

Information lifecycle governance for the oil and gas industry



Contents

- 2 Executive summary
- 3 Business challenges with enterprise information
 - 3 Veracity of data is a critical concern
 - 4 Information silos increase data management issues
 - 4 Regulation and litigation
 - 6 Data privacy
 - 6 Defensible disposal and data preservation
- 7 Address data overload with information lifecycle governance
- 9 Use case: ILG for data in upstream operations
 - 9 Challenge
 - 10 Solution
 - 10 Benefits
- 11 Conclusion

Executive summary

This white paper discusses the importance of information lifecycle governance (ILG) for oil and gas companies, particularly with respect to big data environments. The insights promised by a big data deployment can help organizations make business decisions and set business direction; however, in practice the sheer volume of noise in data sets promoted for analysis can lead companies to miss important insights. Worse, the data being analyzed may come from an untrusted source and lead companies to make the wrong decision. A properly implemented ILG program can help organizations avoid both these scenarios while reducing corporate risks and costs.

An ILG program helps shepherd data to the right place throughout its useful lifecycle while simultaneously acting as a scale to balance cost and risk with value. Costs include the expense of data storage in the present and future, but also the cost of producing data for any potential legal matters, which applies to data in its native location and data held in big data environments. In addition, there is the risk of important intellectual property (IP) and personal information becoming part of a data breach.

On the other side is the value the data provides. Is it needed for a legal proceeding or regulatory statute? Does it provide the business with current value or carry the potential to inform future insights? If the cost and risk of keeping a data item ever outweigh the value, the organization should consider that data for disposal.

This paper will demonstrate that ILG is an important step on the road to deriving higher-value insights from big data, achieving business efficiency and ensuring regulatory compliance in the changing and increasingly global oil and gas industry.

Business challenges with enterprise information

Oil and gas companies face many challenges related to the growth, management, regulation, disposal and effective use of their enterprise information. These companies need to ensure vital documents, including approvals for drilling, building and maintaining wells, are available throughout the enterprise and across enterprise boundaries to minimize risk and ensure regulatory compliance. They have to be sure this documentation, along with its audit and transactional process history, is preserved per company policy and regulatory requirements. In addition, they require timely reporting on operations and accidents as well as risk mitigation plans for critical operations such as drilling. Finally, they must manage all the data generated across different operational units and produce actionable insights to improve business efficiency.

Faced with the business and technological realities of addressing these challenges, many oil and gas companies are investing in big data environments in an effort to reduce operational risk, and enhance and optimize processes across the enterprise. However, big data on its own is not a cure-all solution; it must be complemented by an ILG program for full effectiveness.

Veracity of data is a critical concern

Oil and gas companies accumulate petabytes of information that, with analytics, can improve the efficiency, safety, productivity and cost-effectiveness of their operations. The volume, velocity and variety of that data is driving deployments of big data environments—and the analytical synergy of the combined data sources in these new systems provides valuable insights that inform the actions businesses take.

However, one aspect of big data that is not always incorporated in industrial-sector big data environments is the fourth “v,” the concept of veracity. How trustworthy is the data that is being included in the analysis? Real-time analytics requires close synchronization between operational data and harvested, historical data. Data generated by sensors on equipment and in wells needs to be verified, as erroneous data can lead to poor conclusions during surveillance and impair decisions based on models.

Besides the lack of trust in the data, the level of risk is also often unclear. How great is the risk if you include this data from the beginning of analysis? How does that risk change if you then retain the data? These are crucial assessments: How can you take an action if you cannot trust the data that led you there?

“Trust, combined with risk assessment, is essential before action can be taken.”

—From “Harness the Power of Big Data” by Paul C. Zikopoulos, Dirk deRoos, Krishnan Parasuraman, Thomas Deutsch, David Corrigan and James Giles

Information silos increase data management issues

In the oil and gas industry, information architecture is often fragmented across legal entities, systems and processes. Most companies have multiple information silos such as data warehousing, data marts, Microsoft SharePoint servers, content repositories and custom point solutions that support disparate yet parallel activities—for example, separate systems for upstream and downstream operations. In the past, industrial operations tended to invest in departmental solutions that were heavily customized, creating additional overhead and information silos that were disconnected from the rest of the enterprise.

These legacy approaches mean aggregating risk information across the organization is typically an ad hoc task, which makes it difficult to provide timely information on risk exposure to business managers. In addition, traditional siloed approaches to discovery, records management and data management are inadequate for the current business environment's high volume of data and its distributed nature.

Because each business area typically considers its concerns and potential solutions in isolation, locating, managing and establishing a single point of truth in multiple disconnected enterprise data silos becomes a real challenge. Unfortunately, it often results in information loss and/or data duplication. Compliance, commercial disputes and claims are also difficult to handle in environments that have multiple sources for the types of information required by regulations and standards.

As a consequence, the percentage of understandable data—and decision makers in the organization who trust the information—is on the decline, as is the number of users who can identify the right information and its location.

Regulation and litigation

The oil and gas industry is considered to be a critical and complex system, linked to many other systems including transportation, retail, environment, research and development, distribution, water and food. All of these factors add to the businesses' regulatory burdens. Additionally, various countries have fundamentally different regulatory approaches. This often drives oil and gas companies to segment their business and systems to address regulations—a process that ends up making information integration much more difficult.

Litigation adds further pressure for the industry, and any data kept for big data analysis is subject to litigation requests. There is no sign that this pressure will ease in the years to come:

- The development of new oil fields is continuous, so there is a constant potential for new litigation regarding the terms of leases in the oilfield, new development, or environmental laws and regulations.
- Increasing globalization and tighter regulations increase the complexity of eDiscovery, which is also exacerbated by data growth.
- Poor data management processes and old technology can significantly impede efforts to make sure the right information is easily available for eDiscovery.



Human and environmental safety and health protection remain top priorities for the oil and gas industry. Timely reporting on operations and accidents is required, with risk mitigation plans in place. Oil and gas companies must ensure all vital documents, such as approvals for specific operations, are available for regulatory review. Regulations from the US Environmental Protection Agency (EPA) and Occupational Safety and Health Administration (OSHA), or from Canada's Environmental Protection and Enhancement Act (EPEA) and Environment and Sustainable Resource Development (ESRD) are just some of the regulations with which oil and gas companies must comply. The 2010 BP blowout on the Deepwater Horizon rig and the subsequent investigation reminded the industry about the importance of ensuring the correct information is readily available in the event of a disaster, as it affects a company's position with regulators, the general public and its stockholders.

The impact of regulation is easy to overlook in the excitement of big data implementations. If the information is regulated in the source systems, it will likely be regulated in the big data environment as well. Even if the information is not regulated on its own, the business decisions driven by the insights produced in such an environment could cause the data and analytical model to be regulated. This is why it is important to have the right technologies to facilitate environmental, health and safety-related compliance activities (such as reporting and inspections) and to automate information management across the organization.

To comply with increasing regulations, oil and gas organizations must ensure all information relevant to various regulations is available whenever and wherever it is required. Crucial information pertaining to areas such as design, construction and modifications—process activities that are critical for regulatory and legal compliance—must also be appropriately archived.

Data privacy

Individual information privacy, workplace information privacy and data breach notification have important implications in both traditional data environments and in big data environments. Companies need strong security and the ability to identify sensitive information, safeguard it and dispose of it when required by regulations, no matter where it exists.

In the US, OSHA requires organizations to capture information for cases that are caused by conditions or exposures arising in the work environment, but it also requires special handling for some types of data to protect workers who have experienced occupational injuries and illnesses that raise privacy concerns. Even with the substantial reporting requirements imposed, OSHA establishes stringent requirements for records access to make sure the injured or ill worker's privacy is not compromised. To decrease cost and risk, organizations should stop keeping all their data forever. Instead, they should hold on to the data only as long as required, or as long as its business value outweighs its cost and potential risk.

Defensible disposal and data preservation

Because of oil and gas companies' heightened regulatory compliance responsibility, it will become even more important for businesses to determine what information meets "defensible disposal" criteria. Defensible disposal is a process (manual or automated) to discard unneeded or valueless data in a way that will stand up in court as reasonable and consistent. Eligible data includes information that does not need to be maintained for current or foreseeable legal matters, or for regulatory purposes, and information deemed unnecessary for the ongoing transaction of business.

Identifying and defensibly disposing of this information prevents it from being exposed in a data breach, helping to reduce risk. It also removes the chance that this information may be subpoenaed for litigation and subject to eDiscovery processes. Systematic deletion of non-record data must still be accomplished in a manner that ensures protection of any private data as defined by the law.

Oil and gas companies involved in unconventional exploration (such as "fracking," drilling operations for oil found in low-permeability rock and sandstone) must navigate not only the challenge of differing state and local laws, but also a constantly evolving legal landscape as landmark cases make their way through the courts. Some of the most common allegations in such cases involve environmental or product liability regarding negligence, nuisance, personal injury, property diminution or damage to property during drilling.

Damage to the company, both financial and reputational, is exacerbated when the company fails to preserve important evidence, or over-preserved—creating an abundance of non-relevant data when the company's attorneys search for relevant information. For example, there were several instances of evidence destruction in the wake of the 2010 Deepwater Horizon disaster, including a Halliburton technology director and a senior program manager who destroyed computer models, leading to a criminal conviction against Halliburton¹ and a former BP engineer who was convicted of destroying incriminating text messages and emails.²

It's important to have technology that will identify such information and archive it based on an approved retention schedule, as well as delete information that does not have to be preserved.

Efforts to remove digital information must be thorough. Multiple copies of the same data often exist throughout an oil and gas organization, as well as copies in backup files and archives. This duplicate data can cause great difficulty for attorneys who review the information collected for a case page by page. If different attorneys on the case review this information because it appears multiple times in the collected data (for example, one copy in the archive, one copy on a hard drive and one in a backup tape), it could be marked by those attorneys in contradictory ways and be inadvertently produced to the opposing side. That is why it is critical to develop proactive document preservation and destruction programs that not only identify what information should be preserved (and for how long) and what should be deleted (and when), but also suspend destruction and send notifications to responsible parties when required.

Address data overload with information lifecycle governance

Oil and gas companies collect massive amounts of data, including data from instrumented well and field operations, well logs from an initial drill, asset supervisory control and data acquisition (SCADA) reporting, design and construction engineering data, and engineering documents from projects that feed operations. The big data environment could also contain a massive collection of business data, such as internal financial results and news on energy and petroleum competitors bidding on leases and making major capital investments.

Because of the huge amounts of data involved in today's operations, organizations must change their existing organizational processes and the way they treat information by instituting proper governance in the big data environment with an information lifecycle governance (ILG) program that supports the following goals:

- Balance the need to have the right data for current and future analytical models with the need to reduce the cost of litigation support
- Apply appropriate retention schedules to mitigate the risk of privacy or other exposures
- Improve data quality and the trust in analytical results by removing redundant, obsolete and trivial data
- Improve regulatory compliance with proper document classification and retention
- Streamline litigation response by automating custodial notification and data preservation while applying analytics early in the process, leading to more-informed decision making and review of the smallest legally defensible set of data

What is ILG?

Information lifecycle governance (ILG) is the process of analyzing data to determine its current or potential business value as well as the appropriate action to take: continue normal usage, archive to lower-cost storage, retain for regulatory purposes, hold for litigation or defensibly dispose of data that has outlived its purpose.

To get maximum effectiveness and benefit from a big data environment, follow these fundamental information governance steps:

1. Obtain executive sponsorship early, emphasizing the importance of ILG across the organization as a foundational layer for trusted analytics and improved insights.
2. Determine data sources that need to be governed and what that governance must include. Is the data sensitive from a business, legal or privacy perspective? Does the data need to be standardized?
3. Identify the ownership of the data sources at the executive level.
4. Establish an organizational blueprint that includes business, IT, legal, records and risk/privacy departments.
5. Define metrics and targets and implement measurement of those metrics. The metrics tracked and reported will tie closely to the business value and trustworthiness of the information.
6. Establish and enforce governance policies. Policies without enforcement cannot withstand scrutiny in the event of litigation or regulatory audits.
7. Implement governance technology to achieve the highest maturity level, automation and user transparency. This will drive adoption of data sources as trusted repositories.

The IBM ILG solution helps oil and gas organizations define, manage and automatically execute governance policies across the enterprise. This includes legacy data cleanup, value-based archiving, records retention, efficient eDiscovery and defensible disposal policies. Applying the IBM ILG solution across different data sources—such as structured data in databases,

unstructured electronically stored information and emails—helps lower legal risks and improve information economics by addressing the following needs:

- **Engineering relies on unstructured data, while other parts of the business depend on structured data systems.** All types of information need to be integrated to present a unified view across the entire project lifecycle. ILG creates an environment to integrate the right data, eliminating duplicates and sharing the information across the enterprise, with accountability and ownership based on company policies.
- **Users need to know what data can and can't be trusted.** Oil and gas companies must prevent information loss between phases in the asset lifecycle and eliminate the practice of sending incomplete packages of procurement information to vendors. The IBM ILG solution establishes a single point of truth for engineering and other types of information, helping to ensure that all approved workers can access accurate, timely information from an approved location.
- **The risk associated with storing and retrieving documents in environments where regulatory compliance relies on individual employees is too high.** The IBM ILG solution provides automatic document lifecycle management features, including data identification, data classification and automated policies, that support the creation of consistent, repeatable records management processes that help oil and gas organizations demonstrate compliance.
- **Immature business processes related to document disposal lead to increased storage costs.** The IBM ILG solution allows information stakeholders to identify and analyze data in place as well as safely dispose of data, improving process maturity and slowing the rapid rise of storage costs.

The IBM ILG solution also helps reduce data duplication by curating data prior to ingestion, turning huge volumes of data into relevant information that can help companies make business-critical decisions.

Ultimately, a properly implemented ILG program will provide trusted data as a foundation for analytics by eliminating non-relevant data and keeping relevant, higher-quality data. This is a critical element in optimizing data analytics, and strong governance is a key contributor to improving quality.

Use case: ILG for data in upstream operations

During pre-license prospecting for reserves—one of the most important events in oil and gas companies' operations—an oil and gas company collects geological, engineering, economic and topographical information, including satellite imagery, aerial photographs, gravity-meter tests, and magnetic and other observations or measurements. The company also conducts various types of explorations, which are likely to include topographical, geological, geochemical and geophysical seismographic studies and exploratory drilling. These upstream operations at a global oil and gas company involve multiple phases that all generate large numbers of documents.

Typically, each phase depends on information generated in previous phases, requiring a lot of information to be retained. For example, engineering drawings are used for many purposes during upstream operations, from conception to completion of construction, through facility maintenance to demolition and eventual return to the natural site. All company operations are highly regulated, with industry bodies mandating tight controls and proof of process for any changes on production platforms.



Challenge

An oil and gas company had traditionally used file shares and departmental or divisional content repositories for upstream operations data. But as the organization expanded and as new processes and tools generated more and more data, it needed to gain control of its ever-increasing information. Manual search, retrieval and modification of thousands of engineering documents from internal and external sources resulted in delays, reworking and production downtime. Some legal matters required discovery across the entire global organization, involving over 5 PB of data.

In addition, the existing stored data posed a serious risk because it contained project documents, customer information, personally identifiable information (PII) and other documents that should have been disposed of after projects were completed and sites were decommissioned and abandoned.

Solution

An enterprise implementation of the IBM ILG solution gave the company insight into exactly what information was being stored, where it was stored and how much of it should be defensibly disposed. During the initial discovery, the organization concluded that about half its data hadn't been accessed for over five years and could either be discarded or moved to lower-cost storage media. It identified this old and obsolete data and scheduled it for deletion. To mitigate risk, the company also identified sensitive documents and documents containing PII, such as hazard cases that involved company employees, and marked them for defensible disposal.

Leveraging IBM ILG technology, the company established a fully automated engineering documentation creation, editing, control and monitoring process, with proper records classification according to the associated retention schedule. This helped ensure total consistency of information across departments and locations, and improved compliance and audit capabilities. With an integrated view of enterprise documents and efficient searching, staff could complete projects faster and more accurately.

By deploying the IBM ILG solution to identify relevant content in place, and by collecting only responsive data, the company met US Department of Justice compliance deadlines. The solution also improved visibility into storage usage, ensuring a proactive position for eDiscovery. Automated management of electronic records gave the company's legal department confidence that appropriate documents were placed on hold when necessary, and were not subject to disposal or alteration.

Benefits

- **All records are properly identified and classified**, with the correct corporate retention schedule applied, resulting in significant improvements in the oil and gas organization's ability to demonstrate regulatory compliance and provide a single source of truth for big data initiatives.
- **Redundant, obsolete and trivial data is defensibly disposed**, resulting in less litigation and storage costs, reduced risk and improved ability to find needed documents quickly—while also decreasing the noise presented to big data systems.
- **Legal response is streamlined**: the legal team gets insights faster, improving decision making and reducing costs.
- **Data analysts, engineers and staff have more confidence in the integrity of the documentation they use**—an important factor in the success of any engineering project.

Conclusion

ILG is an important step for oil and gas companies that want to derive higher-valued insights from big data, achieve business efficiency and ensure regulatory compliance in a changing and increasingly global world. When ILG is properly applied, information deletion can and should not only be defensible, but also economically desirable from a business standpoint. In doing so, big data becomes right-sized, *useful* data.

IBM has implemented ILG solutions across the oil and gas industry—encompassing people, process and technology—with the aim of identifying and classifying corporate records, streamlining the legal discovery process, archiving infrequently utilized data, curating high-value data for big data analytics and finally disposing of data when cost and risk outweigh value. The IBM model for improving information economics provides the framework to align and reconcile retention, privacy and security obligations globally and enforce them locally.

For more information

To learn more about IBM information lifecycle governance solutions, please visit: ibm.com/ilg

Additionally, IBM Global Financing can help you acquire the software capabilities that your business needs in the most cost-effective and strategic way possible. We'll partner with credit-qualified clients to customize a financing solution to suit your business and development goals, enable effective cash management, and improve your total cost of ownership. Fund your critical IT investment and propel your business forward with IBM Global Financing. For more information, visit: ibm.com/financing

About the authors

Elizabeth Koumpan is an IBM senior certified IT architect for IBM Industrial Section Solutions Architecture.

Scott Sumner-Moore is an executive IT architect in the Enterprise Content Management Practice for IBM Global Business Services®.

Amir Jaibaji is the Program Director for IBM Information Lifecycle Governance Product Management.



© Copyright IBM Corporation 2015

IBM Analytics
Route 100
Somers, NY 10589

Produced in the United States of America
April 2015

IBM, the IBM logo, ibm.com, and Global Business Services are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at ibm.com/legal/copytrade.shtml

Microsoft and SharePoint are trademarks of Microsoft Corporation in the United States, other countries, or both.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.

¹ The United States Department of Justice Office of Public Affairs. “Halliburton Agrees to Plead Guilty to Destruction of Evidence in Connection with Deepwater Horizon Tragedy.” July 25, 2013. www.justice.gov/opa/pr/halliburton-agrees-plead-guilty-destruction-evidence-connection-deepwater-horizon-tragedy

² Associated Press. “GUILTY: Ex-BP engineer convicted of destroying evidence after Gulf oil spill.” December 18, 2013. <http://fuelfix.com/blog/2013/12/18/ex-bp-engineer-convicted-on-1-obstruction-charge/>



Please Recycle