The Real Cost of “Free” Data

Why Your Data Approach Is Quietly Undermining Your Biggest Energy Initiatives

By Pete Veenstra, *Director of Data Governance at Terracon*



“You always get what you pay for.”

It sounds simple. But when it comes to geospatial and environmental data, most energy organizations haven’t done the math. The numbers are more alarming than most executives realize.

The Illusion of “Free”

In the era of information overload, regulatory uncertainty, and increasingly inconsistent access to government data, the assumption that quality project data is free is widespread in the energy sector. Government portals, federal agencies, and open-source mapping platforms make it easy to believe your teams can pull whatever they need, whenever they need it, at no cost or for a nominal fee.

That belief often costs more than a premium data subscription ever would.

Hidden Costs

Free data carries real hidden costs. It is like an iceberg: what is visible above the water represents only a small part of the story, while the far larger and more consequential risks, complexities, and liabilities remain hidden beneath the surface. Every hour your analysts spend finding, downloading, normalizing, verifying, and re-pulling data is an hour they are not doing the high-value work you hired them to do. When that data is wrong, outdated, or incomplete, the downstream consequences — such as missed permitting windows, derailed site selections, costly rework, and project delays — outweigh whatever was saved by avoiding a subscription fee.



The Stakes Have Never Been Higher

The U.S. energy sector is in the middle of a once-in-a-generation buildout. Wind, solar, battery storage, electric transmission, pipelines, and data centers are all competing for sites, permits, and talent at the same time. The companies that move fastest with the most reliable information have a clear advantage.

Consider what is working against speed right now:

- Federal permitting delays held up **11 gigawatts** of new renewable energy deployment in the U.S. in the last year.[1]
- A single solar project can be stalled **up to 10 months** by a late-stage biological survey; a storage project can lose **six months** to an Endangered Species Act consultation.[1]
- The average clean energy project takes **4.5 years** just to obtain NEPA reviews; for transmission, it stretches to **6.5 years.**[2]
- About one-third of wind and solar projects were canceled in the last five years, and **half experienced delays of six months or more**.[3]
- Project cancellations resulted in average non-recoverable sunk costs of more than **$2 million per solar project** and **$7.5 million per wind project.**[3]

In this environment, every week lost to bad data is a week closer to a permit denial, a missed interconnect queue position, or a competitor locking up the best site. The cost of information failure is increasingly avoidable. The organizations pulling ahead in today's energy buildout share one thing: they've made managing data a strategic priority.

---

[1] Crux, "The Impact of Federal Permitting on Clean Energy Development" (2025). Survey of 50 renewable energy developers. Reported by E&E News/Politico (April 2025) and PV Tech (April 2025).

[2] American Clean Power Association (ACP), "U.S. Permitting Delays Hold Back Economy, Cost Jobs" fact sheet.

[3] Lawrence Berkeley National Laboratory, developer survey on wind and solar project outcomes (published/reported August 2024)

What “Free” Actually Costs Your Team

Most executive conversations about data come up short. The conversation stays at the surface: “Our GIS team can pull that data.” What gets left out is what that actually means in practice.

When you compile a database for use, you’re going to have to take the time to research the data, acquire it, standardize it, and come up with some way to maintain it — all while constantly asking yourself about the quality.

- Is it always available?
- Is it the most up-to-date?
- Can the agency be trusted?
- Is it documented?

That is not a one-time exercise. It is a continuous operational burden because each time data is published, you need to discover it, research it again to understand the deltas between the recent and previous publication, and the whole process repeats.



The Hidden Cost Stack

When your team relies on free or public data to evaluate sites, assess environmental constraints, and build the early-stage picture to move an initiative forward, there are costs worth understanding:

1. **Research** — Identifying which federal agencies, state offices, local governments, open-source platforms, and private vendors hold the data you need, tracking down where it actually lives, and asking the harder questions: Can this source be trusted? Is the data collection and creation process documented? And what is the provenance — where did this data actually come from before the agency published it?
2. **Acquisition** — Downloading data from disparate sources, formats, and portals (a process that can take hours to days per dataset).
3. **Normalization** — Standardizing data from shapefiles, KMZ, GeoJSON, and dozens of other formats into a consistent, usable structure is only the beginning. When multiple sources cover the same topic, like federal wetlands data, state wetlands data, and a third-party layer all describing the same geography, your team must decide which source to trust, how to reconcile conflicts, and how to document that decision in a way that holds up to regulatory scrutiny.
4. **Quality Assurance** — Constantly checking for geometry errors, missing attributes, incorrect projections, and incomplete and stale records.
5. **Hosting and Security** — Storing, securing, and making data accessible to distributed teams.
6. **Ongoing Maintenance** — The responsibility for accuracy doesn't end once data is loaded. Skilled staff must proactively check sources on a scheduled basis, since agencies rarely notify you when data has been updated, revised, or moved. Moved is the operative word: a dataset that lived at one URL last quarter may have been reorganized, renamed, or migrated to an entirely new portal. Your team knows the data exists somewhere, but finding it again takes real time and expertise. Every update cycle means re-locating, re-pulling, re-normalizing, and re-validating, a revolving door that never stops.
7. **Redistribution** — Pushing updated data to remote stakeholders who may be working from outdated versions.
8. **Documentation** — Proving data provenance for regulatory, legal, and permitting defensibility.

The moment a dataset is downloaded, it begins going stale.

Unlike a one-time report, an energy initiative is a living process. Data pulled at the screening phase is already aging by the time permitting begins, and stale again before construction starts. Your team has to start over at every milestone. A government shutdown or budget reduction can pull datasets offline without notice. But the quieter risk is that even available data may not be current, since many public sources are maintained inconsistently, updated infrequently, and carry no timestamp to tell your team how old the information actually is. Finding a data team that has experience with federal datasets going offline during government shutdowns is why data redundancy and on-premise storage matter so much.

The Compounding Risk: When Data Quality Falls Short

Poor data quality does not just slow things down. It tends to compound across the project lifecycle.

Gartner research finds that poor data quality costs organizations an average of **$12.9 million annually**[4]. Forrester's 2023 Data Culture and Literacy Survey found that over a quarter of organizations estimate they lose **more than $5 million per year** from poor data quality, and 7% report losses of **$25 million or more**[5]. For utility and energy companies, where a single bad site selection decision or missed permit condition can cost tens of millions, the stakes are higher still.

Bad construction data, defined as data that is incomplete, inaccurate, inconsistent, or untimely, was estimated to have caused **$1.8 trillion in losses worldwide** in 2020, responsible for 14% of avoidable rework totaling $88 billion in costs, according to an Autodesk and FMI Consulting study.[6]

The **1-10-100 rule** of data quality captures why early investment pays off. Fixing a data error at the point of entry costs **1x**. Fixing it once it has propagated through the system costs **10x**. Addressing the consequences at the decision-making stage costs **100x**. For energy infrastructure, the decision stage means you have already committed capital to a site, a route, or a corridor, and only then discover the constraint that blocks the permit.

---

[4] Gartner, "How to Improve Data Quality" (2021)

[5] Forrester "Millions Lost in 2023 Due to Poor Data Quality (2024)

[6] Autodesk and FMI Corporation, "Harnessing the Data Advantage in Construction" (2021)

Real Scenarios Where Bad Data Becomes Real Loss

Scenario	What Goes Wrong	Consequence
Stale wetlands data	Team selects a site believing it to be clear; field survey reveals protected wetlands	Permit denial, $2M+ in sunk costs
Outdated endangered species layer	Route approved based on old habitat data; new species designation triggers Section 7 consultation	6 to 10 month delay, cost escalation
Incorrect parcel ownership data	Outreach sent to wrong landowners; negotiations stalled	Weeks of schedule loss, competitive exposure
Missing regulatory zone boundaries	Environmental impact assessment incomplete at submission	Permit rejection, full re-submission required
Stale infrastructure data	Transmission route planned near undocumented pipeline	Redesign mid-project, rework costs

The Analyst Time Tax

Every major survey of data professionals over the past decade has reached the same conclusion: the majority of their time goes to finding, cleaning, and organizing data rather than analyzing it. Anaconda's 2022 State of Data Science survey found that data preparation and cleansing alone accounted for 39% of a data professional's time, and that figure likely understates the reality for GIS and project teams. Research on data munging in AI and analytics workflows consistently puts the time spent collecting, cleaning, and preparing data at closer to 70% of total work. For GIS and project teams pulling from dozens of fragmented public sources, that overhead is compounded further, and it represents real budget spent on work that adds no value to your project.

Example cost scenario: For an energy company with a team of five GIS analysts at a loaded cost of $100,000 or more each, you are paying north of $500,000 per year and getting less than half of it in productive analytical output. The rest is overhead that a well-curated data platform eliminates.



The Project vs. Product Mentality

One of the most important framing shifts for executive decision-makers is moving from a *project mentality* to a *product mentality* around data.

With a project mentality, data is treated as a one-time task: pull what you need, use it, move on. This works when a project needs only two or three datasets. But modern energy projects routinely require 50, 100, or more datasets. At that scale, the project mentality breaks down fast.

A product mentality means your data is always available, always current, always documented, and always accessible to the right people, regardless of which analyst downloaded it or where they stored the file. That is what institutional-grade data infrastructure looks like.



A Framework for Executive Decision-Making

If you are evaluating your organization's approach to project data, consider these four questions:

1. What is the true loaded cost of your current data process?
Add up analyst salaries, benefits, management overhead, and hosting infrastructure. Then apply the 60 to 90% rule: that is the portion of that budget going to data collection rather than analysis. Compare that figure to the cost of a data or platform subscription.

2. What did your last project delay actually cost and what caused it?
If a permit was denied, a site was abandoned, or a corridor had to be rerouted after substantial investment, did you trace back the root cause? In many cases, there may be a data quality or timeliness issue at the origin.

3. Can your team prove the origin of your current data for every layer in your last permitting submission?
Regulators increasingly require documentation of data sources, acquisition dates, and transformation methods. If your team cannot produce that documentation reliably, you carry regulatory risk that a curated platform resolves by design.

4. Are your analysts doing analyst work or data work?
The highest-value output from your technical staff is interpretation, risk analysis, stakeholder communication, and decision support. If they are spending most of their time finding and formatting data, the organizational leverage of a data platform far exceeds its cost.

The Value of Getting It Right: Data and Platform You Can Trust at Every Stage

Pivvot Data: A curated geospatial foundation

Pivvot aggregates, curates, and continuously maintains over 450 million records of data from approximately 900 sources spanning land, environmental, regulatory, and infrastructure data across the mainland United States. Pivvot makes this data accessible through a subscription for your own GIS platform or Pivvot's cloud-based platform.

Every dataset is sourced from trusted agencies and private providers, put through rigorous QA, and delivered with full metadata and provenance so you know what you are looking at, where it came from, and how current it is.

Pivvot Platform: Turning data into decisions

On top of that data foundation sits the Pivvot platform, which supports the full lifecycle of energy initiatives: from early-stage site screening and suitability analysis, to interconnect routing and corridor evaluation, to detailed environmental reporting for permitting submissions. Your teams are not choosing between speed and accuracy; they get both, in a single environment.

For energy and infrastructure projects, Pivvot has relevant data layers including:

- **Environmental and ecological:** NWI wetlands, critical species habitat, floodplains, protected waterbodies
- **Land and parcel:** property parcels, land ownership conflicts, land cover, federal and tribal lands
- **Regulatory and jurisdictional:** EPA regions, FEMA zones, BLM administrative units, U.S. Army Corps districts
- **Infrastructure:** electric transmission lines and substations, existing pipelines, fiber optic, wind turbines, data centers
- **Geotechnical:** depth to bedrock, hydric soils, slope, landslide risk, karst topography
- **Societal:** environmental justice indices, social vulnerability, cultural resources

You can license this data directly into your own GIS and engineering platforms or access it on demand through Pivvot's cloud services. In either case, your team starts from a standardized, map-ready data foundation instead of rebuilding it project by project.

Because every layer is curated, normalized, and delivered map-ready with current metadata, your staff is freed from the work of sourcing, cleaning, and maintaining data and can focus on interpreting it, comparing options, and moving initiatives forward.

Revisiting the Project vs. Product Mentality

Pivvot solves this project vs. product problem because Pivvot's data process reflects these disciplines:

- Every federal dataset is verified **every 30 days.**
- Every state dataset is verified **every 90 days.**
- Full metadata (source agency, acquisition date, transformation log) is documented for every layer.
- The last three data loads are preserved for instant rollback, so if a QA issue surfaces after publication, or a source agency fundamentally changes its data, Pivvot can restore the previous clean state immediately without clients missing a beat. A separate three-year deep archive provides an additional layer of protection against data going permanently offline.
- Every dataset is delivered with formally governed metadata that provides full documentation of provenance, lineage, and curation pedigree. Your team can take this documentation to demonstrate that every data layer you relied on has a complete, auditable chain of custody from original source to final delivery.

Now, the question isn't whether your organization could build this infrastructure; **it's whether that's the best use of your data team's time, budget, and expertise.**



The Business Case for Getting the Right Data

Building the justification for premium, curated data comes down to four ROI levers:

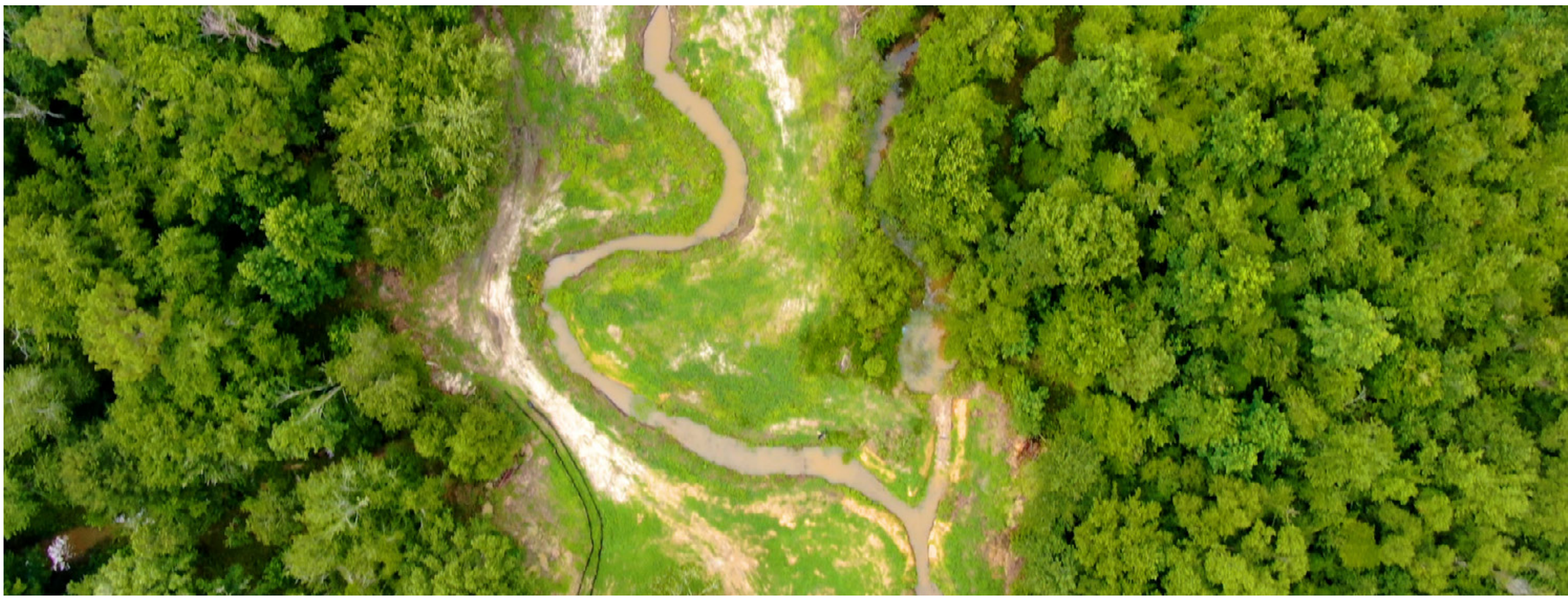
1. Schedule Compression

Pivvot clients consistently report a significant reduction in project planning time. For a wind or solar project that normally takes four to six years from announcement to commercial operations, cutting the planning phase in half translates directly into earlier revenue recognition, earlier PPA activation, and earlier return on capital.

For a transmission project, an EPC used Pivvot's data to reroute an existing corridor, generate multiple route options, and select a defensible least-cost, minimal-impact path in hours instead of weeks — turning what used to be a manual data-gathering exercise into a fast, data-rich decision.

2. Earlier Risk Identification

Identifying a serious flaw in a site or route during the desktop phase costs a fraction of what it costs during field investigation or permitting. Pivvot's data surfaces environmental constraints, protected species habitat, regulatory zones, and land ownership conflicts before your team commits resources to a site. Just as importantly, it surfaces the absence of data, flagging where information gaps exist so your team knows when to dig deeper before making a commitment, not after. Sending crews to a bad site is a fully avoidable cost.



3. More Accurate Permitting

Permit applications built on accurate, current, and documented data are more defensible and less likely to trigger requests for additional information or outright rejections. Engineering and construction firms using Pivvot data report more accurate permitting estimates, better cost projections, and faster RFP responses because the underlying data is reliable.

As Max Hengst, VP of Pipeline Engineering at Audubon Engineering, stated: “Pivvot’s technology and its comprehensive approach to big data analytics reduces our design time and enhances project delivery, allowing us to provide transformational value and quality to our customers.”

4. Analyst Productivity and Retention

When your best analysts are hunting and gathering data instead of analyzing it, you are underutilizing one of the most expensive parts of your workforce. Freeing GIS analysts and project engineers from data wrangling and directing their time toward interpretation, stakeholder engagement, and decision support improves throughput without adding headcount. In a talent market where experienced professionals are increasingly scarce, that productivity gain can double as a retention tool.



The Bottom Line

“Free” data is rarely free. It tends to show up as a recurring operational cost that compounds quietly across projects and permitting cycles.

Competitive advantage increasingly belongs to the organizations that can evaluate sites in hours instead of weeks, generate defensible permit applications on the first submission, and redirect their most valuable people toward work that moves your business forward.

Pivvot was designed and built specifically for this moment — to give energy companies the data infrastructure they need to make faster, more accurate, and more defensible decisions at every stage of the project lifecycle. Pivvot data and Pivvot platform slash project planning time, eliminate the hidden cost stack of free data management, and ensure your teams are always working with information they can trust.

The question is no longer whether your organization can afford premium data; but can you afford not to have it.

Learn more about Pivvot

Ready To See What Better Data Means for Your Next Project?

Visit **pivvot.com/data** or contact our team to schedule a discussion about Pivvot's data and platform, and see which is right for your projects and workflows.



pivvot.com
816.859.5131
info@pivvot.com