Leveraging Energy-specific GenAI to Transform Decision-Making

Executive Summary

The modern energy industry does not suffer from a scarcity of subsurface data; rather, it is constricted by a severe consumption bottleneck. Geoscientists spend a disproportionate amount of operational hours wrangling data files and scripting workflows, rather than performing high-value geological interpretation, structural mapping, and risk mitigation.

This white paper analyzes the core geospatial and geological integration challenges facing geoscientists across the entire asset lifecycle—from regional exploration, lease evaluation, and target generation, to well planning, production surveillance, and infrastructure optimization. By examining real-world enterprise deployments, we demonstrate how shifting from rigid dashboards to interactive, natural-language data dialogue, and why that shift, not more data or more tools, may be the real unlock for the industry.

The Problem: The Subsurface Lifecycle Data Tsunami

For decades, the primary objective of exploration workflows was the acquisition of high-fidelity data. Today, explorationists, structural geologists, and reservoir engineers are routinely inundated with diverse datasets that arrive faster than any one person can organize, cross-reference, and interrogate.

The technical bottleneck has fundamentally shifted from data acquisition to interpretation velocity. The distance between an urgent geological question—such as defining an underexplored fault block or tracking a localized sweet spot—and a defensible, data-backed drill location remains unacceptably wide. Every hour an explorationist spends executing manual data transfers, formatting text files, or waiting on localized data loaders to build static map views is an hour lost to economic exposure and could result in poor choices in operations, with problems during drilling, completions, etc.

The modern geoscientist faces challenges at each phase of their workflow.

Exploration and Appraisal: Evaluating Leases, Mapping, and Target Identification

The exploration phase requires geologists and geophysicists to synthesize vast regional spatial trends to identify target formations and rank lease blocks. This demands the rapid generation of structure maps to define depth constraints, isopach maps to map true stratigraphic thickness, and isochore maps to establish vertical intervals.

Furthermore, identifying commercial opportunities relies on geoscientists isolating and blending seismic attributes, production histories, and petrophysical log suites. Traditional software architectures keep these views segregated: A geologist must inspect a regional structure map in one mapping package, pick key faults and horizons in a seismic interpretation application, and run standalone log cross-plots in a separate petrophysical package.

This fractured workflow disrupts the interpreter's train of thought, introduces format-conversion risks, and discourages the iterative, multi-variable querying required to uncover subtle indications of economic petroleum systems.

Well Planning, Drilling, and Completions

Once a target location is identified, geoscientists must construct accurate cross-sections and real-time structural maps to guide the wellbore through the sweet spot.

Critical geological context is frequently trapped in unstructured formats, such as daily drilling reports or geological mud logs buried in PDFs. Manually mining these documents to cross-reference reported drilling metrics against real-time mechanical torque or lithological changes introduces massive operational delays when immediate geosteering or completion adjustments are required.

Production, Workovers, and Infrastructure Optimization

During the active development phase, geoscientists may need to combine time-series production overlays, structural fault maps, and historical drilling and workover logs to diagnose an issue. Because these diagnostic datasets live in isolated relational databases separate from the spatial infrastructure charts, establishing an integrated asset view demands extensive manual schema mapping.

The Data Wrangler's Tax

Because traditional analytics interfaces rely on rigid pre-built dashboards or complex, programmatic query languages (such as SQL, Python, or R), a strict technical barrier exists. Geoscientists face a compounding 'data tax':

- **The Programming Gap:** An experienced structural geologist understands the physical meaning of a fault network or play fairway analysis, but he/she may lack the programming familiarity to write a custom script to automate data cleaning across thousands of digital well logs.

- **The Dashboard Dead-End:** Traditional static dashboards are built for predictable, recurring questions. When an anomalous event occurs—such as a sudden change in reservoir pressure—the pre-built dashboard cannot adapt to investigate the anomaly. The interpreter is forced to request a new view with additional data from an IT or data engineering team, introducing delays that stall immediate remediation efforts.

Closing the Gap: Conversational Analytics and Agentic Orchestration

Instead of forcing geoscientists to adapt to the rigid syntax of software development, the software environment must adapt to human geological reasoning through an intelligent, domain-aware conversational orchestration layer.

An effective solution framework must possess **five foundational characteristics:**

1 Direct Natural Language Interpretation

The platform must allow geoscientists to interact with multi-variate subsurface data using standard domain vernacular. An interpreter should be able to give a complex instruction—such as "plot porosity against resistivity for the Woodford penetrations, colored by operator"—and have the analytics platform immediately compile, validate, and render the appropriate cross-plot or spatial display without manual chart-building.

2 Multi-Step Agentic Orchestration

Subsurface mapping workflows are rarely solved by a single chart. The solution must function as an orchestration engine capable of coordinating multi-step analytical pipelines. When tasked with a multifaceted problem—such as evaluating and ranking underdeveloped landing zones across mature oilfields—the system must interact with external data functions, trigger specialized machine learning models, and immediately return interactive, spatial visualizations within a single, unified workspace.

3 The "Apprenticeship Effect" and Reusable Assets

A critical flaw of standalone chatbots is that their outputs are ephemeral and opaque. To provide true enterprise value, the conversational layer must generate transparent, governed, and reusable analytical assets.

- When a user requests a complex geological calculation, the underlying code (e.g., Python or R data functions) must be explicitly visible, editable, and saved to a central repository.
- This creates an "apprenticeship effect." As geoscientists inspect the scripts generated by the AI in response to their explicit data questions, they organically absorb domain-specific programming patterns, accelerating their personal professional development.

4 Model-Agnostic, Governed Architecture

Because energy data includes highly sensitive, Proprietary Areas of Interest (AOIs), fracture models, and reserves estimates, data privacy is paramount. The conversational architecture must be entirely decoupled from public large language models (LLMs). Organizations must be capable of deploying a private orchestration layer within their own secure cloud or on-premises infrastructure, routing queries strictly through enterprise-vetted, secure LLM instances where data never leaves corporate boundaries.

Beyond security, organizations need model-agnostic tools that can work with AI platforms such as: OpenAI, AzureOpenAI, Anthropic, Google, AWS, and/or Cohere. They also should be able to switch between models when needs or strategy shifts.

5 The Ability to Scan Visual Data

Geoscience data often spans charts, graphs, seismic, and other visual formats. An effective AI-driven analytics tool will incorporate the ability to analyze qualitative data equally as effectively as quantitative data.

Software that combines these five characteristics, such as Spotfire, into an AI solution dismantles data silos, transforms data into usable insights, drives operational agility, and delivers data-driven customer value.

Spotfire is a visual industrial analytics platform designed to hold these capabilities in a single environment, uniting interactive geospatial mapping, streaming and batch data ingestion, and embedded Python and R analytics against live enterprise sources rather than static extracts.

Spotfire Copilot™ is the conversational layer built on that foundation. A geoscientist can pose a question in plain language and receive an interactive visualization, a data function, or a multi-step analytical workflow in return—together with the underlying code, which remains visible, editable, and available for reuse by colleagues. Because Copilot is model-agnostic and can be deployed inside an organization's own secure cloud or on-premises infrastructure, proprietary AOIs, fracture models, and reserves estimates never leave corporate boundaries. The deployments below illustrate what this architecture delivers in practice.

Case Studies and Results

For example, oil and gas operator Wintershall Noordzee aimed to eliminate the friction that historically separated exploration teams from data engineers. What started out as a couple of data tables in Spotfire evolved into a fully functional, AI-driven data platform called "Crossfire."

The platform integrated and analyzed Wintershall Nordzee's massive amounts of data from its North Sea area operations, revealing new insights and streamlining future decision-making. It shifted the time needed for data analysis from eight hours to roughly real time—a huge savings and win for the company.

"Before Spotfire analytics, when a user had a question, I would have to plot the data on a map after using five different tools to get the desired information from our data. Now, they can do it themselves, and they're just one click away from considering other variables (such as the rock unit, or borehole geometry), which previously would have cost me hours of extra work. Crossfire is a huge time saver with a lot of transparency and new insights," said Dejan Zamurović, data engineer at Wintershall Nordzee.

Similarly, oilfield services provider Liberty Energy constructed Atlas, a real-time operational dashboard built on the Spotfire visual industrial analytics platform. Atlas ingests telemetry from frac equipment, including pumps, sensors, and blenders, at one-second intervals. This data is streamed via a Kafka event broker, processed and standardized, and displayed in Spotfire data applications that update continuously. Streaming, Automation Services, and Python Services are all deployed in a containerized Kubernetes environment to ensure scalability and high availability.

Atlas facilitated Liberty's analysis of more than 50 billion data points in 2024 and 2025, quickening innovation and increasing real-time operational awareness and efficiency. Now, Liberty has also added Spotfire Copilot™, a large-language-model, conversational AI assistant to Atlas, to further simplify its analysis and cross-team adoption.

At the enterprise level, an AI-driven tool with a conversational layer, such as Spotfire Copilot™, can transform data analysis across departments, reducing uncertainty company-wide. As the largest independent pure-play oil and gas operator in the Permian Basin, Diamondback Energy has more than 1,300 Spotfire users deploying the technology across facility planning, operational assignments, financial cost footprint analyses, land data browsing, production surveillance, and cross-system well comparison.

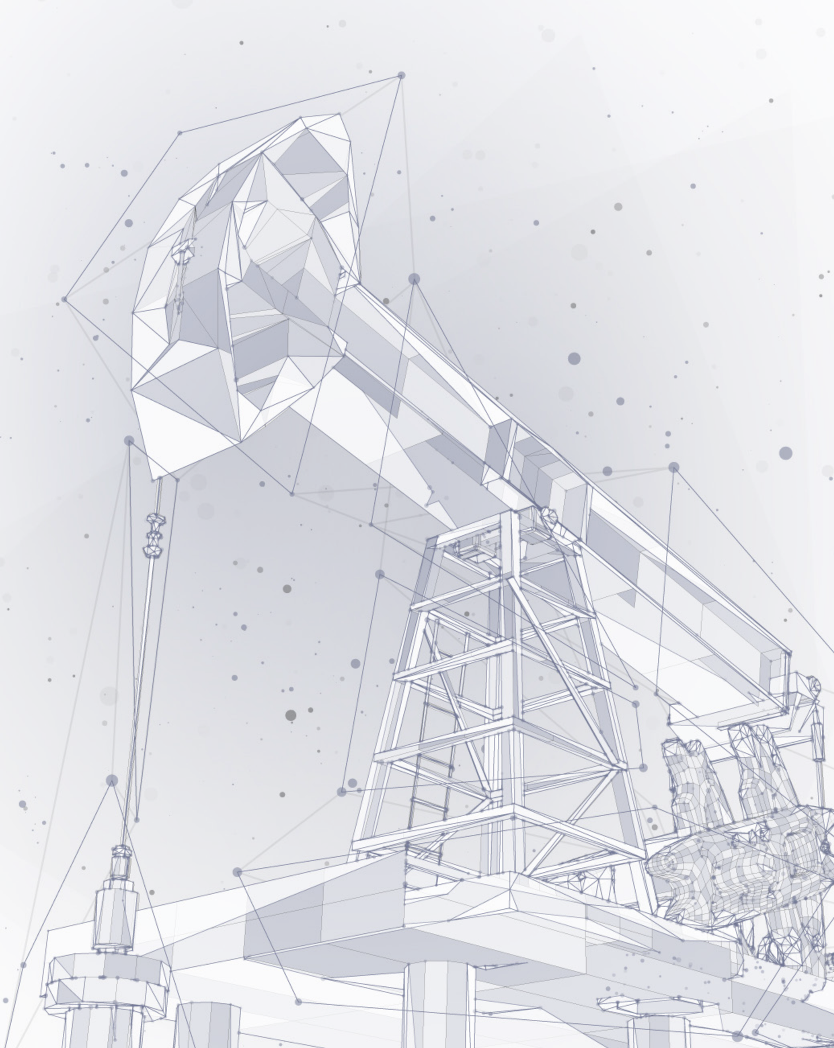
Finally, for B2B information services companies, these tools create immediate added customer value. Analytics and information services giant S&P Global needed to provide data-driven answers for every stage of the oil and gas exploration and drilling workflow, so it integrated Spotfire with its own data analytics tools to automate processes and reduce analysis time.

"With Spotfire, the user can start the computation running and go to lunch and come back and have everything ready. The solution has reduced the time needed by about 80 percent," says Camilo Rodriguez, associate director of project management at S&P Global. Conversational AI like Spotfire Copilot™ can compound these gains by turning multi-day script-writing tasks into seconds-long queries.

Conclusion

The proliferation of complex, heterogeneous data confronting today's energy geoscientists and engineers is not a problem to be solved by better filing systems or more disciplined curation alone. It is, instead, an invitation to rethink how professionals interact with their data in the first place — moving from a model built around selecting the "right" data set to one built around workflows sophisticated enough to make productive use of the full, messy inventory of information a modern energy company holds.

Conversational AI layers, such as Spotfire Copilot™, mark a genuine inflection point in that shift. By combining decades of geospatial and visual analytics expertise with model-agnostic generative AI, agentic multi-step workflows, and built-in guardrails, Spotfire has moved beyond simple query-and-visualize functionality toward something closer to a collaborative analytical partner.



To learn more about how Spotfire can streamline data analytics and efficiency at your organization

spotfire.com/industry/Energy