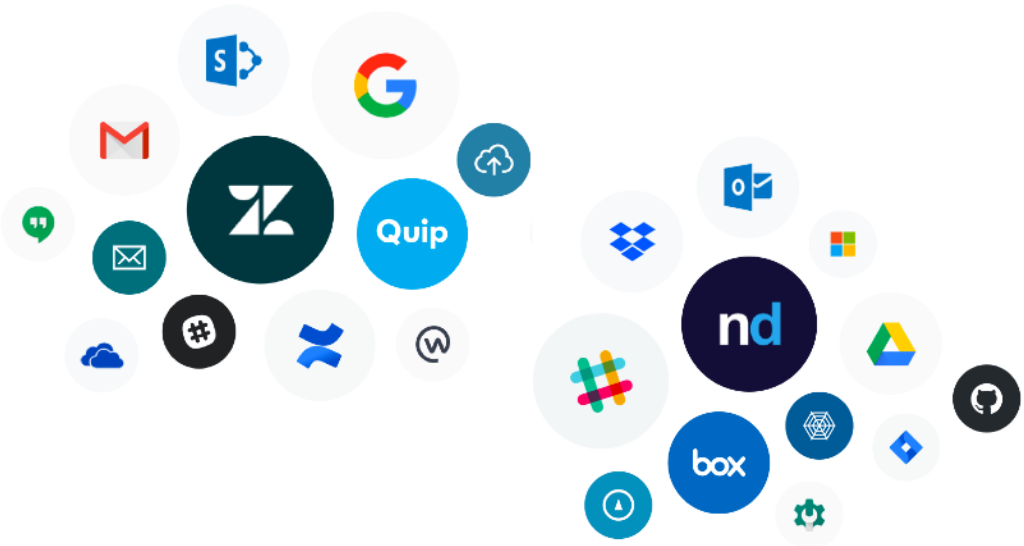
White Paper

Defensible and Proportional Precision Collection



www.onna.com

Defensible and Proportional Precision Collection

The Google It World

We live in a world where people have a “just Google it” expectation to quickly find answers to questions. The presumption that information can be found quickly has been a challenge for many when it comes to identifying electronically stored information for preservation. The cross section of varied data sources, from email messages to social media posts to cloud file sharing services, has often required using different applications for collecting different sources of data.

Onna is a knowledge integrator that connects to different data sources used by organizations to enable enterprise wide search. When collecting the files from these sources, Onna automatically processes and indexes the data. It uses deep learning during the processing phase to help identify types of documents, features embedded on images, clusters of information, relationships between files, sentimental analysis, language, and finds all possible searchable text extracting all entities and metadata inside all kind of files.

This white paper explores how Onna can bring Google expectations for identifying relevant data from multiple sources for the preservation of electronically stored information. Knowledge managers responsible for records information management, whether in-house or a managed service provider, can defensibly issue litigation holds on diverse sources of electronically stored information, and collect the data with precision.

Preservation and Proportionality

Parties have a duty to preserve data from “employees likely to have relevant information —the key players in the case, and applies to unique, relevant evidence that might be useful to the adversary"[1] when litigation is “reasonably foreseeable.”[2]

The scope of discovery stated in Federal Rule of Civil Procedure Rule 26(b)(1) is for "any nonprivileged matter that is relevant to any party's claim or defense and

---

[1] *Ala. Aircraft Indus., at *27,* citing *In re Ethicon, Inc. Pelvic Repair Sys. Prod. Liability Litigation, 299 F.R.D. 502, 517-518 (S.D. W.Va. 2014).*

[2] *Ala. Aircraft Indus. v. Boeing Co.,* No. 2:11-cv-03577-RDP, 2017 U.S. Dist. LEXIS 33527, at *27-28 (N.D. Ala. Mar. 9, 2017).

proportional to the needs of the case."[3] The proportionality requirements in Rule 26(b)(1) were moved forward in the 2015 Amendments to Federal Rules of Civil Procedure to address the “explosion” of data that “has been exacerbated by the advent of e-discovery.”[4]

This explosion of data can be seen from the fact there are 269 billion emails sent and 100 terabytes of data uploaded to Facebook every day.[5] Companies spend a significant amount of time just trying to manage their own data. For example, reports from IDC and the Butler Group state that knowledge workers can spend up to 8.8 hours a week simply trying to find information.[6] Assuming a knowledge manager’s annual income is $80,000 a year, that means $14,209 of their income goes towards “finding” information.[7]

Proportionality analysis requires considering, “the importance of the issues at stake in the action, the amount in controversy, the parties' relative access to relevant information, the parties' resources, the importance of the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit."[8]

Onna’s artificial intelligence technologies or “AI” can help attorneys working with knowledge managers to determine the proper scope of the lawsuit that is proportional to the needs of the case. Onna is a central point for over 20 different types of enterprise information platforms, such as Zendesk, Slack, NetDocuments, Google Drive, Sharepoint, or Dropbox. Its machine learning and natural language processing technologies can classify data across diverse applications, empowering a records information manager to execute a litigation hold on specific custodians across the universe of data.

---

[3] Fed. R. Civ. P. 26(b)(1).

[4] *Benavidez v. Sandia Nat'l Labs.,* No. CIV 15-0922 JB/LF, 2017 U.S. Dist. LEXIS 47724, at *65 (D.N.M. Mar. 30, 2017), citing Fed. R. Civ. P. 26(b) advisory committee's note to 2015 amendment.

[5] Waterford Technologies, Big Data Statistics & Facts for 2017, citing THE RADICATI GROUP, INC., Email Statistics Report, 2017-2021 and The Big List of Big Data Infographics, July 24, 2012.

[6] The Costs of Not Implementing an Enterprise Search Solution, WhereDat, June 13, 2016, citing The High Cost of Not Finding Information, by Analysts: Susan Feldman and Chris Sherman, and Enterprise Search Tools Move From Luxury Item to Business Essential as Data Builds Up, by Gary Eastwood.

[7] *Id.*

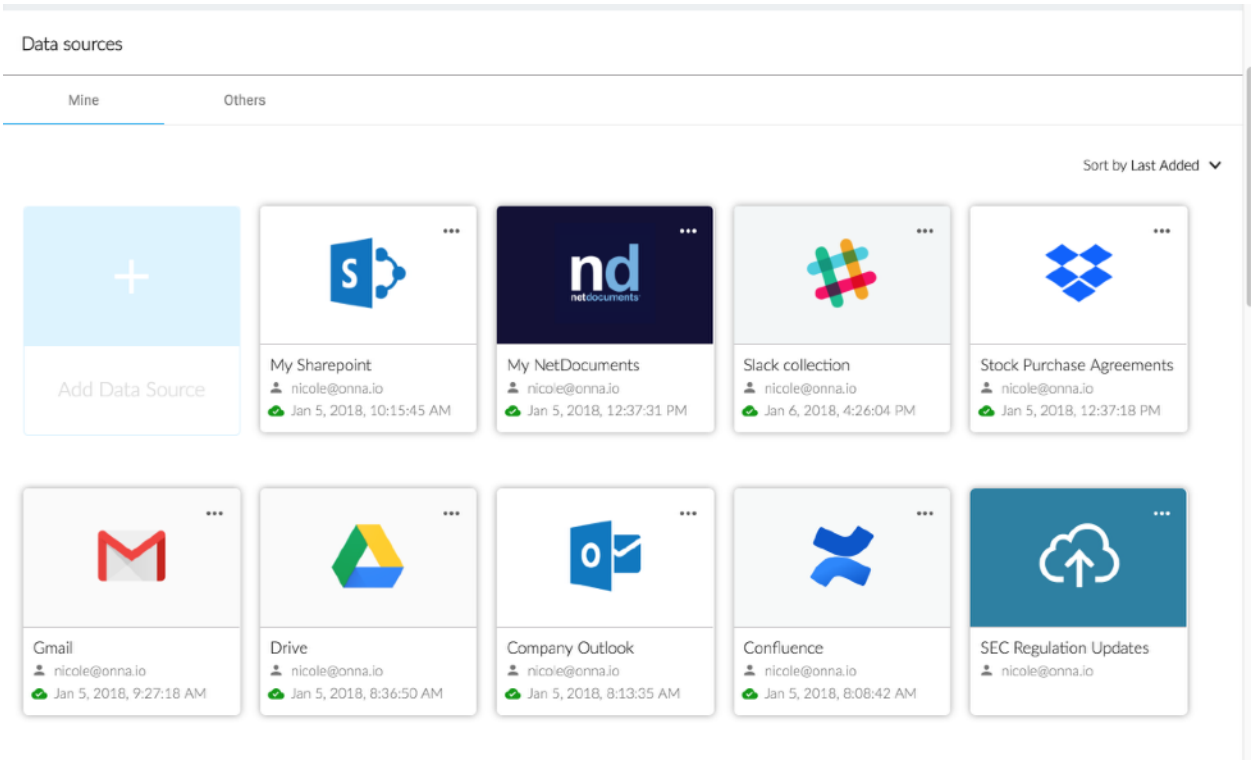
[8] *Prusin v. Canton's Pearls, LLC,* No. JKB-16-605, 2017 U.S. Dist. LEXIS 53453, at *3 (D. Md. Apr. 7, 2017), citing Fed. R. Civ. P. 26(b)(1).

What does this mean practically for issuing a litigation hold?

1. Onna can be set up with pre-trained models to identify relevant information. For example, the first step in a breach of contract case is to train the system to identify a small set of contracts based on the claims or defenses of the case;
2. Define the labels that you would like for the litigation hold (contracts with client ABC) and label the set of contracts.
3. For enterprise users, a training of the documents can be scheduled to enable the model to go through all previously included files within the platform and all future documents.
4. Enterprise users can then confirm predicted labels, reinforcing the model and improving its understanding of relevant documents.

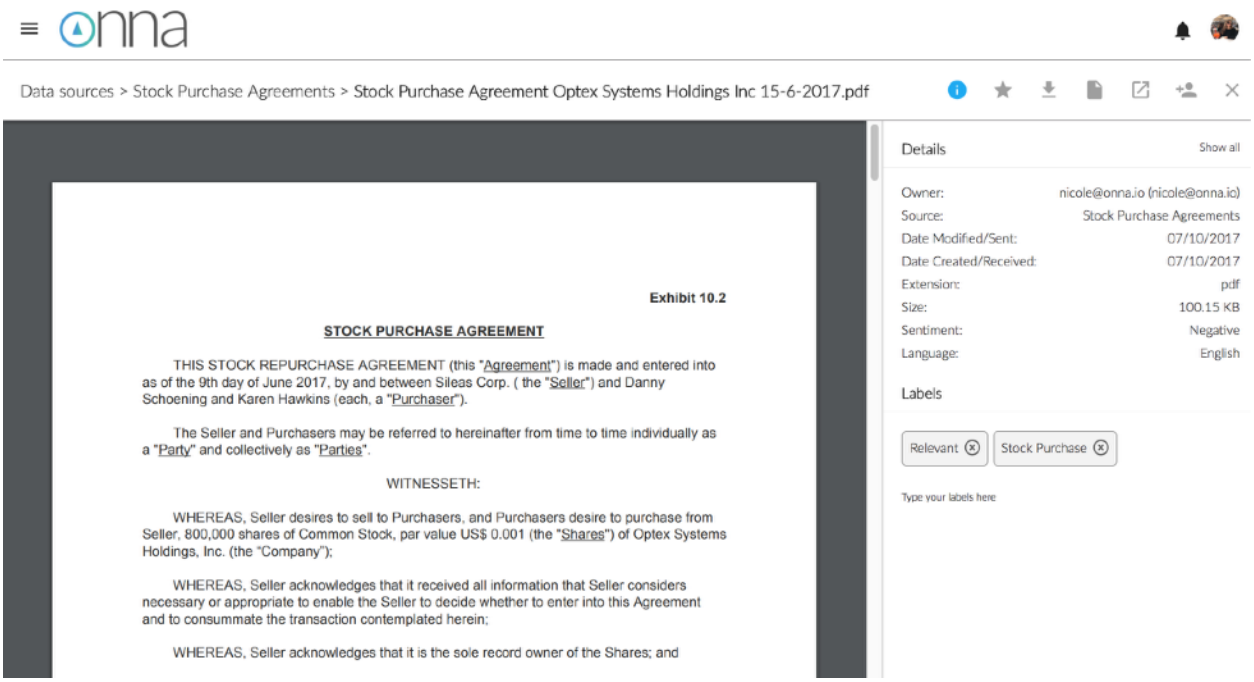
Dynamic models in enterprise environments are based on deep analysis of new documents to detect outliers on the global model plus reinforcement to provide a an accurate tuning. Onna uses a mixture of both models to provide an initial prediction (pre-set training models to identify types of contracts) and then provides records managers the ability to create specific training for their models.

The practical application of this technology is that a records manager can set the specifications of the litigation hold, which then is searched across a company's data sources. Its AI algorithms learns to identify what is relevant, so in a dispute involving breach of contract, Onna's search technology will find all versions of the contract and negotiations in an enterprise environment.



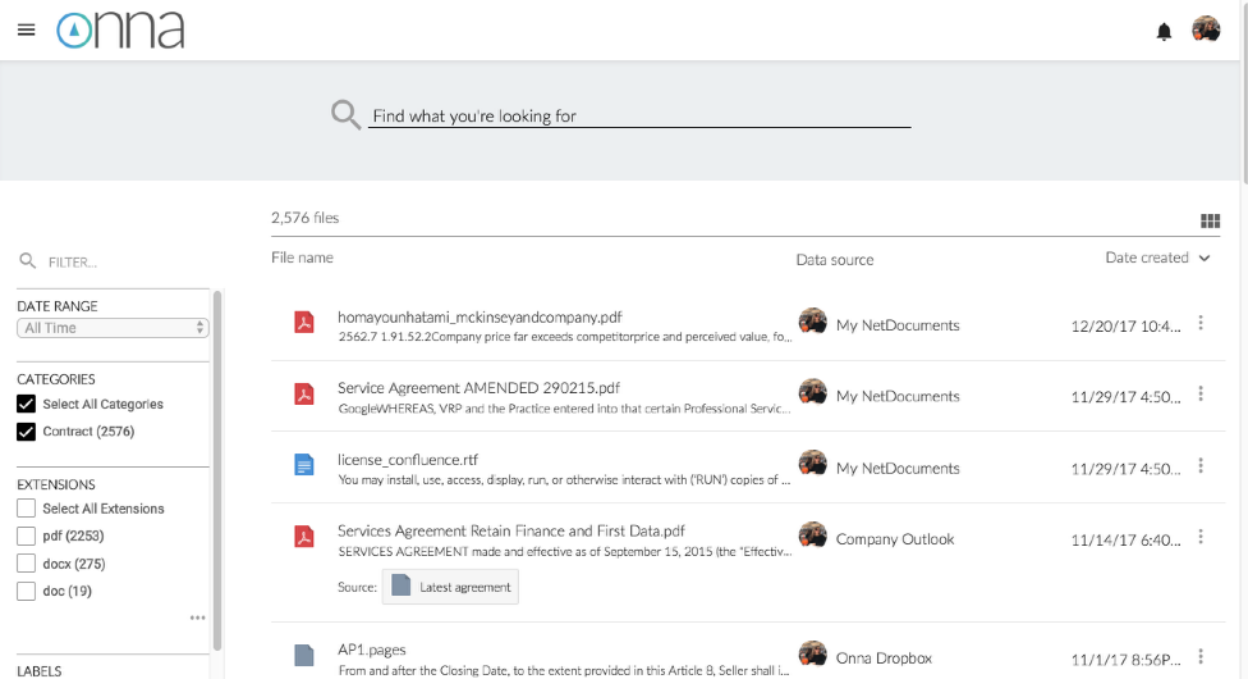
Consider the following hypothetical: In a patent infringement case, a litigation hold is issued over the “widget” at issue in the litigation. The records manager first identifies the relevant employees who developed the widget. The data sources include cloud-based email, Google Docs, CAD files, and social media from the party’s product launch.

Onna proactively indexes information on an ediscovery level at all times. This allows for Google-type indexing of data. If an email with a CAD drawing for an attachment is identified for preservation, Onna will recognize and issue code all similar CAD files and email messages for preservation.



The first element of proportionality is to be able to explain "the importance of the issues at stake in the action."[9] The topic modeling of Onna's deep learning empowers lawyers to perform "Instant Case Assessment" on the data in enterprise environments. This allows attorneys to identify data that can support their claims or defenses for their initial disclosures under Rule 26(a). If used proactively before a case is filed, attorneys could conduct a reasonable inquiry as required by Rule 11.

[9] Fed. R. Civ. P. 26(b)(1).



The speed of identifying relevant information with deep learning enables attorneys to focus on the real question of proportionality: What is the benefit of the data? Does the data help move the case forward? If not, then requiring the data to be exported for review is not proportional to the value of the case, because the expense of reviewing irrelevant data would outweigh its benefit.

Improve Document Review Efficiency

Data identified for preservation can be exported to a review application such as Everlaw or Relativity for document review. Onna's leverages natural language processing and OCR to make all data searchable with eDiscovery-level metadata integrity.

The goal of precision collection is to eliminate as much irrelevant data as possible. Moreover, relevant data can be loaded into a review application with relevancy and other issue coding, which in turn can be leveraged by predictive coding technology in a review application. This will help expedite review, so attorneys can focus on the merits of the case, and less time separating relevant ESI from irrelevant data.

Conclusion

Identifying relevant data from disparate data sources can be a challenge when litigation is reasonably foreseeable. Onna's deep learning search technology empowers knowledge managers enacting a litigation hold to apply new search parameters across

indexed data to quickly train the system for what is relevant to a new lawsuit. This allows for searching for relevant electronically stored information with precision, so only what is truly relevant to the claims or defense of a lawsuit can be exported for document review. This can allow attorneys to focus on the merits of the case, not driving up costs with irrelevant information in document review.

About Onna

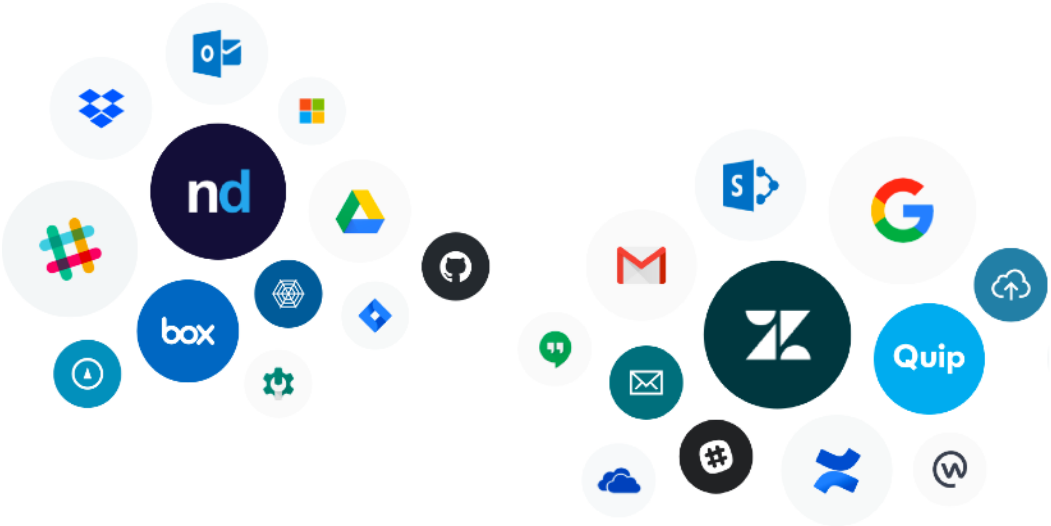
Onna is a platform for real-time search across multiple repositories that aids in eDiscovery and finding high-value items across legal departments. Integrations are built with major cloud services and storage repositories, such as G Suite, Zendesk, Dropbox, and Slack. Data is collected in a defensible manner and can be exported in standard eDiscovery format for assessment in document review platforms. Sources can be set to sync continuously, creating an up-to-date fully searchable environment. The platform can be hosted on the cloud or deployed on-premise.

Onna was named by Gartner as a Cool Vendor in AI for Legal Affairs.

For more information, please email contact@onna.com or visit https://onna.com

onna

Connect anything, so you can find everything



www.onna.com