

WHITE PAPER

# Preparing for the *Intelligent Web*

The New Digital Frontier

# The new frontier is here

For years, the formula for online visibility was clear: create impactful content, optimize it for keywords, build backlinks, and earn the click. Today, that path is being disrupted.

Now, AI has become the primary interface users turn to for research, drastically changing how information is surfaced in a short span of time. Gartner has predicted that traditional [search engine volume will drop 25% in 2026](#) as search marketing loses ground to Large Language Models (LLMs), chatbots, Google's AI Overviews (AIOs), and other AI-powered tools. Instead of perusing a list of links from a search engine, users are opting for direct, synthesized answers.

As the very foundations of online discoverability continue to change, a question emerges: How do you ensure your most recent,

relevant, authoritative data is accurately scraped and surfaced by LLMs and AIOs? The answer lies in the transition to the **intelligent web**. Brands have to rethink what job their websites need to do.

A digital property must act as both a conversion destination and a source of information, serving both human visitors seeking connection with a brand and agentic visitors that require clear, structured data.

Static content management is no longer enough. Architecture that supports personalized, interactive engagement for humans while providing the structural clarity AI systems need to interpret and trust your data is what will ensure your brand remains visible, citable, and authoritative on the intelligent web.

“

*The intelligent web is fundamentally rewriting the rules of the digital economy. While it's creating massive opportunities, it's also introducing profound complexity, ambiguity, and some anxiety for users, enterprises, and the workforce at large.*

*For WP Engine, these challenges represent a definitive shift in how we must equip our customers. Our mission is to liberate digital content owners from technical complexity. We do so by creating tools that can help them unlock the full economic potential of their content and provide delightful experiences for both humans and agents.”*

*In essence, experiences on the intelligent web must be simultaneously human-first and AI-native.*



**Thierry Muller**

VP of AI Products, WP Engine

# Welcome to the intelligent web

**The intelligent web** is a totally new frontier in which AI, through LLMs and features like AIOs, has disrupted both the traditional path users take from research to solution selection and the strategies website owners must use to maintain visibility.

## The birth of SEO

Traditional SEO was born from a mix of necessity and frustration. While digital curators began tinkering with terms to manipulate rankings in the early 1990s, the first recorded use of the term “search engine optimization” seems to have come from marketing materials for Webstep Marketing Agency in 1997.

[Internet lore also credits Bob Heyman](#), a music lawyer and marketing executive, for starting the practice. Legend has it, a manager for the rock band Jefferson Starship called him, furious that the band’s official site was buried on page two of search results behind fan pages.

The industry shifted from simple “keyword stuffing” to a more sophisticated discipline with the [launch of Google and its PageRank algorithm](#) in 1998. It was the first search engine algorithm to take inbound links into account, setting content authority as the new gold standard for visibility.

For over two decades, SEO was a game of matching specific terms to indexed pages and gaining credibility through link building. It was an era defined by a short, fragmented search string (think “top restaurants Austin Texas” or “best hosting provider”) that led users on a linear journey from their initial query to a list of links containing potentially relevant information, which they would peruse until finding their ideal destination.

## The disruption: LLMs and zero-click enter the chat

In the last few years, the traditional search model has been disrupted by a new way of interacting with the internet. Users are entering longer queries into new interfaces and getting direct answers from familiar search engines without ever clicking a link. [Data from 2024 by SparkToro](#) reported that nearly 60% of all searches in the United States ended without a click.

As AI, including LLMs and AIOs, becomes a primary layer of discovery, the traditional “click” is becoming a premium commodity. [A 2025 study by Amsive](#) found that click-through rates (CTR) drop by more than 15% when an AIO is present, while more recent [research from Ahrefs](#) in 2026 puts that number at a staggering 58%.

Research from 2026 shows  
click-through rates *drop 58%*  
when AI Overviews appear.

Now pair that loss of CTR with the fact that the sheer number of searches surfacing AIOs is ballooning. [Conductor](#) found that the total number of keywords generating AIOs grew 88% between April and September 2025. As users find the gist of the information they need without ever leaving the search engine results page, clicks to your site are harder to come by.

This large-scale digital disruption will likely persist as bots continue to gather information for use by the tools powering the next generation of search. Whether it’s an answer in ChatGPT or an AI response in a search engine, the traditional funnel is being compressed. This raises the stakes for what your website must deliver in this new frontier.

# Building an intelligent web experience

It's tempting to view this widespread traffic disruption as a sign that websites are facing obsolescence, but AI is not replacing websites; it's intensifying what they must deliver. Modern sites must provide the information structure and authority for AI systems to cite your content, as well as the engaging, personalized experiences human users still expect.

## The agentic audience

In the intelligent web, the clarity of your site's structure is what ensures you are cited by the AI, and the quality of your experience is what ensures the user stays once they arrive. The agentic audience, including crawlers and LLM bots, values data efficiency, accuracy, and structured semantics.

For bots, context is everything. The average [ChatGPT prompt is 60% longer than a typical Google search](#). To best serve their users, bots must go beyond serving a blue list of potential destinations; they must serve personalized, definitive answers based on user queries.

Longer queries indicate users are looking for *answers*, not options.

To serve this audience, site owners must prioritize machine readability. This involves:

- **Structured data:** Implement [Schema.org](#) markup and tools like [Advanced Custom Fields](#) (ACF) so bots can instantly identify the “who, what, and where” of your content.
- **Semantic consistency:** Ensure your site's internal linking and entity relationships are clear. Tools like [Google Search Console](#) and [Schema Markup Validator](#) are essential for verifying that your agentic experience is technically sound.

## The human audience

The human audience, on the other hand, values personalized experiences, emotion, and brand connection. Research by 6sense shows [users have already completed two-thirds of the buyer journey before ever engaging with a seller](#). These high-intent users have already conducted preliminary research using AI and narrowed down their top vendor choices by the time they arrive on your site. At that point, they're merely looking for the trust signals that will help them make a final purchase decision.

Because McKinsey projects that by 2028, [\\$750 billion in US revenue will funnel through AI-powered search](#), showing up in generated answers is the best way to ensure your brand stays in front of your audience as the funnel changes. Once they get to your site, there are a few ways you can optimize the experience you're providing:



**Personalization tools:** Tools like [Smart Search AI with AI-Powered Recommendations](#), [Optimizely](#), or [Adobe Target](#) help you tailor content based on a user's previous interactions. Visitors who come to your site and use your on-site search tools are [two to three times more likely to convert](#) once they find what they're looking for.



**Core Web Vitals:** Site speed and performance are still vital to user retention. If a site takes too long to load or loads improperly, users will see it as less trustworthy and find another provider. Tools like [NitroPack](#) are built to cut load times and improve [Core Web Vitals](#) quickly, addressing speed issues without technical complexity.

## Basic principles of intelligent websites

As individuals building digital experiences think about future-proofing for the intelligent web, there are three guiding principles that can make all the difference:



**Audience awareness:** Sites must be able to immediately recognize the end user, whether human or agentic.



**Intelligence balance:** Site content must be designed to equally prioritize human understanding and interactions as well as artificial reasoning and agentic interactions.



**Modular adaptivity:** Sites should adopt a modular content structure to enable seamless content dissemination and interaction based on context.

# Elevating from SEO to GEO

As this new frontier continues to change, so does the vocabulary of visibility. Whether you call it Generative Engine Optimization (GEO), AI/Answer Engine Optimization (AEO), or Generative Search Optimization (GSO), the objective is the same: to ensure your brand's information both shows up and remains accurate when mentioned by LLMs and AIOs. While traditional SEO tactics and keywords laid the foundation, visibility is no longer just a matter of keyword density.

Your information surfaces through two primary channels when it's picked up by an LLM crawler or AIO. Models are trained on massive repositories of historical data compiled by scrapers, which is why your traditional SEO

tactics remain valuable. They also utilize retrieval bots to find the most current, relevant data on the live web to ground their responses.

Retrieval-augmented generation (RAG) bots are playing a major role in facilitating the shift toward AI-enabled search. These are crawlers specifically designed to scrape your content in real time and feed it to LLMs to inform generated responses. Data from TollBit showed [traffic from these retrieval bots grew 49% in the first quarter of 2025](#) compared to Q4 2024.

Information gathered from both traditional scrapers and RAG bots works together to inform visibility in AI-generated answers.

	Scraper Bots	Retrieval-Augmented Generation bot
<b>Data Focus</b>	Massive, historical snapshots of the web	Current, live, and context-specific information
<b>Interaction style</b>	Scrapes data in bulk, crawling entire sites	Focuses on targeted snippets specific to user prompts
<b>Timing</b>	Runs in the background, periodically consuming large amounts of data to refresh its dataset	Runs on demand in response to a user query
<b>Primary goal</b>	Data collection and storage	Answering questions in real time
<b>Key concerns</b>	Pre-existing domain authority and link quality	Technical semantic clarity
<b>Output</b>	Raw data	Context-aware, natural language answer

## Digital DNA: The entity map

To ensure data accuracy on the intelligent web, site owners have to start building contextual relationships across their content through a process called entity mapping. An entity map is an interconnected web of your brand's knowledge, where the relationships between the information on your site are as clear to a machine as they are to a human. Just as DNA is the blueprint that directs your body's development, an entity map provides the blueprint that directs the intelligent web as it learns how your content works together.

**An entity refers to a single, unique, well-defined concept that appears in your site content. Entity maps provide a visual representation of how these entities relate.**

For example, if you have content that references "New York City," "NYC," and "The Big Apple," an entity map would help provide the context that those three terms refer to the same location. If your site mentions Frank Sinatra, an entity map would add the additional context that he sang the song "New York, New York."

The primary language of an entity map is schema markup. It provides the explicit context that LLMs need to interpret and trust your data. By using standardized markup, you tell the AI exactly what it is looking at rather than forcing it to guess.

According to a controlled test from [Search Engine Land](#), the integration of advanced schema was the only differentiating factor between pages that showed up in an AIO and those that did not.

## The technical pillars of GEO visibility

There are three technical pillars that can help ensure your entity map is retrievable by both training sets and RAG workflows:

### Vectorization

Traditional search relies on exact keyword matching, but AI users tend to use more natural language prompts. Think about the difference between searching for "best running shoes" or asking an LLM to "provide a short list of good shoes for long-distance runners. I prefer if they are available in black, I need good arch support, and I wear a women's size 8."

Vectorization is the process of converting textual data into numerical values that bots can easily parse. When your content is vectorized, an AI can understand the concept of a query, even if the exact keywords don't match exactly. That information is stored in a [vector database](#) that indexes vectors to optimize the fast, accurate retrieval of information most relevant to the query.

This is the foundation of site-based conversational interfaces, like embedded chatbots. A vector database of your website's information grounds chatbot answers in the specific content of your site, including product catalogs, company announcements, events, programs, and more. A high-quality, well-maintained vector database can reduce inaccurate chatbot answers, facilitate better user experiences, and improve conversion rates.

## Server-side rendering (SSR) and agent-specific servers

For a bot to cite your content, it must be able to see it instantly. SSR ensures that the full on-page content is delivered in the initial HTML response, preventing bots from getting stuck on empty JavaScript shells.

That said, we are already seeing a shift toward [Model Context Protocol](#) (MCP). Instead of just rendering a page for a browser, MCP servers provide a standardized way for site owners to remove UI elements machines don't need and expose structured data and APIs for faster, clearer AI consumption.

MCP is also paving the way for Universal Commerce Protocol (UCP), which is designed to standardize how eCommerce transactions and product data are exchanged between different platforms. When a UCP checkout experience is delivered via MCP, an agent is able to communicate directly with a retailer to complete transactions. Where MCP gives an AI agent the legs to run out and fetch data for its user, UCP gives the agent the wallet it needs to purchase products directly.

**Where MCP gives an AI agent the legs to run out and fetch data, UCP gives it a wallet to make a purchase.**

## Clean, machine-readable markup

The key to making your content understandable for AI, search engines, and all kinds of bots is the use of clear, structured markup.

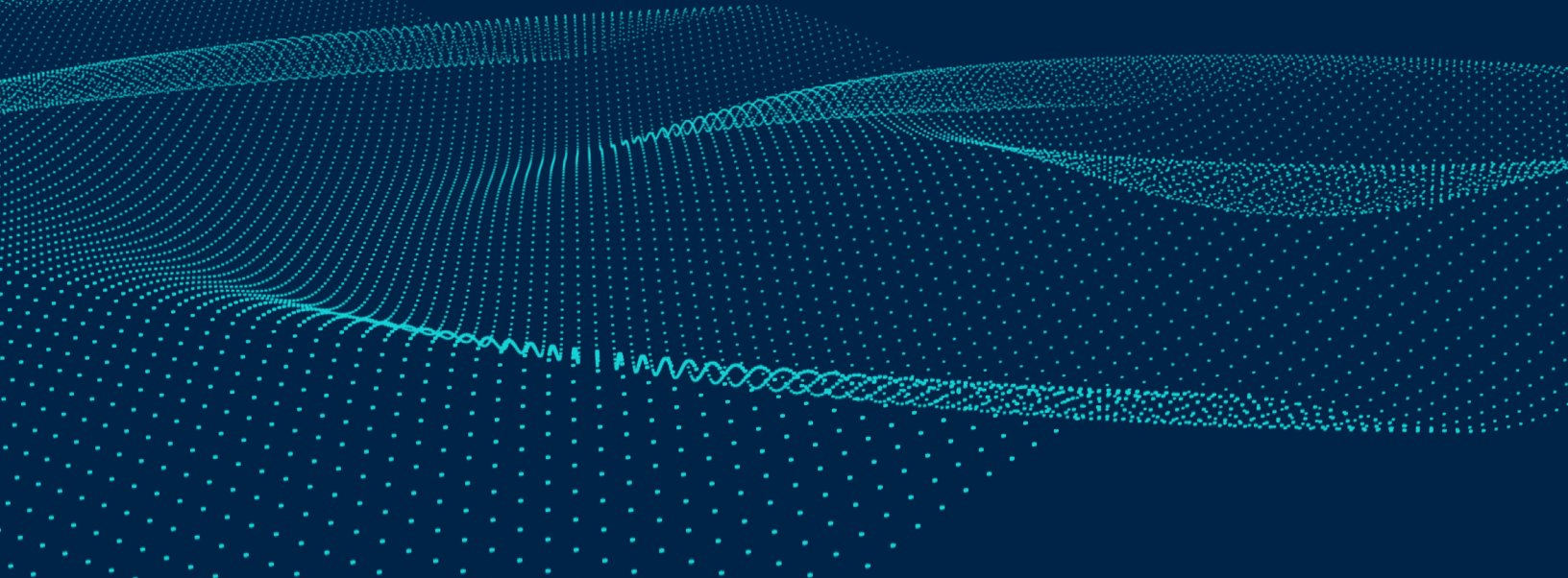
Semantic markup indicates the structure of content. It tells a browser, "this is a paragraph," or "this is a heading," to ensure your content hierarchy is clear.

Schema markup is a subset of semantic markup that indicates the nature of content. It goes beyond semantic markup to tell a bot, "this is a recipe," or "this is a product."

JSON-LD (JavaScript Object Notation for Linked Data) is the recommended technical format for implementing your markup. It is the language developers use to effectively translate your schema for bot parsing, so it doesn't get confused by any additional code used in your site design or UX elements. [Google recommends](#) JSON-LD as it is the format typically considered easiest to implement and maintain at scale.

AI parsers are designed for efficiency. Excessive code bloat, like heavy scripts and messy CSS, creates friction for AI agents. When properly structured, digital content becomes instantly machine-readable, providing the clarity search engines and bots use to prioritize the most valuable elements within a piece of content.

Prioritizing these three architectural pillars will help you change your website from a static asset into a queryable source of information for the bots that are training AI models. They provide the technical foundation that makes your brand impossible for AI to ignore.



# Structural changes for the intelligent web

To move from a passive digital property to an active participant in the intelligent web, shift focus from purely aesthetic updates to foundational changes.

Prepping a site for AI-driven discovery requires a three-pronged strategy: adopting an agile infrastructure, promoting authority through structured trust signals, and creating high-conversion destinations for already-informed visitors.

## Use composable architecture for agile content distribution

The problem facing most current sites is that they are monolithic: the front and backends are inextricably linked. This makes it much more difficult to distribute different versions of your content to the right audience at the right time. Monolithic sites often use slow, heavy

code that's difficult for bots to parse, and that code is often too rigid to offer meaningful personalization to humans.

In contrast, composable architecture—often referred to as [headless](#) or modular web design—decouples content from its presentation.

By packaging content into modular blocks, you can deliver the same information as a visually rich page or app experience for a human and as a raw data stream (such as Model Context Protocol or GraphQL) for an AI agent. [Gartner research](#) has posited that composable architecture is now key to building AI-native business operations, enabling rapid deployment of specialized data to various interfaces in a way that remains readable for agents.

## Establish authority to encourage citations

AI agents don't just look for information; they need information from authoritative sources to keep users happy. Experience, Expertise, Authoritativeness, and Trustworthiness (E-E-A-T) are no longer just SEO guidelines. The E-E-A-T model is now a filter AI models are using to ground their reasoning and avoid hallucinations. For example, [Google has rooted its AI Mode and AIOs](#) in its existing core Search, safety, and anti-spam systems, which the company has been refining for more than 20 years.

To bake authority into your structure, you must provide explicit trust signals that machines can verify. Examples include:

- **Verified author entities:** Use "[Author](#)" schema to build authorial integrity
- **Transparent sourcing:** Use schema like "[ClaimReview](#)" or "[Citation](#)" to prove the accuracy of your claims.
- **Third-party validation:** Highlight industry certifications and user feedback through "[Certification](#)" or "[Review](#)" markup.

## Craft high-intent endpoints to finish the funnel

The role of the website has shifted. Because AI agents now handle the education and solution research phases of the buyer journey, the users who eventually click through to your site aren't looking for a basic definition; they are looking to convert quickly.

This requires turning your most citable pages into conversion-optimized endpoints for high-intent customers. This includes:

- **Streamlined conversion paths:** Reduce friction for informed users who are ready to buy or book.
- **Personalized interactions:** Use [Smart Search](#), [chatbots powered by your site's existing knowledge base](#), or other internal AI tools to provide deep-dive answers that build on information the user already received from their AI-enabled search.
- **Substantive destination content:** Shift focus from top-of-funnel articles to in-depth white papers and case studies that provide the value AI-educated visitors demand.



# The foundation of discovery: Why your platform choice matters

The infrastructure supporting your digital presence is no longer a background player. At WP Engine, our position is clear: we aim to produce solutions that are human-first and AI-native.

Our tools and platform liberate content owners from technical complexity while still providing the developer-centric tools that enable true customization and AI-readiness. On our platform, customers can unlock the full economic potential of their content, providing delightful experiences for both humans and agents.

## Security and speed

The table stakes of the intelligent web remain the same. Security and speed still affect user retention, and they are the prerequisites for AI crawling. If a site is slow or unresponsive, RAG bots and scrapers will deprioritize it to save compute resources. A high-performance environment makes sure your site loads quickly enough for successful bot parsing when an LLM seeks real-time data.

Plus, our bot mitigation efforts successfully blocked more than 75 billion malicious bot requests in 2025 alone. Our platform is intuitive enough to keep the bad bots out while letting the good ones in to help you build your authority in AI-generated responses.

## Intelligent web readiness

Supporting the strategies that enable Generative Engine Optimization (GEO) requires a platform that's more than fast. WP Engine provides solutions to help our customers [create composable sites](#), [support advanced schema](#), [build unique vector databases](#), and [expose clean APIs](#).

Our teams are focused on building the tools that will help our customers stay ahead of the moving goal posts as the race toward AI preparedness continues to mature. In fact, we already offer a complete suite of tools designed specifically to help brands navigate the transition to the intelligent web:



### Managed Platform for WordPress®<sup>1</sup>

Our [best-in-class platform](#) is engineered to help customers create, manage, and publish content at scale and without technical complexity. We ensure the underlying code is clean and performant, so you don't have to.



### Headless Platform

Through [our headless solution](#), brands can serve a visually rich experience to humans while simultaneously delivering a raw, structured data stream to AI agents from a single content layer. Developers still get the freedom they need, but your entire headless tech stack comes from a single vendor.



### Advanced Custom Fields (ACF)

[ACF PRO](#) is the primary tool for transforming unstructured WordPress content into the structured, queryable data that AI agents crave. It allows you to define the Digital DNA of your brand. For headless builders, [WPGraphQL](#) makes [GraphQL](#)'s widely used query language and schema definitions readily available for hybrid WordPress builds.



### Managed Vector Database

Built specifically for technical teams, this fully managed service eliminates the need for custom pipelines or reindexing. Our platform automatically extracts, cleans, and vectorizes your site's data, syncing changes made to your live site in real time to keep your content AI-readable as it changes. With public API access for easy integration and a chatbot blueprint included, our [managed vector database](#) simplifies your team's process for plugging into the intelligent web.

<sup>1</sup> WP Engine is a proud member and supporter of the community of WordPress® users. The WordPress® trademarks are the intellectual property of the WordPress Foundation. Uses of the WordPress® trademarks in this ebook are for identification purposes only and do not imply an endorsement by WordPress Foundation. WP Engine is not endorsed or owned by, or affiliated with, the WordPress Foundation.



### Smart Search AI with AI-Powered Recommendations

Connect users to the information they'll find most valuable with a tool that provides fast, accurate results and personalized product recommendations. [Smart Search](#) is an easy-to-integrate tool that processes natural language queries and accounts for the semantic inconsistencies and typos that can render traditional search tools ineffective.



### Global Edge Security

This add-on feature combines Advanced DDoS Mitigation (Layer 3, 4 & 7), Origin IP Protection, a managed WAF, and Cloudflare's Global CDN and Automatic Image Optimization. With [Global Edge Security](#), our team handles the setup, ongoing management, and rule hardening for you, delivering set-it-and-forget-it protection in the face of [rising automated traffic](#).



### NitroPack

[NitroPack](#) helps WordPress site owners deliver faster, more stable site experiences and improve Core Web Vitals right out of the box. With more than 60 features working together to optimize images and videos, preload fonts, minify code, distribute content through extra CDN resources, deliver advanced caching, and more, NitroPack is an all-in-one performance solution that enables lightning-fast speeds without requiring technical expertise.

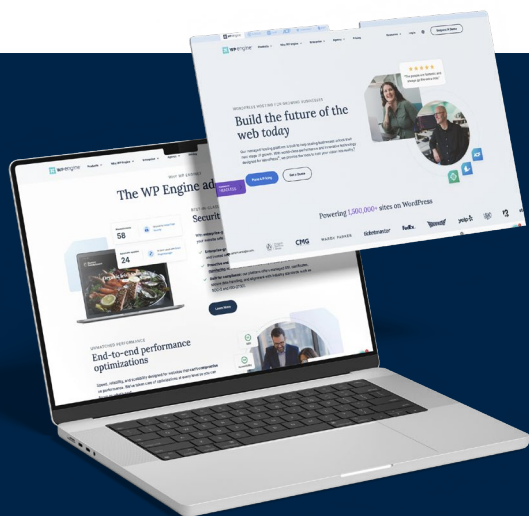
# WP Engine prepares digital properties for the intelligent web

Bots have irrevocably changed the face of the internet, but the web isn't dying; it is entering a more sophisticated, intelligent era. The transition from a link-based to a synthesis-based economy makes the quality and structure of your data more important than ever before.

To thrive in the intelligent web, brands must stop viewing their websites as static brochures and start seeing them as authoritative hubs of information. By prioritizing the agentic audience through clean schema and vectorization, while

rewarding the human audience with highly personalized, valuable content, you ensure your brand is citable and conversion-ready.

This new frontier rewards the prepared. Visibility is no longer a matter of luck or keyword density; it is a matter of architecture. Brands that build for intelligence today will be the ones that maintain visibility, build deeper trust, and grow market share in this new future. The intelligent web is here; it's time to make sure your brand is part of the revolution.



## How does your CMS investment compare?

Get a personalized TCO assessment with [WP Engine](#) and uncover key cost-saving opportunities, security enhancements, and performance optimizations tailored to your business. Future-proof your digital presence.

[Start your assessment today](#)



*WP Engine empowers companies and agencies of all sizes to **build, power, manage, and optimize** their WordPress websites and applications with confidence.*

The global technology company serves and powers more than 5 million sites, providing premium, enterprise-grade solutions, tools, and services, including specialized hosting platforms for websites built on WordPress, industry-tailored [agency](#) solutions, and developer-centric tools like [Advanced Custom Fields](#), [NitroPack](#), [Local](#), and more. Innovative technology and industry-leading expertise are why Pandora, Experian, Equifax, VMware, HubSpot, and Dell choose WP Engine. Learn more at [wpengine.com](http://wpengine.com).