaboration systems. Visit us at ThetaLake.com; LinkedIn; or Twitter at @thetalake.