

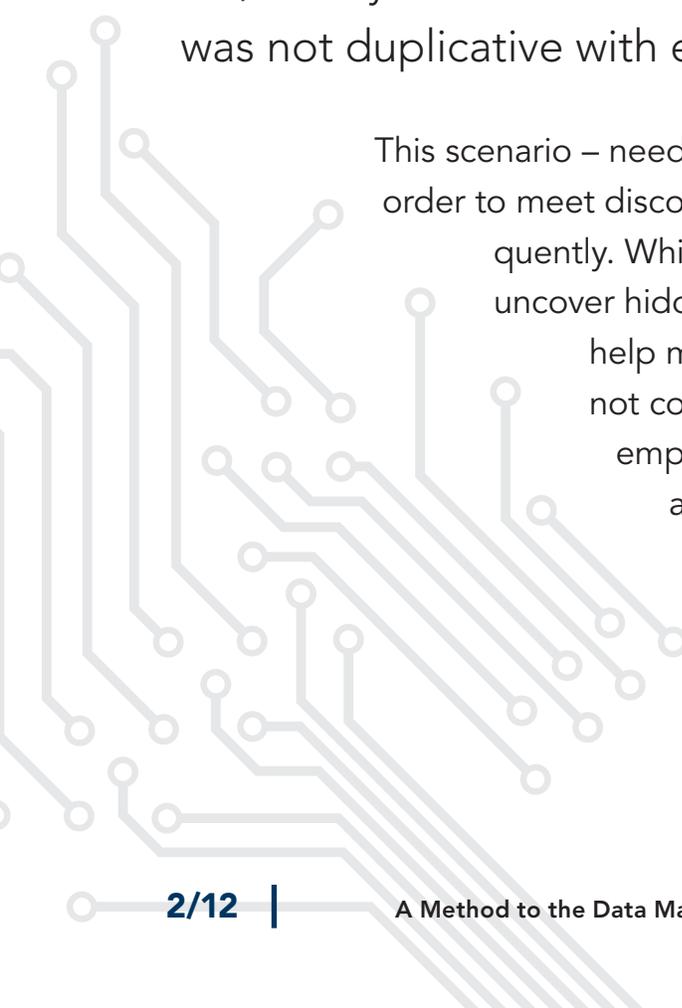
# A Method to the Data Madness: *E-Discovery in the App Age*

By Tim Anderson and JR Jenkins



**F T I**<sup>™</sup>  
CONSULTING

TECHNOLOGY



**W**

hile we were assisting a law firm on the discovery process during a recent legal matter for a large company, something interesting happened. During one of the final custodian interviews, after we had exhaustively detailed the myriad of data sources in play for this matter, we ended the interview with a question that was almost an afterthought: Are there any other sources of data that we haven't already covered? The custodian said "no," then after a pause added, "wait – did anyone mention the Google Apps pilot we've been running?" No one had, not the CIO nor the legal team nor any of the other custodians. It was far enough outside the normal information ecosystem that no one had thought about it. As it turned out, 30 key custodians were on the pilot, and some of the data was not duplicative with email or any other data source.

This scenario – needing to collect and review data from new apps in order to meet discovery requirements – is likely to be seen more frequently. While some companies have "data scientists" to help uncover hidden treasures in their archives, or at the very least, help map the organization's data landscape, this is still not common for the vast majority of organizations. As employees migrate to an ever-increasing number of apps, how can legal teams ensure that this data is picked up as part of the usual e-discovery process?



# How did we get here?

---

Most legal teams understand that in today's fast-paced, global business climate, the vast majority of litigation events and regulatory investigations involve electronic discovery. The collection, processing and review of large amounts of "traditional" electronically stored information (ESI) – think Word documents and emails – have dramatically driven up the costs of litigation. With more mobile devices and the groundswell of cloud-based business apps, the challenges of "too much data" will only continue to grow. A by-product of the explosion of cloud computing – first as backup, then as convenient storage and file exchange – these cloud-based business applications are impacting standard business processes from project management to basic communication and bring with them new complexities for today's legal teams.

There are a number of reasons for the growth in these types of applications. People and businesses are increasingly comfortable using different apps, rather than simply email, to communicate. This "app culture" may have started on smartphones, where texting is ubiquitous and

it is common to download apps to do everything from reading magazines to buying airline tickets to communicating across borders.

Then the app culture migrated to businesses. Chief information officers are beginning to understand the limitations of enterprise email systems and shared drives and are looking for tools that allow easier communication and sharing of information while supporting process and workflow. Many of these systems are cloud-based, which makes them easy to implement. In fact, many are so easy work with that ad hoc groups of workers are downloading them themselves and using tools that are not supported or well understood by IT.

Additionally, a growing percentage of data is created away from the workplace. Today's mobile workforce does a great deal of its work on mobile devices, often enabled by the cloud. The result is more data of more types being stored in different places.

# The Discovery Issue

---

Because these tools are becoming a vital source of information, employees are using them even before policies are developed for them. Thus they are not always incumbent in the traditional discovery workflow. The CIO may not know who is using what and how and what data exists where, and new tools are showing up all the time.

As another example of this scenario, we assisted on a recent matter in which we collected hundreds of thousands of messages and documents from Slack, Dropbox, Google Mail, Google Drive and Evernote. In fact, data from these systems made up the overwhelming majority of the data acquired for review in the matter. This was in addition to the documents collected from the standard laptops and mobile devices.

There are many new tools which circumvent the corporate information ecosystem and allow data to be stored and moved

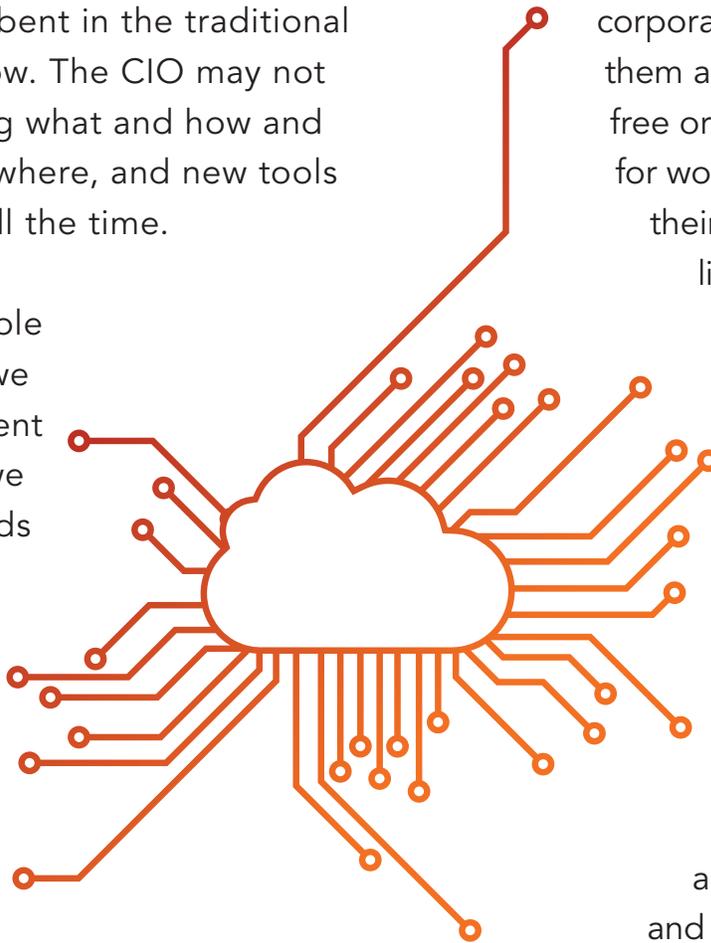
without it ever touching corporate servers. All of them are easy to install and free or inexpensive enough for workers to sign up on their own. Big players like Google Apps and Microsoft

Office 365 allow documents to easily be shared in the cloud.

Project management tools like Atlassian and Trello support communication and workflow

among teams. Dropbox and Box store files of all types in the cloud. And col-

laboration workspaces like Yammer and Slack are growing by leaps and bounds: As of early 2015, Slack alone had grown from 15,000 to half a million users in one year, and those users now create more than 300 million messages per month.



# Assembling a Complete Picture

---

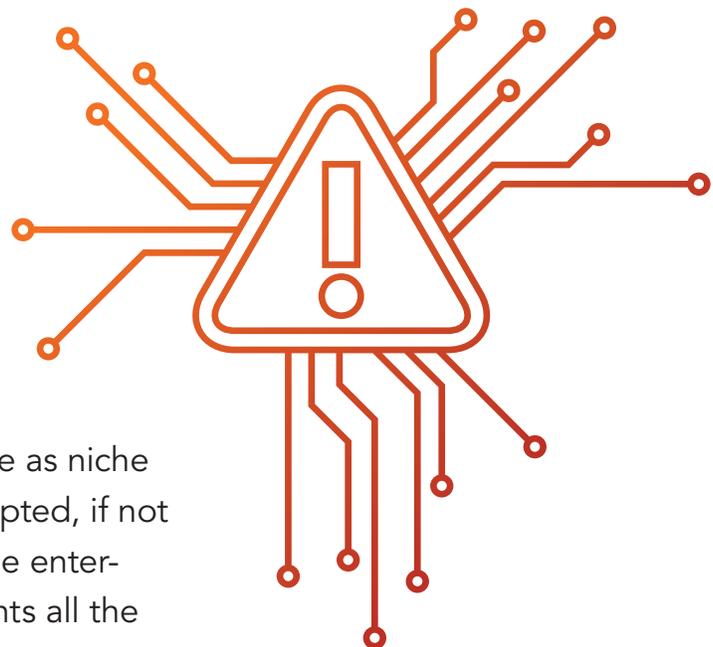
Given the growth of these new, less traditional sources of information, it is critical for organizations to develop a predictable methodology for ensuring that this data is included in internal investigations and e-discovery. But to do this, teams must overcome challenges on both business and technical fronts.

## CHALLENGE ONE:

### Awareness

It can be difficult to identify these systems since new ones are popping up all the time. Once they are identified, it can be a challenge to understand who is using them and for how long, especially considering that there may be many installations of the same tool throughout the organization.

The current discovery culture treats these as niche systems, but they are being broadly adopted, if not by the enterprise then certainly inside the enterprise. We see these tools on engagements all the time now, even in Fortune 500® companies.

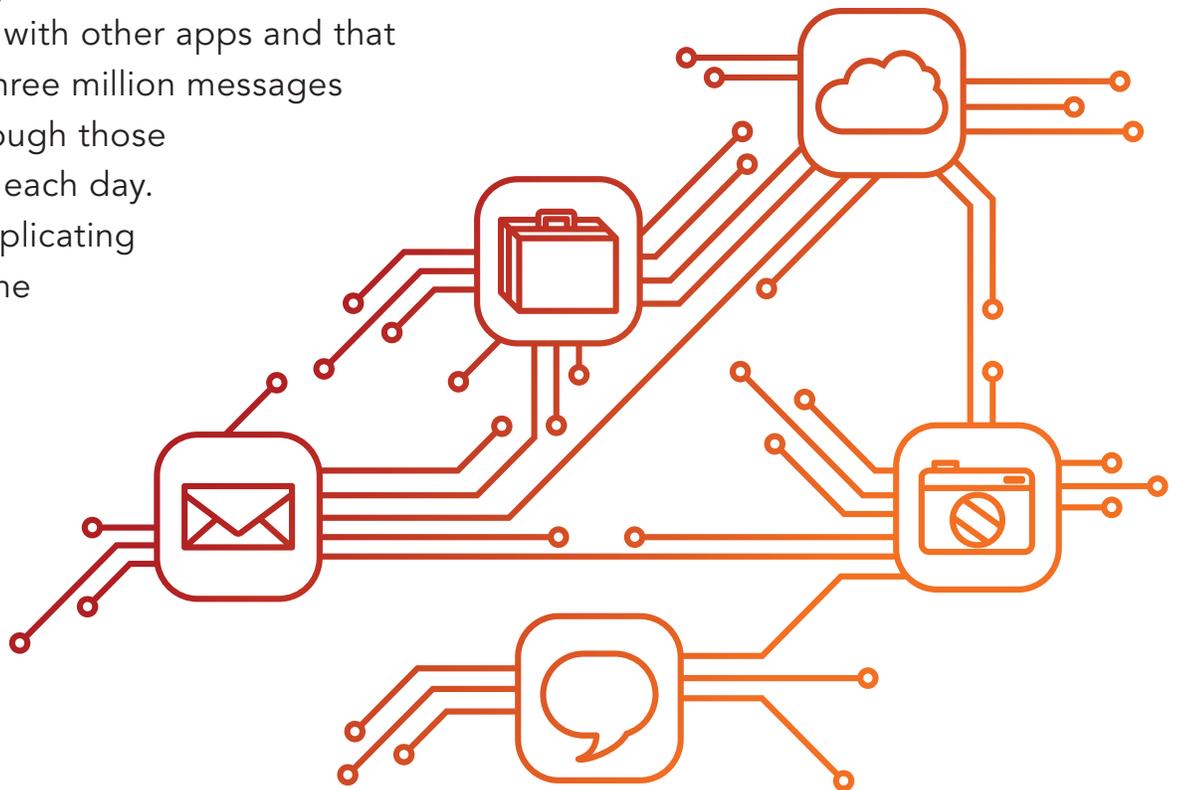


## CHALLENGE TWO:

# Connecting to the data

A discovery project typically requires legal and IT teams to be able to access data quickly, which requires an understanding of how these platforms function and the connection methods available. Most modern platforms provide RESTful API access to allow developers to make applications that interface with data; this is one of the factors that is driving the adoption of these systems. One of Slack's biggest strengths, for example, is its simple integration with third-party apps. The company says it has 800,000 individual integrations with other apps and that more than three million messages are sent through those integrations each day. Further complicating matters, some of the more

enterprise-friendly systems like Google Apps and Office 365 provide administrator capabilities that allow multi-user connectivity, while other systems are single user access only. These systems sometimes have some compliance systems, but the system may not always function as expected or provide data in formats easily readable by traditional electronic discovery tools.

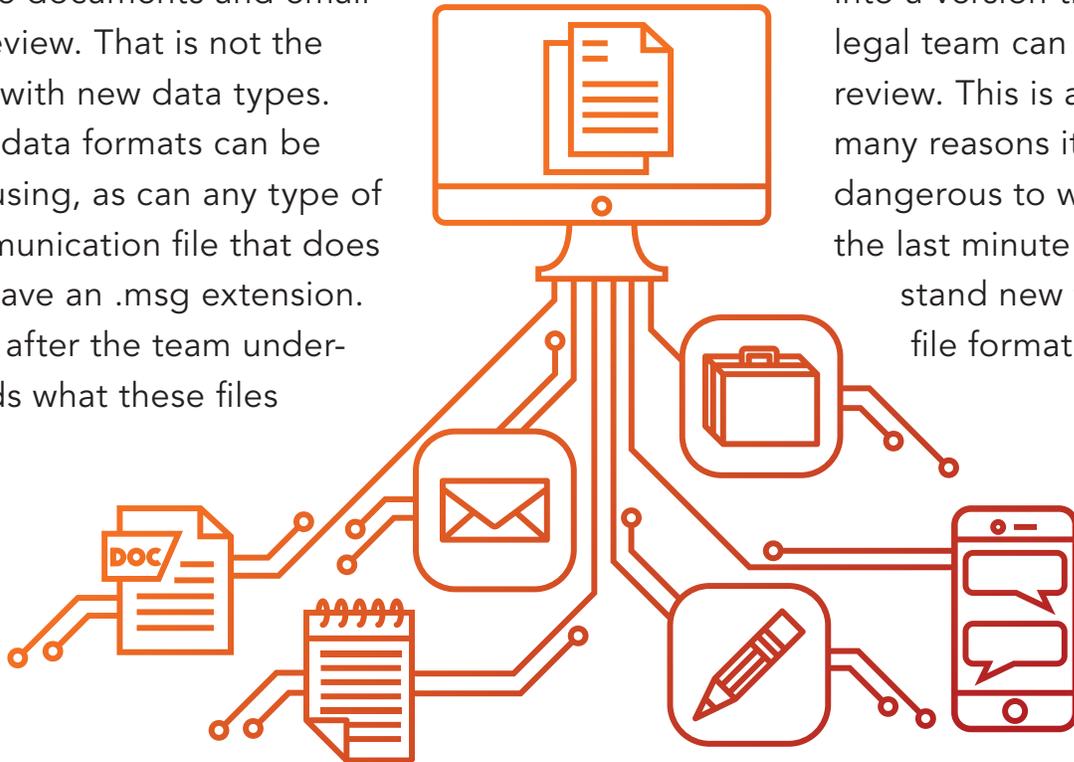


## CHALLENGE THREE:

# Parsing the data

Discovery professionals have years of experience preparing Microsoft Office documents and email for review. That is not the case with new data types. New data formats can be confusing, as can any type of communication file that does not have an .msg extension. Even after the team understands what these files

are, there can still be a steep learning curve for how to transform them into a version that the legal team can quickly review. This is among the many reasons it can be dangerous to wait until the last minute to understand new tools and file formats.



## CHALLENGE FOUR:

# Weaving new data into your time-tested workflow

Data from collaboration tools often does not look or act like data from legacy email or word processing systems, with distinct files for distinct communications.

As such, it likely does not fit into the workflow that has been built for well-understood sources of data. It may also have a meaningful impact on what must

be done to properly display the data for review and production.

Finally, do not ignore other contexts besides litigation, such as investigations and data breaches. These communications tools often create many copies of the same documents, with some copies archived, some emailed and some sitting in a variety of inboxes and outboxes. For example, an employee may have attached a salary

spreadsheet to a Slack message. Now there is another copy of that spreadsheet that can create risk from an information governance or data privacy perspective. So it is important to do risk assessments to understand where data is located and how it moves throughout the organization. This is especially true in today's era of compliance, with so many enforcement issues from regulators around FCPA and other types of issues and investigations.



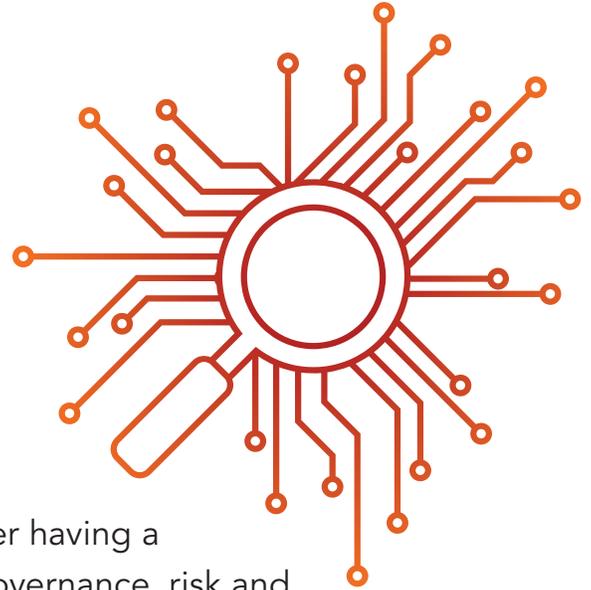
# Key Takeaways

---

While next-generation collaboration tools do create some new challenges, none of them are insurmountable. Here are some tips that will help:

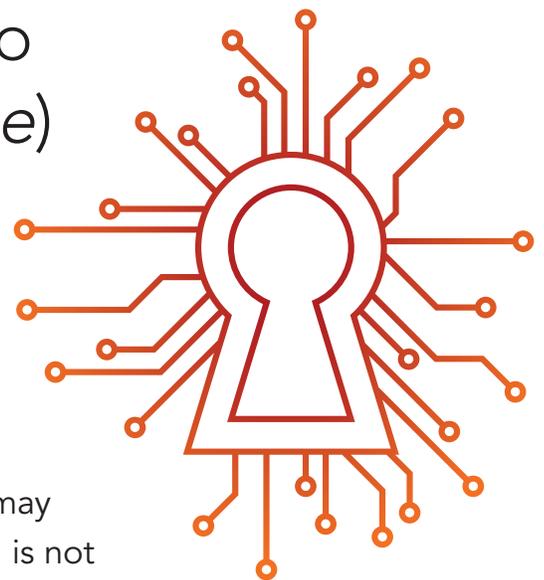
## 1. Fail-safe for Today and Tomorrow

Legal teams must have the ability to work with collaboration tools at the “big data” scale. Make sure yours is ready and able to find data and use it to uncover malfeasance, target custodians for collection, assess exposure and derive business intelligence on short notice. Also consider having a repository available to manage this data from a governance, risk and compliance perspective, as well as for business use.



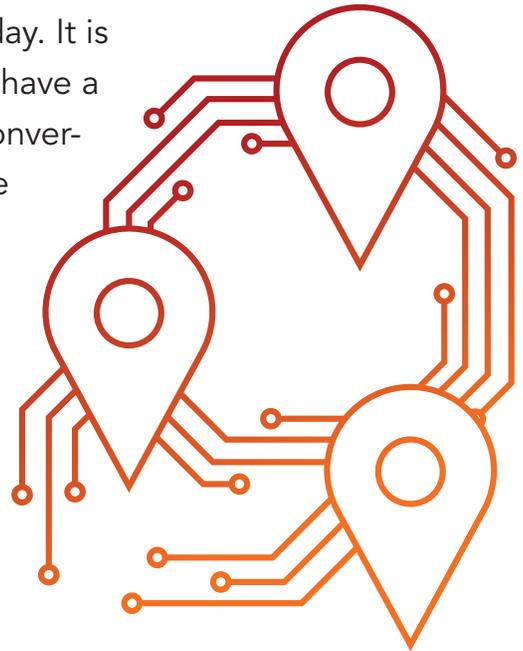
## 2. Know How to Connect to Critical (*and Discoverable*) Information

Understand the connectivity models that are incumbent in these systems. Some (like Google Apps and Office 365) have administrator access or allow for access via OAuth and similar protocols, but others (like Slack) require user authentication. So you may to gain access from the custodian, but if the custodian is not available, access can be difficult.



### 3. Proactive Transformation and Enrichment

We communicate on many different platforms every day. It is not uncommon to start a conversation on email, then have a phone call and later follow up with a text, all in one conversation. So it is important to have the ability to take the underlying data and view it in the context in which it was created – and not get hung up on the platform. The data must be standardized across platforms. For unstructured content, legal teams should be able to understand additional information by using algorithms and technologies that automatically identify people, places and things. New technology that allows the data to “speak for itself” can be extremely advantageous.



### 4. Draw a Coherent Picture

Concept clustering and visualization can turn a million data points into a single coherent picture that can allow attorneys to determine facts or fact patterns and parse through the details of what they are working with. This is necessary in order to understand what is important, because the numbers are so large pictures are required for people to understand the information.

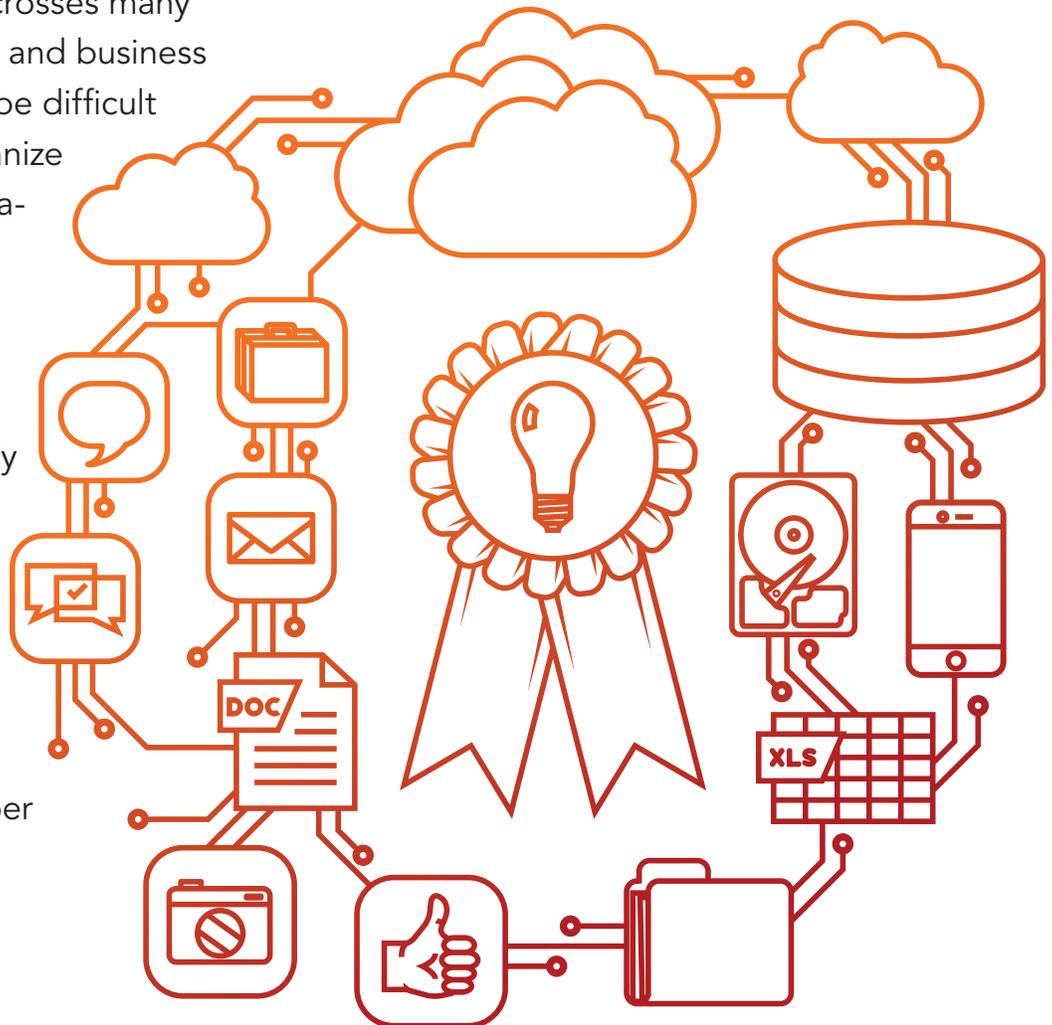


## 5. Do Research to Understand the Data Universe

It is absolutely critical to understand who is using next-generation collaboration tools before you need access to the data. Create an environment where this can be done immediately and regularly, and well in advance of a discovery project.

The information governance and discovery continuum crosses many different departments and business processes, and it can be difficult and expensive to organize all the data. Many organizations are seeking pragmatic ways to do this analysis and understand risks and mitigate them, not only for current projects but into the future. Many organizations have started to rely on data scientists to help. It's a good start. However, as the number

of sources increases dramatically and data is subsequently less organized throughout the company, information governance will no longer be the domain of the few, and new policies, processes and tools must come on line that will allow subject matter experts throughout the organization to themselves act as data scientists.



# About FTI Technology

FTI Technology helps clients manage the risk and complexity of e-discovery, compliance and information governance. We collaborate with clients to develop and implement defensible data governance, investigation and discovery solutions that utilise our award-winning visual analytics software. Clients rely on our software, services and expertise to address matters ranging from internal investigations to large-scale litigation with global e-discovery requirements.

## FOR MORE INFORMATION:

[ftitechsales@fticonsulting.com](mailto:ftitechsales@fticonsulting.com)

[www.ftitechnology.com](http://www.ftitechnology.com)

<b>North America</b>	+1 (866) 454 3905
<b>Australia</b>	+61 (2) 9235 9300
<b>Europe</b>	+44 (0) 20 3727 1000
<b>Hong Kong</b>	+852 3768 4500