The State of Emerging Data 2023: An Overlapping Arrival and Departure

EXPERTS WITH IMPACT™

FTI CONSULTING™

Emerging data sources passed a tipping point this year. It was an expected turn for experts watching the evolution of systems, applications, technology adoption and communication habits over the past several years. Some understood that this year would bring an arrival of a new data reality. Indeed expectations have been validated — **71% of CIOs from the world's largest organizations now say that the explosion of data is beyond human ability to manage.** Simply, traditional data norms have been replaced by a complex, vast, dynamic and constantly changing data universe.

In 2020, the sudden shift in work environments ushered in an explosion in the use of cloud-based applications and collaboration tools. Then, over the following two years, there was a steady process of slowly waking a sleeping data giant — in which growth in data volume, variety and velocity began creating new and complex facets of corporate risk. Today, emerging data sources have fundamentally transformed how data is created, shared, stored and relied upon. At the current data growth trajectory, most data within an enterprise (and by extension, in scope for litigation and investigations) will be from cloud-based, emerging data sources by the end of this year.

Emerging Data Sources Defined:
An emerging data source is any cloud-based platform or collaboration application used for business purposes. The most commonly known are Microsoft 365, Google Workspace and chat tools such as Slack and WhatsApp. There are thousands of emerging data sources currently in use around the world, with more arriving all the time.
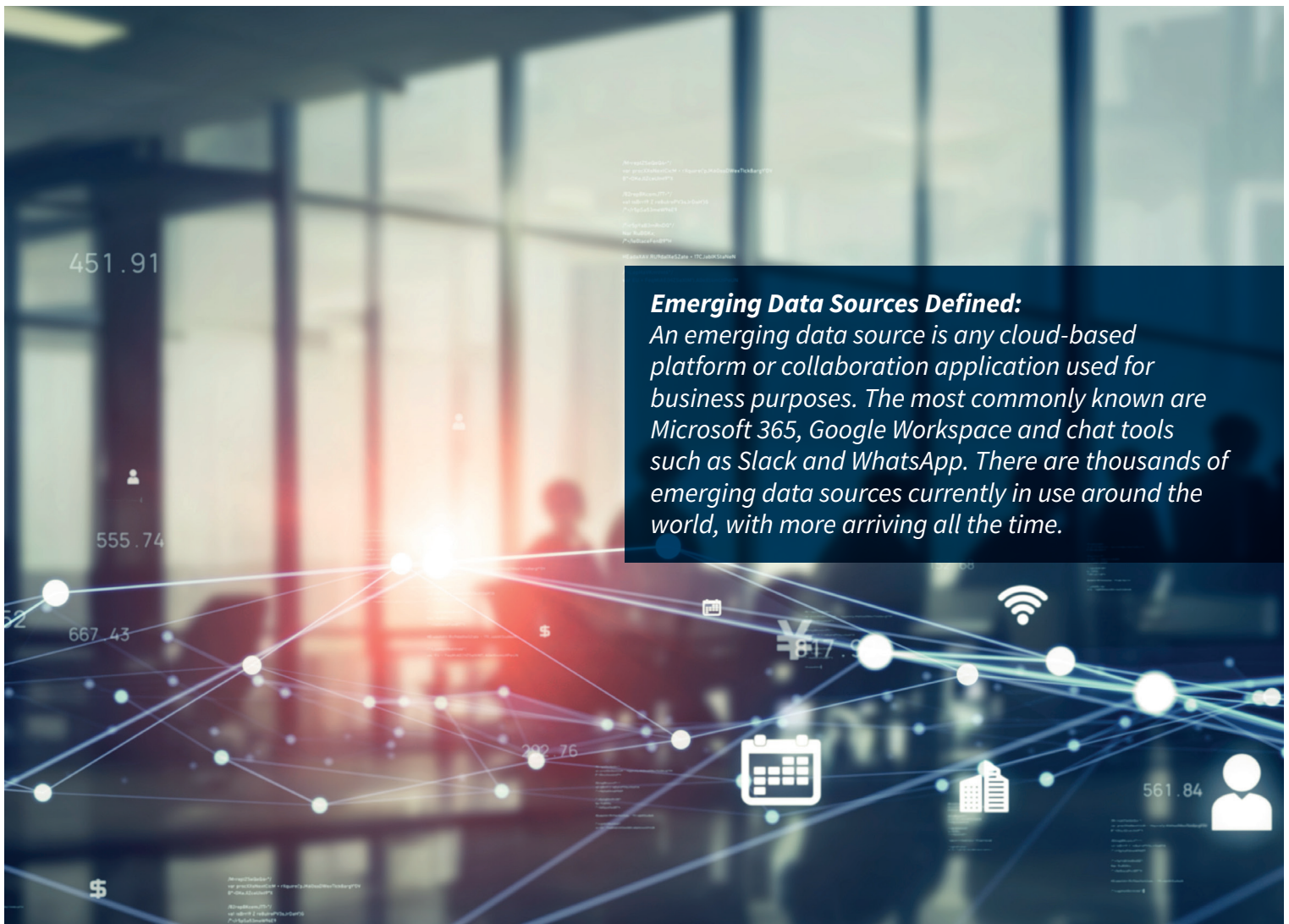
A Look at Data Trends Around the Globe

More than 40% of general counsel have reported that their legal departments are experiencing new challenges relating to emerging data sources, according to *The General Counsel Report 2023* from FTI Technology and Relativity. A study of global CIOs found that 77% of organizations say their IT environment changes every minute or less.

FTI Technology's Digital Insights & Risk Management report found that one-third of senior business leaders expect operational and/or compliance risks relating to emerging data sources to increase in the year ahead. Additionally, 21% of those respondents listed the growing number of data types and formats as one of their top three risks. A closer look at trends by region presents an even starker picture of just how pervasive data challenges have become.

NORTH AMERICA

— Analysts expect the use of persistent messaging at work to continue to grow, with one estimate that daily active users will reach 247.6 million by 2025.

— Also by 2025, 44% of data created will be driven by analytics, AI and deep learning, and by an increasing number of internet of things devices feeding data to the enterprise.

— In 2023, the number of connected devices will reach 15 billion, a number that is expected to double by 2025.

LATIN AMERICA

— In Brazil, a single enterprise application transaction crosses 47 different technologies from beginning to end.

— Messaging applications have become ubiquitous, with WhatsApp the most used messaging tool in the region, with penetration across more than 92% of internet users in selected countries (Brazil and Chile top the list).

U.K. AND EUROPE

— The U.K. is the third largest market (behind the U.S. and China) for adoption of collaboration software.

— Data has become so central to business and society that the European Union has introduced or implemented numerous regulations to govern its use, including the Digital Services Act, the Digital Markets Act and the Data Governance Act.

MIDDLE EAST AND AFRICA

— 83% of Middle East organizations are planning to invest significantly more in their digital culture in the coming year.

— 112.7 million mobile connections were active in South Africa in early 2023, a figure equivalent to 187.4% of the country's population.

AUSTRALIA

— 80% of organizations say their IT environment changes every minute or less and 78% say cloud data is now beyond human ability to manage.

ASIA

— Use of real-time mobile messaging tools are increasingly being used for a mix of personal and business purposes in Asia — WeChat has more than 1.2 billion users and DingTalk has more than 500 million, as well as more than 19 million institutional users.

— Indian companies spent nearly $2 billion on data analytics in 2022.

Emerging Data Dominates Workplaces, Courts, Regulatory Investigations, Data Breaches and More

FTI Technology's disputes and investigations work deals with large volumes of data every day. This data is handled for a number of reasons, including e-discovery for litigation, regulatory inquiries, internal investigations, data breach response, data governance projects and more. Emerging data sources have fundamentally shifted the nature of this work on all fronts.

For example, consider these anecdotal figures:

— 35% of data now managed for legal and regulatory purposes falls within the emerging data category, a nearly 20% increase over the previous year.

— In 2022, 20% of the data provided in standard Outlook PST format was data from Microsoft Teams, even when clients were adamant that no Teams data was present.

— Data volumes grew 75% in the last year.

— More data was processed in the first quarter of 2023 than in the entire year of 2017.

— The number of data formats now in play in e-discovery has increased 10x from five years ago.

— 100% of regulatory investigations now involve at least one emerging data source, compared to only roughly 10% in 2019.

— A significant volume of data breach response matters over the past three years have involved data from Microsoft 365. Google Workspace and Slack are also common sources in these types of matters.

With the arrival of these new conventions, there's been a simultaneous departure of governance, compliance and e-discovery activities that were previously considered standard practice. These emerging data platforms were not designed with e-discovery in mind. Nevertheless, courts and regulators have made clear that if emerging data systems contain relevant communications, documents or other records, they must be collected, reviewed and produced just as traditional sources must be, or the company will face risk of sanctions or other penalties.

To mitigate the resulting costs and risks, organizations must understand that legacy workflows and traditional expertise no longer apply.

Several recent matters have underscored the critical nature of these issues. In one case involving a digital

assets organization, our team identified and collected dozens of terabytes of data from platforms including Slack, Google Workspace, Dropbox and others. Among the many challenges associated with these emerging data sources, one notable issue was the use of emojis. Not just emojis used in the context of a message, but emojis as reactions to messages. The client organization used specific emojis, with pre-determined meanings, to react to Slack messages. These reactions would then drive specific workflows based on the emoji used. Having the ability to capture the reactions and represent them visually were key as was the ability to search and batch by reaction. With this capability, the team was able to more quickly and effectively track and analyze relevant activities across Slack and the related systems, which were key to informing the facts of the investigation.

Key Issues to Watch in 2023

Ongoing Disputes and Investigations Implications

In *The General Counsel Report 2023*, e-discovery risks relating to emerging data sources rose by more than 10 percentage points from the previous year report. One respondent said, "It is a huge discovery mess and has prompted us to look for more legal hold and Slack-related discovery tools."

These risks have played out in the courts. For example, in the U.K. last year, the England and Wales High Court issued a more than £500,000 penalty to a defendant for failing to fully produce relevant documents and fulfill e-discovery obligations in a patent litigation. In part, these issues resulted from a lack of knowledge and expertise in the complex legal and technical issues that arise in matters involving dynamic data.

Moreover, rulings in recent cases including Nichols v. Noom, Inc., Fast v. GoDaddy.com, Porter v. Equinox, Drips v. Teledrip, and others, have shown how unprecedented issues, such as whether hyperlinked documents should be considered as attachments and preservation obligations relating to messages sent via apps such as Facebook Messenger and Signal, are introducing a new wave of debate about the nature and parameters of e-discovery.

Lack of Standardization

In the former data reality, most companies — and subsequently, most e-discovery tools, best practices and case law — had standardized on a few main systems for email and file sharing. However, there is no such homogeny in emerging data platforms. Each one is unique, many are updated regularly (which can change how data is

collected after each update) and they were not built to have data reviewed in an e-discovery software platform. This, combined with the fact that case law to date has been limited, there are no standard workflows that solve for the nuances in emerging data sources.

Similarly, IT may not be familiar with the e-discovery nuances, and in rare instances where there is a preservation process, workflows break down when it comes time to export the data at scale or in a review-ready format.

Velocity of Change Persists as a Significant Issue

By nature, cloud applications are constantly changing, improving and adjusting to business and productivity needs. While a fast pace of change is hugely beneficial from an end-user perspective, it is equally problematic in the context of governance, compliance, investigations and legal discovery. When new or adjusted platform functionality is continuously introduced (which often occurs without notification), a never-ending game of catch-up ensues for teams working to minimize, preserve, collect, analyze and review data from those platforms.

The e-discovery and compliance functions that do exist in certain platforms (such as Google Workspace and Microsoft 365) also change frequently. This can impact export formats, changes in export options, the inclusion of linked content and attachments, relevant metadata and more.

Regulatory Attention Across Anti-Trust Enforcement and Compliance

Global regulators are aware of the shift in the corporate data landscape. Authorities in the U.S., Europe, the U.K., Brazil and other regions are increasingly including information from emerging data sources in their inquiries. With the way the lines have blurred between personal and business communication, and the channels through which these communications take place have multiplied, there is now significant regulatory risk if emerging data sources are not properly managed.

In a competition law context, one recent case in Colombia resulted in a company receiving a fine for obstruction after an employee refused to turn over a personal device containing company data during a dawn raid. In another raid, the Netherlands Authority for Consumers and Markets (ACM) issued a €1.84 million obstruction fine to a company whose employee deleted WhatsApp messages during a dawn raid inspection.

In the U.S., the Department of Justice and the Federal Trade Commission have issued numerous statements and guidance implicating emerging data sources — such memorandums have been issued in recent months, and dating back to 2020 when the nature of the traditional workplace was first disrupted. This includes the requirement that companies govern all communication

The Four Pillars of the Emerging Data Paradigm Shift

Shared Access

Nuanced shared access roles and complexity, for custodian identification, while also increasing data volumes

Chat Messaging

Chat, channel and short form messaging increase volume and remove content

Linked Content

New ways of sharing data via hyperlinks challenge the existing document attachment paradigm

Versions

Access to multiple versions of files provides a previously difficult-to-access historical view of content

channels as part of their compliance programs, and be prepared to produce information from those channels in the event of an investigation.

Moreover, guidance indicate that data and devices must also be investigated for irregularities, after end of use or after employees leave the company, before any data is deleted. These are no small feats, but have become unavoidable realities.

A Path to Insight

Emerging data sources have arrived as the primary category of corporate data. With that, there must be a departure of previous convention, workflow and best practice across numerous functions. While the challenges are great, and organizations may face significant work in adapting to the current requirements and solving for new challenges, there is an upshot. When technical, defensible and innovative solutions are implemented, issues can be predicted and insights can be derived. Ultimately, emerging data sources have the potential to give organizations faster and more actionable information about their exposures and overall risk profile.

TIM ANDERSON
Senior Managing Director
+1 (415) 293-4483
tim.anderson@fticonsulting.com

COLLIN MILLER
Managing Director
+ 1 (915) 801-8997
collin.miller@fticonsulting.com

F T I
CONSULTING