b. **Non-custodian Data:** Data is non-custodian, and a traditional custodian interview is not feasible. Appendix C illustrates the adaptations to the New Framework.

c. **International Custodians:** If international custodians are involved, local data privacy law must be considered. Appendix D provides an example scenario that takes privacy laws into account.

Section 02: Defining Data Source Burden and Effort

I.   Introduction

The New Framework's "standard approach to frame proportionality assessments" requires the designation of data sources by the degree of burden. Section 2 provides guidance on how to rank eight of the most common data sources in four broad categories, based on the degree of burden incurred in accessing information. Five primary variables are identified, which directly affect the degree of burden for each data source. The specific degree of burden associated with each data source is not determined, which is dependent on circumstances.

Although the ranking of data sources by burden must be done on a case-by-case basis using the five variables, the New Framework provides a "model set" of rankings for a "typical" case. The model rankings are based on a literature review and the collective judgment and experiences of the New Framework's experts, applying the five variables as they most commonly appear in a typical case. They are intended to provide guidance and reference points, especially to those unfamiliar with ESI, when ranking the data sources. A graphic illustrating the "model set of rankings" of the eight data sources is set out at the end of this section.

Individual circumstances will require adjustments to the model rankings. The clearest example are emails, which typically are located on a centralized server, providing relatively easy access and modest burden, compared with the less common use of decentralized email, which is located on individual computers, requiring multiple extractions and significantly increasing the burden.

The following discussion identifies the eight most common data sources and explains the variables that affect the burden assessments.

II.   Data Sources and Types of Data

Listed below are the eight most common sources of ESI, which are considered under the New Framework.

- Collaboration / Messaging Systems (Slack, Teams)
- Computers / laptops
- Email Systems
- File shares (departmental and personal)
- Mobile devices
- Paper documents / physical evidence
- Social Media
- Structured data (HR, finance, marketing databases)

A. Specialized Data Sources

Listed below are specialized data sources, which are not considered under the New Framework, but can arise in an individual case.

- Backup media
- Computer code
- Corporate telephone data
- Ephemeral data
- Geo-location, GPS, IP addresses
- Specialized/proprietary databases, systems, or programs.
- Website data

III. Variables Affecting Burden Assessments

Five major variables affect the burden assessment, including:

      A. The location of and accessibility to the data

      B. The availability or state of readiness for collection

      C. The methods used to preserve and collect data

      D. Any specialized resources needed to effectuate collection

      E. Any legal or regulatory constraints (e.g., data privacy laws) that may impact collection

These variables are central in determining the burden associated with collecting ESI.

A. Access to and Location of the Data

The accessibility and physical location of data sources can significantly impact the effort and burden of collection. The more centralized the data is, the easier it is to access the information, which lowers the burden.

For instance, if the data is centralized in a single server room or a single cloud instance, it is more readily accessible than if it is spread across multiple data centers or cloud-storage locations that are not interconnected. The geography of US-based and international locations may also place a greater burden on coordination and collection efforts.

It is important to determine whether sources like desktops, laptops, and mobile devices can be accessed remotely from a central location, or whether the collection can only be accomplished by physically traveling to the actual data source.

Many organizations now allow employees to use their own personal devices to connect to their organizational networks and access work-related systems and potentially sensitive or confidential data. BYOD situations present a new challenge to logistical collection efforts and also raise significant privacy concerns.

B. Availability of Data

The state of the data and its ready availability impact the degree of burden. Accessing data from a departed employee increases the burden.

Thus, if the data belongs to a current employee, it is more readily available than data belonging to a former employee, which may have been archived or dispersed. A current employee's data is typically in an active state and readily available for collection by the custodian, IT, or vendor. If the employee is no longer employed, the data source is no longer in use, or has been deactivated, availability to the data becomes more challenging.

C.   Preservation and Collection Methods

The preservation method utilized may impact the burden as well. If special processes are required to preserve information before it can be accessed, the burden is higher.

In some instances, the only way to defensibly preserve data without altering it is through collection. In other circumstances, like those in which an advanced document management or filing system is in place, it may be possible to preserve data in place with little or no effort.

The method of collection can also impact the burden. A decentralized organization might require collection by the custodian, which would most likely entail custodian training to ensure against spoliation. On the other hand, some organizations may have a native collection tool that allows for in-place or near-line collection.

Finally, specialized data sources may require forensic or customized queries and system knowledge to collect data efficiently and safely.

D.   Specialized Resources to Collect

There are instances in which a data source that would typically be considered easier to collect information may impose added burdens because of age, corruption, or malintent. In such circumstances, it may be necessary to engage specialized tools or resources that significantly increase the degree of burden of collection.

Additionally, there may be corporate resources such as structured databases that require expert knowledge and understanding of the system in order to correctly and safely extract the pertinent data, leading to an increased burden of collection.

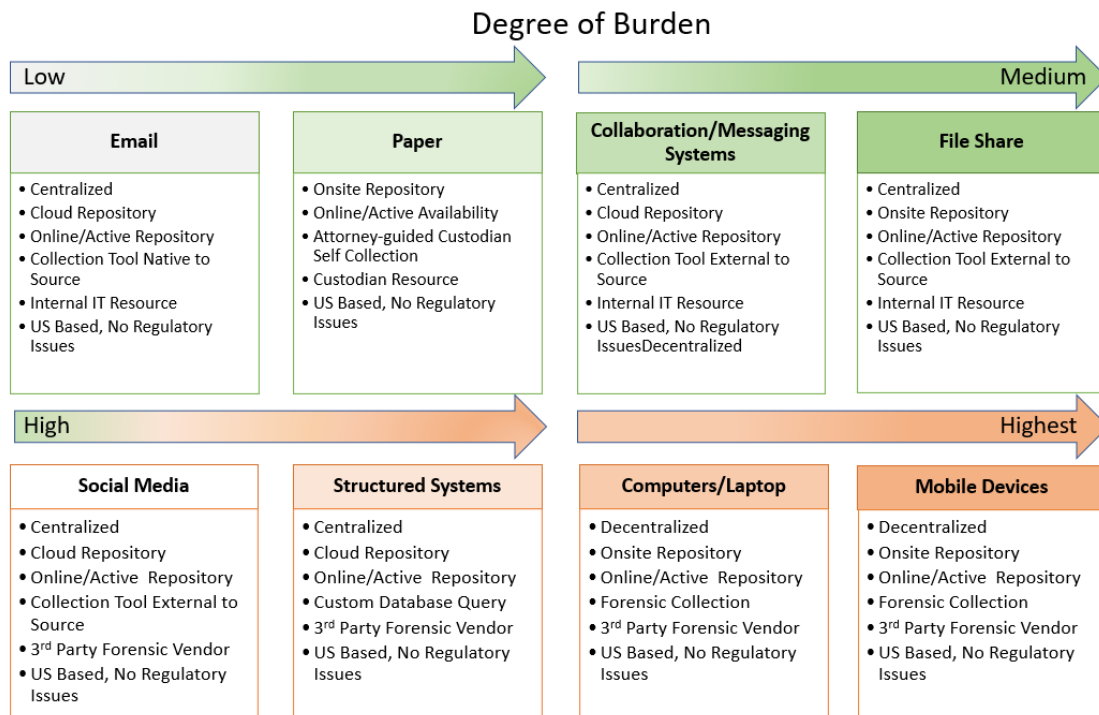## E.   Legal and Regulatory Restrictions and Requirements

Legal requirements can also add to the burden of collection. Privacy legislation and regulations may regulate or require a party to take certain steps before the data is collected. The E.U. General Data Protection Regulation ("GDPR") applies to any processing of personal data when a European company is the controller. The GDPR has specific legal requirements for the processing of personal data (preservation, collection, processing, and production are separate data-processing activities). For instance, each data-processing activity needs a legal basis,[11] and you have obligations towards the individual.[12]

In addition, the transfer of personal data for use in litigation-discovery in the U.S. is only possible if certain conditions are met.[13] Moreover, there are labor-law regulations, blocking statutes (foreign and domestic) and state-secrets law in other nations that should be considered. Failure to observe or properly comply with these legal and regulatory restrictions and requirements may lead to fines or even criminal consequences. Therefore, compliance with these legal and regulatory restrictions and requirements results in, among others, additional

consultations and assessment steps involving local counsel or authorities, which can add to the burden of collection.

IV.   Assigning Burden

The New Framework's model set of rankings of the eight most common data sources is based on variables that typically apply to provide baseline reference points, which are graphically displayed below. Accordingly, the low degree of burden associated with typical email is ranked at one end of the spectrum compared with messaging, which is ranked at the other end of the spectrum under the New Framework. Custodians with high-priority information at lowest discovery burden, are quickly identified.

Degree of Burden

Low ⟶⟶ Medium

Email	Paper	Collaboration/Messaging Systems	File Share
• Centralized • Cloud Repository • Online/Active Repository • Collection Tool Native to Source • Internal IT Resource • US Based, No Regulatory Issues	• Onsite Repository • Online/Active Availability • Attorney-guided Custodian Self Collection • Custodian Resource • US Based, No Regulatory Issues	• Centralized • Cloud Repository • Online/Active Repository • Collection Tool External to Source • Internal IT Resource • US Based, No Regulatory IssuesDecentralized	• Centralized • Onsite Repository • Online/Active Repository • Collection Tool External to Source • Internal IT Resource • US Based, No Regulatory Issues

High ⟶⟶ Highest

Social Media	Structured Systems	Computers/Laptop	Mobile Devices
• Centralized • Cloud Repository • Online/Active Repository • Collection Tool External to Source • 3rd Party Forensic Vendor • US Based, No Regulatory Issues	• Centralized • Cloud Repository • Online/Active Repository • Custom Database Query • 3rd Party Forensic Vendor • US Based, No Regulatory Issues	• Decentralized • Onsite Repository • Online/Active Repository • Forensic Collection • 3rd Party Forensic Vendor • US Based, No Regulatory Issues	• Decentralized • Onsite Repository • Online/Active Repository • Forensic Collection • 3rd Party Forensic Vendor • US Based, No Regulatory Issues

02-09-2021

Although this model set of rankings will apply in many, perhaps most, cases, the variables in an individual situation may be atypical and require adjustments, sometimes significant adjustments, to accurately reflect the actual burdens. Recognizing the possibilities that the five variables may be different for a particular data source in an individual case and require adjustments, a "Burden Assessment Calculator" has been developed to assist in leveraging the New Framework to account for the atypical variable and adjusting the degree of burden for the affected data source. The calculator can be used to select specific factors that are then evaluated to assign an approximate burden level of Highest, High, Medium, or Low to each data source.

The Burden Assessment Calculator is contained in Appendix E.