Schlumberger

Hess Corporation Streamlines Creation of Geographic Metadata

Case study: MetaCarta software saves significant resource time by automating geotagging

Challenge

Add geographic metadata tags to unstructured documents being migrated from shared drives into a content management system (CMS).

Solution

Use MetaCarta® geotagging capability to

- automatically discover geographic locations embedded within documents
- autopopulate document metadata with geographic locations, including information such as well names, latitude and longitude coordinates, and unique well identifiers
- simplify the CMS migration workflow.

Results

Increased the value of a CMS solution by including geographic information; automated a metadata process that would have consumed the valuable time of highly trained E&P professionals.

Migrate from file servers to CMS

The trend of accessing critical E&P information in a collaborative environment has led many oil and gas companies to consider content management systems (CMS) for data storage, management, and discovery. Hess Corporation decided to migrate its unstructured documents—previously stored on traditional file servers—into a CMS environment.

Add metadata to documents

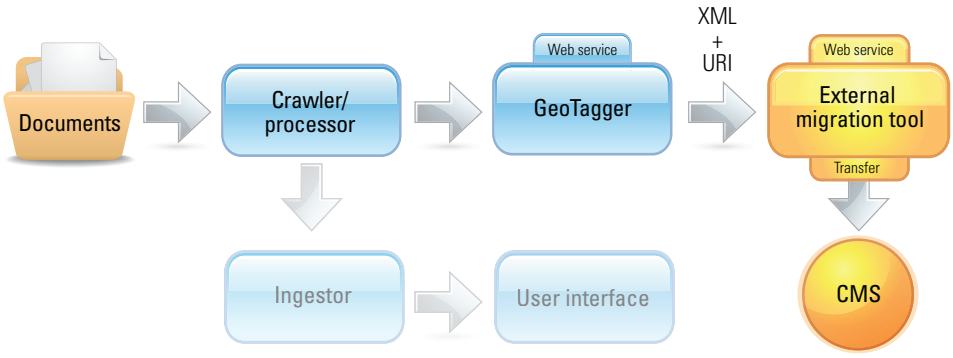
Schlumberger Information Solutions (SIS) and Hess engineers worked together to automate the creation of metadata as part of a data migration into the CMS.

Hess chose a basic data migration tool that was flexible enough to allow application programming interface (API) tie-ins to databases of metadata tags, but it also needed a way to automate the creation of geographic-focused metadata. Initially, project leaders at Hess planned to have experienced professionals add this metadata document-by-document. However, this manual method would be too time-consuming for the initial sample set of about 5,000 documents.

Use MetaCarta for geographic search and referencing

Hess was an early adopter of MetaCarta for oil and gas†. The company primarily used the MetaCarta GTS geOdrive™ portion of this geographic text search and referencing platform (GSRP) for its exploration business groups. GSRP is comprised of several different components, including

- crawlers that scan various repositories for documents
- processors that convert those documents into plain text



Data migration process showing the role of MetaCarta GeoTagger software.

Case study: MetaCarta software saves significant resource time by automating geotagging

- ingesters that seek geographic terms within the natural language of the document
- a default mapping user interface with location APIs that can be integrated in a wide range of mapping systems.

Another component of the GSRP, named GeoTagger, enables place names to be assigned to documents within the new CMS. When fed raw text, GeoTagger returns marked-up XML that identifies geographic locations found in the text, as well as other pertinent location-based information.

Industry-specific locations

A key differentiator of MetaCarta's GeoTagger versus dictionary-based gazetteers of place names is the patented Natural Language Processing (NLP) engine. Additionally, millions of industry-related locations are preprogrammed in the Oil and Gas Geographic Data Module (GDM), including basins, fields, well names, and even UWIs and location APIs—information that would require experienced E&P resources to insert manually. For Hess-specific proprietary locations, the advanced gazetteer functionality built into MetaCarta allowed further customization. By combining the NLP and Oil and Gas GDM capabilities, Hess could be confident that place names extracted would be both accurate and relevant.

Migration dataflow

MetaCarta's plug-in architecture was used to modify the direction of data flow. First, the system's connector framework crawled the 5,000 documents and decoded them into pure text. Once this text was created, it was redirected into the GeoTagger framework rather than the traditional flow of passing through the Geosearch engine. The GeoTagger XML was then exported via Web services and stored in a database.

Then, an API call from the migration tool with a specific query extracted the unique document identifier (UDI) and required data fields as metadata that accompanied the document. Finally, the migration tool swiftly moved each document from the file server repository, embedded the geographic metadata fields without manual intervention, and kicked off the next step of the migration workflow.

Gained higher-quality metadata

Incorporating the MetaCarta GeoTagger tool into the migration process enabled more efficient automation, thus enabling less user input and higher-quality metadata on unstructured content. Based on the time savings and cost-effectiveness of the E&P document migration project, Hess plans to use the CMS solution for other groups within the organization.

Schlumberger Information Solutions

Schlumberger Information Solutions (SIS) is an operating unit of Schlumberger that provides software, information management, IT, and related services. SIS collaborates closely with oil and gas companies to solve today's tough reservoir challenges with an open business approach and comprehensive solution deployment. Through our technologies and services, oil and gas companies empower their people to improve business performance by reducing exploration and development risk and optimizing operational efficiencies.

"MetaCarta technology helped us embed geographic information associated with our documents in a seamless manner, requiring little to no human intervention."

Stephanie Byrd
IM Project Specialist
Hess Corporation

E-mail sisinfo@slb.com or contact your local Schlumberger representative to learn more.

www.slb.com/sis

[1]Schlumberger has an exclusive sales and support agreement with MetaCarta, Inc. for oil and gas.

Schlumberger