**WHITE PAPER**

**Leveraging Cloud-Based Delivery Capabilities for Do-It-Yourself (DIY) eDiscovery**

Sponsored by: Kroll Ontrack

Vivian Tero
February 2012

**IDC OPINION**

eDiscovery is a business process whereby the technical controls for data management are supported by the implementation of various technologies. To date, there are primarily two technology architecture approaches for automating these core processes end to end: the single-vendor platform approach and the multivendor, multiplatform hybrid approach. The emerging do-it-yourself (DIY) eDiscovery approach is a permutation of the hybrid approach and leverages technology advances pioneered by the cloud or software-as-a-service (SaaS) delivery model. The key facts to note about DIY eDiscovery solutions are:

- DIY eDiscovery solutions enable an organization to utilize the infrastructure and enabling technologies of a third-party service provider while still employing the organization's own internal business processes and domain experts to execute and manage the eDiscovery process end to end.

- DIY eDiscovery solutions support multiple use cases under the single-vendor platform and multivendor hybrid eDiscovery infrastructure models.

- There are a myriad of options available for DIY eDiscovery, and not all offerings are created equal, especially with cloud or SaaS-based technologies quickly becoming viable eDiscovery platform alternatives. An organization should therefore scrutinize its options very closely, keeping in mind its litigation profile, the maturity of its processes and technical capabilities, resource constraints and IT strategy objectives, and the features, functionalities, and service capabilities of the third-party service provider. Buyers should also pay close attention to the differences in pricing and licensing models in the context of the features and capabilities of the third-party service provider as well as the provider's financial stability and security protocols.

- Legal matters change throughout the life cycle of the case. An eDiscovery matter that starts seemingly small and easy to handle can quickly escalate into something larger and more complex. Given this, DIY eDiscovery solutions will likely be only a part of any organization's approach as the flexibility to migrate from in-house tools to outsourced service provider solutions with minimal disruption and data risk will be critical, depending on the criteria of the matter type.

- An organization needs to engage in a strategic assessment of its needs and resources (internal and third-party enabled) using best practices, given this evolving eDiscovery landscape.
METHODOLOGY

The conclusions of this IDC white paper are based on qualitative interviews with corporate legal and IT stakeholders as well as on data from quantitative surveys of eDiscovery technology trends and investment priorities among business organizations.

IN THIS WHITE PAPER

This IDC white paper discusses how changes in the market and industry dynamics are driving organizations to rethink their current approaches to eDiscovery, focusing on better alignment of eDiscovery business processes with data management practices and technical infrastructures. The document concludes with a recommendation for potential customers to consider cloud or SaaS-based DIY eDiscovery technologies.

SITUATION OVERVIEW

eDiscovery Is in a State of Flux

The activities law firms and corporations employ to execute their eDiscovery obligations are very different from what they were 10, or even 5, years ago. Convergence in information technology, the increased digitization of information, and the 2006 amendments to the Federal Rules on Civil Procedure for electronic discovery were the catalysts that drove corporations and law firms to develop and adopt eDiscovery best practices. Continued shifts in market dynamics underscore the need to ensure that eDiscovery protocols and technologies are able to evolve to meet emerging data management, data segregation and privacy, and chain-of-custody obligations. The most notable of these developments include:

- Aggressive data volume growth. The average volume of electronically stored information (ESI) that needs to be preserved, collected, analyzed, and reviewed per matter remains widely distributed; however, among large matter sizes, IDC finds an increase in the proportion of litigants that average at least 2.5TB per matter. As shown in Figure 1, the proportion of corporations that average more than 2.5TB per matter increased by 17 points from 20% in 2008 to 37% in the beginning of 2011. Organizations must therefore have flexible eDiscovery process and data management capabilities that will allow them to easily scale up or down depending on the size and complexity of the matter at hand. Most organizations today worry that over-preservation means higher eDiscovery collection, processing, and review costs; on the other hand, aggressive disposition and non-preservation of ESI could potentially lead to sanctions — resulting in even higher costs. The rules are not as cut and dry as corporations and their law firms want them to be. The aggressive growth in data volumes and the introduction of new content types and data stores further exacerbate the current challenge for most corporations to strike the right balance between over-preservation and defensible legal disposition.
**Insourcing of eDiscovery activities.** An increasing number of corporate litigants are insourcing core eDiscovery process capabilities, specifically in the areas of identification, preservation, collection, processing, and first-pass review. These organizations are leveraging more mature processes and using a combination of tools and services for data retention, preservation, collection, and legal holds (see Figure 2). As a result of demands for flexible eDiscovery process and data management capabilities, organizations increasingly find themselves more actively involved in managing the eDiscovery life cycle, building out their ecosystem of law firm partners, third-party service providers, and technology partners. It should be cautioned that sometimes this results in a fragmented eDiscovery approach, with data being jockeyed from tool to tool and team to team, resulting in possible data loss or in the least inconsistent data handling practices. The entire life cycle should be considered when evaluating the total cost of litigation, balancing risk, security, defensibility, and cost efficiencies.

**Litigation trends.** The Gibson Dunn 2011 Mid-Year eDiscovery Update, which evaluated eDiscovery decisions in the federal courts from January 1 to June 15, notes that eDiscovery decisions are growing at an aggressive rate. The report finds that the number of decisions during the first half of 2011 rose 82% over the number of decisions in the prior-year period. The report also notes that "the number of instances in which litigants sought sanctions in the first half of 2011 was more than double the number in the same period last year (68 mid-year in 2011 versus 31 mid-year in 2010)." The increase in the volume of costly eDiscovery-related sanctions, such as the $1 billion court-ordered sanction in *E.I. du Pont de Nemours and Company v. Kolon Industries, Inc.* (E.D. Va., Jul. 21, 2011), is fueling demands for changes in the Federal Rules of Civil Procedure for electronic discovery in order to curtail the costs of data preservation and collection.

**Introduction of new technologies in the corporate infrastructure.** Emerging technology trends like social media applications, cloud computing, "bring your own device" (BYOD) policies, and smart mobile devices pose new challenges in the segregation and collection of ESI for eDiscovery. In the absence of a more definite legal standard for these new technologies, prudent organizations will proactively address the impact these media have on discoverability, regulatory compliance and corporate management, building business processes, and seeking technology partners to meet legal and regulatory obligations.
Figure 1

Average ESI Volume Collected per Matter

- Less than 500GB
- 500–999GB
- 1–2.5TB
- 2.51–5TB
- More than 5TB
- Don't know

Source: IDC, 2012
Established eDiscovery practices will enable organizations to execute information management and document management as well as comply with preservation obligations and discovery deadlines while ensuring security, confidentiality, and chain-of-custody requirements. A sound eDiscovery program is based on the following best practices:

- Policies, business processes, and technical protocols for information management and governance, including records management, retention, archiving, and disposition and data destruction. Organizations should also consider maintaining an inventory of ESI sources that describe the type of data, physical and logical location, retention periods, and native capabilities for search. This inventory information gives auditors, eDiscovery practitioners, and legal counsel the ability to determine and plan reasonable searches as well as better estimate the eDiscovery efforts and associated costs.
• Policies, business processes, and technical protocols for legal holds, including life cycle management of the legal hold and eDiscovery process, preservation management of associated ESI, and protocols and processes for releasing the legal hold. These practices will take into account jurisdictional and industry-specific requirements for privacy and data protection as well as IT operational objectives like business continuity, application upgrade, and hardware decommissioning plans.

• eDiscovery response plans, including the identification of an incident response and eDiscovery SWAT team and Rule 30 (b) (6) witnesses. This would also include plans for data collection, processing, analysis, and review, taking into account the type of matter, data volume, target sources, compliance and security requirements, and potential data protection roadblocks. Having the ability to understand capabilities and gaps would enable the organization to pinpoint the areas for instituting process and technology efficiencies, improving the quality of the results and identifying potential benefits or challenges of streamlining technologies in the corporate network.

• To mitigate potential challenges to its retention, legal hold, and eDiscovery response practices, an organization should plan on developing repeatable processes. Policy-driven automation is leveraged to ensure consistent enforcement of these policies and business processes. Automation also means that these activities are documented for audit, benchmarks, and process improvements. Policies and the associated workflow automation should allow for exceptions and permutations in the discovery and regulatory requests.

• Seek areas for effective and cost-efficient continuous improvement, keeping in mind that the landscape for eDiscovery is constantly evolving because of new requirements and the introduction of emerging technologies. Benchmarks and tests are maintained to measure the efficacy and efficiency of processes and enabling tools as well as forecast the efforts and costs associated with each matter. These benchmarks, which can be developed in collaboration with law firm and third-party service provider partners, should measure and monitor both internally and externally delivered eDiscovery capabilities. These benchmarks are also valuable for conducting risk analysis during the early case assessment process.

• Establish a process for conducting early case assessment and early data assessment to further drive data volume and cost reductions early in the eDiscovery life cycle.

IDC finds that close to 90% of organizations have existing information management and governance programs. Investment in the top 5 technologies employed to support information management and governance is at 40% or higher for the following technologies: message archiving, document management and imaging applications, file archiving, records management, and SharePoint archiving. Investments in technology infrastructure can be used as a proxy to gauge the maturity of legal hold and eDiscovery response practices. Here, IDC finds that less than 25% of organizations have invested in applications such as legal hold management applications, early case assessment applications, and search and analytics platforms to accelerate and simplify eDiscovery responses.
Organizations have highly distributed and heterogeneous IT environments, oftentimes managed by decentralized IT organizations. Translating core eDiscovery practices into executable policies, business processes, and protocols poses challenges in finding the right balance among multiple and, sometimes, competing objectives: providing timely responses to eDiscovery obligations, meeting application and business service-level requirements, complying with data protection and privacy rules, and meeting budget objectives. The challenge for most organizations is to find the right balance, which would allow them to optimize their internal and external investments in eDiscovery processes, infrastructure and technical capabilities, and domain experts.

**FUTURE OUTLOOK**

**eDiscovery Process and Architecture Approaches**

IDC views eDiscovery as a business process whereby the technical controls for data management are supported by the implementation of various technologies. The core eDiscovery technologies include archiving, enterprise content management (which includes document management, capture, and records management), search and discovery, and security applications. The main technology architecture approaches for automating these core processes end to end are (see Figure 3):

- **Single-vendor platform approach.** In this scenario, the archiving or enterprise content management platform serves as the on-premise data management and business process back end. For large/complex matters, organizations would typically own and manage the architecture to support the processes for information management to first-pass review but may employ an ecosystem of third-party service providers for attorney review, production, and presentation. For smaller matters, organizations may choose to manage and host the end to end process within their datacenters. Taking on the entire eDiscovery life cycle internally for all types of matters would require organizations to make significant capital and manpower investments. International compliance and data protection requirements and the emergence of new content types and data structures add another dimension to existing data segregation and collection challenges. A permutation of the single-vendor approach is the instance where an organization works with a single provider to execute the data management activities related to identification and preservation, all the way to production and presentation. Here, the organization works with a proven vendor that has the full suite of services and technologies to manage medium-sized to large matters. This approach increases data cohesiveness and reduces the data movements and access points, which can be risky in a multivendor scenario. The service provider should be able to provide the flexibility based on the unique matter requirements. It should be noted that this approach is almost similar to the traditional model, where litigant organizations turn over the entire eDiscovery process to service providers. The major difference with the new model is the ability to employ the DIY eDiscovery capabilities (depending on the profile of the legal matter) in addition to the full suite of services.
**Multivendor, multiplatform hybrid approach.** This scenario is usually the by-product of highly decentralized IT organizations and multiple business acquisitions and divestitures. An organization typically supports multiple archival and content management systems and has heterogeneous database and hardware platform environments as well as multiple geographically distributed IT organizations with complex data segregation mandates. Here, archival and content management systems support policy-based archiving and records management. Legal hold and eDiscovery workflows may be offered as features of the archival or content management platform, but a separate application operated by the organization’s internal team or an external service provider may be deployed to support search, identification, preservation, and collection in distributed endpoints. Potentially relevant ESI is copied to a secure SAN or server farms to be managed and maintained internally or by a third-party provider through a cloud or SaaS-based architecture.

**FIGURE 3**

*Source: IDC, 2012*
**DIY eDiscovery approach.** This emerging solution scenario is a permutation of the hybrid approach and leverages technology advances in the cloud or SaaS delivery model. Here, the organization utilizes the infrastructure and enabling technologies of a third-party service provider while still employing the organization's internal business processes and domain experts to execute and manage the eDiscovery process end to end. There are multiple use cases that could benefit from this model:

- Organizations with low volumes (such as 10 or less matters per year) of relatively simple and small matters (such as raw collection volumes of 50GB or less per matter and involving a small number of custodians) could simply pay for the capacity and processing power that they would consume. These organizations will typically not require the complex workflows and project management associated with fully outsourcing eDiscovery obligations to a third-party service provider.

- Organizations with a broad mix of large/complex matter types (involving one or more of the following attributes: multiple, multimatter custodians; multiparty/multiple jurisdiction; and collection from esoteric data types) and small/simple matter types will have relatively advanced eDiscovery process capabilities and could benefit from the economics of the cloud or SaaS delivery model. In this scenario, the organization would employ the cloud or SaaS service provider's infrastructure and enabling technologies to conduct and manage eDiscovery for the smaller and simpler matters, having the option of upgrading to the provider's full-service model should the need arise.

- Organizations that have invested in the single-vendor platform solution could leverage the DIY eDiscovery platform model for the following scenarios: handling unplanned increases in data processing capacity while still using internal domain experts augmented by short-term contractors; utilizing the DIY infrastructure to manage ESI, post collection and processing, for first-pass review through production and presentation for all matters; and leveraging the DIY infrastructure for first-pass review through production and presentation of large, multiparty, multijurisdictional matters, with the option to upgrade to the full-blown services should the need arise.

**Selecting the Appropriate eDiscovery Infrastructure: Key Considerations**

There are a myriad of options available for DIY eDiscovery, and not all offerings are created equal. Organizations should therefore scrutinize their options very closely. The selection of the appropriate approach is influenced by the following factors:

- **The organization's litigation profile.** This profile includes the types and classes of matters, target data sources and eDiscovery collection challenges, average collection sizes per matter, average duration per matter and per legal hold, number of custodians per matter, complexity of eDiscovery workflows and reporting, jurisdiction and data segregation challenges, number of concurrent users/reviewers per matter, preferences in application delivery models, and geographic footprint. The culture and appetite of the organization's legal and eDiscovery project teams for using disruptive technologies to realize process improvements should not be overlooked.
The organization’s existing process and IT capability maturity for eDiscovery and broader information governance programs. Organizations want to leverage their existing IT investments in information governance and eDiscovery and be able to augment and address capability gaps that arise from unforeseen increases in data volumes and matter volumes, new data architectures, and emerging compliance and data management requirements.

The organization’s long-term IT architecture strategy. This strategy includes current and planned investments for the organization’s applications, hardware, and network environments. Organizations should assess the constraints and requirements that cloud computing, virtualization, and mobility initiatives impose on their eDiscovery capabilities. Organizations should look to form IT teams that can support in-house to outsourced migration with minimal disruption and data risk.

The business and technical attributes of the third-party provider. Organizations should pay close attention to the technical features- including security parameters-offered by the eDiscovery infrastructure provider, the pricing and licensing options available, reputation and industry expertise, and the ability to support data migration. eDiscovery is a collaborative effort, so organization and cultural fit between the client organization and the potential third-party partners should be carefully evaluated.

**Kroll Ontrack’s Verve**

As a traditional full-service eDiscovery provider, Kroll Ontrack — with its high-quality reputation as the "gold standard" service offering at market competitive rates — has long been a dominant player in the market. Historically, the company has focused on building its global services business and enhancing its suite of products that support its services — a review platform (Ontrack Inview) and an early data assessment platform (Ontrack Advanceview). Kroll Ontrack processes approximately 800GB of eDiscovery data daily, with a single day high of 2.3TB, upward of 49 million pages of load file data generated daily, and more than 20PB of data hosted in its U.S., U.K., and Asia/Pacific datacenters. Customers consistently comment that Kroll Ontrack consultants and professional services staff possess deep knowledge of both technology and legal matters.

In late 2011, Kroll Ontrack announced the release of a new technology venture — Verve. This product marks the company's departure from a pure-services business model, bringing to market a product that decouples the Kroll Ontrack service engine technologies and forms a DIY eDiscovery software offering. Key features of Verve include:

Verve offers a unified platform for uploading the ESI, selecting the processing criteria, and conducting early data assessment, review, and production. Customers will have access to data processing, Verve EDA (for early case assessment), and Verve Review (for review, analysis, and production). Verve users have ultimate control over their data, using one remotely hosted platform — instead of several disjointed technologies — to conduct processing, filtering, review, and production. The technology is highly scalable, allowing users to manage small or large matter types without additional infrastructure costs and support burdens.
Verve offers simple and affordable pricing frameworks. The pricing models are intended to give clients predictability in their eDiscovery spend, despite variability in litigation volume month over month: "pay as you go" is based on a combined per-gigabyte model, and the subscription model is based on an annual volume usage consumption. The price only includes access to Kroll Ontrack eDiscovery technologies. Customers that require professional services for data management and case and project management will be referred to the traditional Kroll Ontrack eDiscovery services.

Kroll Ontrack purposefully chose to launch the Verve DIY eDiscovery platform using new, secure cloud or SaaS-based technologies rather than seeking to compete with other on-premise eDiscovery tools that require sizable up-front investments in hardware, installation, and training. Depending on the litigation profile and data management needs of the organization, a cloud or SaaS model could potentially deliver lower total cost of ownership, given that the remotely hosted platform requires no additional hardware, software, or IT support resources.

Verve is designed to complement any existing eDiscovery software investments already implemented by an organization. Users can leverage some or all of the processing, filtering, review, and production technologies within Verve.

Backed by Kroll Ontrack datacenter investments and experience in handling complex data segregation and data security requirements, Verve leverages the most modern security protocols.

Should an eDiscovery matter become out of scope for an organization’s internal experts, Kroll Ontrack has the ability to provide customers with a seamless transition to its full-service eDiscovery offering, obviating the need to transfer data onto a new platform and service provider. Kroll Ontrack professional services agents start up where the organization's internal team leaves off, maximizing consistency, efficiency, and defensibility on the eDiscovery matter.

Verve customers will be automatically upgraded to the latest product versions because of the Web-based nature of the cloud or SaaS platform.

Given the evolving eDiscovery landscape, many organizations face a fragmented solution set that hinders instead of aids in accelerating and simplifying eDiscovery responses. It is imperative that organizations deploy eDiscovery strategies that include a reputable partner with deep DIY and full-service solution sets, augmented with the flexibility to migrate any DIY matter to an outsourced full-service matter with minimal disruption and risk. Further, streamlining third-party service providers and technology platforms will result in less complexity and less cost for organizations. Driving down the total cost of litigation through the continuity of a single solution provider for corporations and law firms with secure, high-quality, and dependable technology and service solutions is the core mission of Kroll Ontrack.
CHALLENGES AND OPPORTUNITIES

Customer Challenges

IDC has identified the following challenges for potential buyers:

- Gauging their internal information governance and process capability and technology maturity (Conducting an objective audit and assessment of internal capabilities and weaknesses can be time consuming and fraught with organizational politics.)

- Aligning competing priorities among the key stakeholders: legal, IT, compliance, and business leaders

- Aligning eDiscovery process and technical priorities with broader IT strategy and operational objectives

- Keeping track of and planning for emerging eDiscovery challenges that arise from the rollout of new corporate initiatives such as social media marketing, cloud computing migration, and "bring your own device" programs

Challenges for Kroll Ontrack

Kroll Ontrack has always been perceived as a high-end full-service eDiscovery provider. Verve decouples the Kroll Ontrack software technologies from its high-value professional services.

Kroll Ontrack is not the first provider to engage in such a strategy, given the changing market dynamics. Past attempts by other competing eDiscovery service providers to delineate their full-service eDiscovery offerings from their technologies/capacities were not well received by potential customers. Kroll Ontrack could potentially do well if it avoided the mistakes of previous competitor offerings, such as:

- Market confusion on the pricing and licensing models (Previous offerings by other competing providers did a poor job in communicating the benefits and cost savings between a full-service and a cloud-based DIY service offering. Potential customers were also unable to gauge the differences in benefits between on-premise software and the DIY cloud or SaaS-based eDiscovery offerings.)

- Inability to clearly articulate the use case scenarios and the associated litigation profiles that were most suitable for the DIY eDiscovery approach

- Inability to clearly articulate the economics of a cloud or SaaS-based DIY architecture relative to other options (such as on-premise software or full-service eDiscovery models)

- Immature features, workflows, and reporting functionalities for various eDiscovery activities (such as early data assessment, early case assessment, and analysis, review, and production)
CONCLUSION AND RECOMMENDATIONS

Today, IDC finds that close to 64% of organizations are using multivendor hybrid approaches, while the rest of the organizations are gravitating to single-vendor platform solutions for their eDiscovery business process and data management activities. The majority of these multivendor eDiscovery architectures are the outcomes of business and IT organizational developments such as mergers and acquisitions, changes in IT strategies, and siloed and reactive purchases of ad hoc technologies. But differences in the matter profiles, information governance and eDiscovery capability and process maturities, geographic footprint, and eDiscovery obligations suggest that there are opportunities for realizing process and cost efficiencies by leveraging DIY eDiscovery tools across these approaches.

Organizations that are in the process of selecting the solution that best fits their needs are exhorted to keep in mind the following:

- Current and future litigation profiles
- Current process and technical capabilities and gaps in the context of potential changes to the organization's litigation profiles
- Ability for the solution to align with and complement the organization's overall IT infrastructure investments and strategies, including strategies for "future proofing" the eDiscovery programs
- Platform options in the context of resource and budget constraints, taking into account the advantages and disadvantages of cloud or SaaS-based solutions versus on-premise software solutions
- Features and functionalities offered by competing solutions along the eDiscovery process value chain (collection, preservation, filtering, and early data assessment, processing, review, production, and security) in the context of pricing and licensing offered by competing solutions
- Pricing and licensing structure, paying attention to subtle differences in pricing specific line items (Buyers should also assess how the pricing models stack up against on-premise software solutions.)
- Alignment of the third-party provider's service capabilities with the organization's vendor management strategies

At the end of the day, organizations need to engage in a strategic assessment of their needs and resources, given this constantly evolving landscape. Organizations should consider leveraging infrastructure and processes to efficiently conduct information and document management, meet preservation requirements and discovery deadlines, and ensure data security and chain-of-custody documentation.

This means possibly utilizing cloud or SaaS-based DIY eDiscovery technologies in addition to service providers for eDiscovery obligations. Engaging in DIY eDiscovery may be new to many organizations, but this option should be fully considered by organizations amid the changing market.
Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2012 IDC. Reproduction without written permission is completely forbidden.