



## BUYER CASE STUDY

### A United States–Based Telecommunications Firm Employs FTI Harvester to Address eDiscovery Challenges in Microsoft SharePoint

Vivian Tero

#### IDC OPINION

Microsoft SharePoint integrates the major information work tools — search, content management, business intelligence, workflow, and messaging — to facilitate collaboration. With this application, the information worker is able to easily define and change users, add and change content, and change access levels and display formats. This highly dynamic environment creates problems when an organization has to preserve and collect data from SharePoint sites in response to internal investigations, ediscovery, and regulatory requests for information. Documenting chain of custody, mitigating spoliation risks, addressing index lag, incomplete search results, and minimizing overcollection (a result of SharePoint sprawl) are issues that organizations must plan for. Organizations should also be aware of the limitations and challenges of point-in-time search and collection tools. To address these ediscovery obligations effectively, organizations are advised to:

- ☑ Assess if the organization has sound data retention and disposition practices as well as legal hold and ediscovery response protocols. These policies and protocols should extend and be adjusted for the unique requirements of dynamic environments such as collaborative applications and workspaces.
- ☑ Understand the configuration of the dynamic application and its implications on meeting chain-of-custody requirements.
- ☑ Evaluate tools and services that would enable your organization to address the limitations and risks when using point-in-time search, preservation, and collection tools in dynamic environments.

#### IN THIS BUYER CASE STUDY

This IDC Buyer Case Study discusses the use of FTI Consulting's Harvester solution by a United States–based telecommunications company. The organization employed Harvester to facilitate ediscovery collections within its SharePoint repositories.

## SITUATION OVERVIEW

---

### Organization Overview

IDC interviewed the ediscovery and litigation associate of Telco X's outside law firm on February 2, 2011. (The names of the corporate litigant and outside law firm are withheld upon request.) Telco X is a multibillion-dollar telecommunications organization with approximately 26,000 employees. The corporation is headquartered in the southern region of the United States. Employee and HR-related disputes, privacy, and antitrust investigations make up the bulk of the Telco X's legal matter events.

---

### Challenges and Solutions: Meeting eDiscovery Obligations for Highly Dynamic Data Repositories

The following key facts prompted Telco X to employ FTI Consulting's Harvester solution:

- ☒ Telco X's dispute with several current and past employees involved corporate policies and practices related to compensation for overtime, travel, and loading and unloading company vehicles.
- ☒ The corporation's business operations policies (including HR practices) were published in the employee SharePoint site. Telco X had initially deployed MS SharePoint 2003 to host and manage the HR site and was in the process of upgrading to MS SharePoint 2007.
- ☒ Telco X was obligated to preserve, collect, and produce in discovery its corporate HR policies and practices that were accessible to the employees at issue through the designated MS SharePoint site.
- ☒ Telco X was not obligated to document individual employee access and views, but it had to document when and where HR policies were published, modified, and made generally accessible to employees. This included capturing information on document version control; the authors, editors, and records managers of relevant files (records custodians); information on the physical and logical system administration of the relevant HR files (system custodians); and information on all documents that employees had access to within the corporate HR site.

### ***eDiscovery Challenges with Microsoft SharePoint***

Microsoft SharePoint poses data collection and chain-of-custody challenges:

- ☒ **SharePoint "sprawl" stymies the efficacy of search and indexing efforts as well as contributes to tendencies among producing parties to overpreserve and overcollect data.** Microsoft SharePoint is designed to address a need for an enterprise platform that integrates the major information work tools — search, content management, business intelligence, workflow, and messaging — into a collaborative work framework. The application allows an information worker to

easily define and change users, add and change content, and change access levels and display formats. These attributes make it very difficult for IT administrators to monitor and manage SharePoint sites and contribute to the problem of SharePoint "sprawl." In addition to taxing storage operations, application performance, and network bandwidth, SharePoint sprawl presents search and indexing challenges. In many instances, indexing may not be turned on or it may not be up to date, resulting in incomplete search results. Some data formats do not contain searchable data. Litigants tend to err on the side of caution and end up preserving more data than necessary. Preservation and collection under these circumstances could easily yield massive volumes of duplicate and near-duplicate data.

- ☒ **Dynamic environments also lead to overpreservation and overcollection, especially when ongoing collection is required.** Microsoft SharePoint is highly dynamic. Since the application is designed for collaboration, it is not unusual for a particular file to be modified several times within a 24-hour period. When a file is deleted or modified, the prior version may no longer exist in the production environment, but copies of the old versions (which may be relevant to the matter) may reside in another site, in the backup application, or "offline" on an employee's computer.

Also, there is typically a time lag between the date for defining the scope of the preservation and collection and the actual date of preservation and collection as procedures, issues, and search terms are defined and tested. During this period, multiple changes are happening to the documents within a SharePoint site. Mapping and tracking the relationships across these relevant data stores over a time series can be onerous and expensive. Some organizations have resorted to preserving and collecting snapshots of SharePoint sites over a time series. This approach is onerous and expensive.

- ☒ **Dynamic environments pose challenges in documenting chain of custody.** SharePoint's integration with active directory provides information on the employees' roles, access rights, and permissions. Employees can easily create, modify, delete, and add content, but the SharePoint management system is not designed to track the relationships between the various roles, identities, and activities performed on a specific document over a time series. For example, if Jane created Telco X's HR policies document in MS Word, Tina makes a PDF copy of the same document and stores this in another SharePoint site, Jack uploads a copy of the document into the SharePoint site, and Sue modifies the document, deletes the original version stored in Telco X's SharePoint site for HR policies and operations, and then loads an updated version of HR policies — at that point in time, SharePoint is only aware of the updated version and will list Sue as the creator of the file. SharePoint will have no knowledge of the prior versions and the contributions of Jane, Tina, and Jack in the creation of the final document. There are challenges in documenting chain of custody and in properly preserving relevant data under this scenario.

- ☒ **New environments are presenting new formats of data, which do not conform to "standard" document types.** SharePoint allows for various types of data, from videos to calendar notes, many of which are common in today's

workplace. Because SharePoint has such a customizable interface, data doesn't have to be a "document." It can be as simple as a survey response, or link to a Web site outside of the environment. These new data formats present challenges in collection and presentation format for review and production to opposing counsel.

***Solution: Employing FTI Harvester to Address Point-in-Time Collection Obligations***

Telco X worked with FTI to craft a solution that would allow it to:

- ☒ Limit the scope of the SharePoint collection to the most relevant custodians and date range. Given the dynamic nature of SharePoint, to balance the proportionality of burden and cost with the obligation to identify, preserve, and collect responsive information, Telco X identified a specific date range from which to collect relevant data from its HR SharePoint site.
- ☒ Address limitations of keyword searches in dynamic SharePoint environments, thus avoid missing responsive documents by focusing on relevant custodians.
- ☒ Document the names and locations of relevant files that all designated Telco X employees had access to, including their rights to change, create, modify, delete, and copy these files.

**FTI Harvester**

FTI Harvester is an integrated technology and services offering to facilitate the preservation and collection of documents from SharePoint. Harvester offers predictable pricing, up front. The software tool can be deployed behind the firewall and connected to the network, managed remotely, or implemented from a backup of the environment. Harvester is designed to offer the following benefits to its clients:

- ☒ Identify documents by custodian-specific metadata such as access rights, authored, or last modified and support custodian-based preservation.
- ☒ Identify relevant file information such as documents by site, file extension, and date range (for site).
- ☒ Use a combination of SharePoint Security, Microsoft Active Directory, and FTI's proprietary technology and services to find and preserve documents relevant to a legal matter without disrupting security policies and network operations.
- ☒ Employ an interactive review process with corporate legal counsel to fine-tune the scope of the collection before actual harvesting of the data is executed.
- ☒ Provide a data export that can be imported into any litigation review tool. The data export is augmented with SharePoint information.

---

**Results**

As a result of this approach, Telco X was able to achieve the following objectives:

- ☒ Identify, preserve, and collect information relevant to the legal matter while minimizing the risk and cost of recollection due to issue and keyword changes, as well as incomplete efforts in a large dynamic environment.
- ☒ Document chain of custody and mitigate spoliation risks by mapping the relationship of relevant documents to specific custodians based on authorship (creation and modification); roles, access rights, and permissions (and associated changes as the custodian's role within an organization evolves); the custodian's network user names; and groups that the custodian belongs to.
- ☒ Develop a legally defensible practice for responding to ediscovery obligations in dynamic environments like SharePoint.
- ☒ Cut down the time and effort required to preserve and collect live SharePoint data.

## ESSENTIAL GUIDANCE

Corporate litigants that face ediscovery obligations involving data captured and stored in dynamic environments are advised to take note of the following:

- ☒ Understand the configuration of the application and its implications on meeting chain-of-custody requirements. Will the organization be able to effectively map the relationship between the documents and all the actors that are authorized to perform various activities? Can this relationship be tracked over the life of the document?
- ☒ Document your organization's data retention and disposition policies. Do these policies and protocols extend to dynamic environments like SharePoint and other collaborative applications? Do these policies and protocols extend to conducting search, preservation, and collection of data across dynamic environments? If not, investigate plans to expand support into these applications.
- ☒ Document and understand the limitations of the organization's existing search and indexing architecture when it comes to investigations and ediscovery in dynamic environments. Investigate tools and resources that can facilitate the creation of data maps in these applications.
- ☒ Define a process (software and expertise) to identify and preserve potentially responsive documents in these environments. Ensure this process extends through the review and production stages.

## LEARN MORE

---

### Related Research

- ☒ *Worldwide Governance, Risk, and Compliance Infrastructure 2010–2014 Forecast: Increased Regulatory Oversight, Privacy, Cloud Computing, and Smart Cities Drive Emerging GRC Obligations* (IDC #222214, March 2010)

## **Copyright Notice**

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, telebriefings, and conferences. Visit [www.idc.com](http://www.idc.com) to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit [www.idc.com/offices](http://www.idc.com/offices). Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or [sales@idc.com](mailto:sales@idc.com) for information on applying the price of this document toward the purchase of an IDC service or for information on additional copies or Web rights.

Copyright 2011 IDC. Reproduction is forbidden unless authorized. All rights reserved.