

Pureinsights

Founded 2020 | HQ Washington, DC | 30 employees (approx.) | <\$5M revenues (est.)

Enterprise search and AI require careful curation, moderation, and management over time. The use of the Pureinsights Discovery Platform along with its associated services could indeed potentially transform a traditional enterprise search experience to work more like Google.



The Company

Pureinsights was founded in 2020 and is headquartered in Washington, DC. The firm is led by CEO and co-founder Kamran Khan. The leadership team at Pureinsights are enterprise search veterans who previously founded Search Technologies and then sold the firm to Accenture in 2017. In addition, the team all held past senior roles at key enterprise search vendors such as Excalibur and Convera. We estimate that the company has less than \$5 million in revenues, with approximately 30 employees. Version 1 of the Pureinsights Discovery Platform (PDP) was formally launched in June 2022 and is the focus of this report.



The Technology

Pureinsights provides several capabilities in the Pureinsights Discovery Platform (PDP). This platform aims to add value and orchestration to the core search functionality of open-source Solr, Elastic, or OpenSearch implementations. PDP leverages open-source components, but

more specifically it allows firms to deploy and operate a knowledge graph with their existing search system. The company claims it built the search components that most enterprises typically neglect to develop or use properly.

Knowledge graphs (aka semantic networks) are a means to represent relationships between objects, events, and concepts. As the name suggests, knowledge graphs are typically stored in a graph database and are similarly accessed and visualized through a graphical structure. They come into play here because an understanding of the relationships between items is generally missing or weak in enterprise search engines. Instead, traditional enterprise search engines rely on an index catalog of data elements that can be searched for, based on keywords. The engine becomes more nuanced by bringing a knowledge graph into the equation. Rather than simply providing a means for a user to “search,” the engine can actually “suggest” answers with a pretty high degree of accuracy.

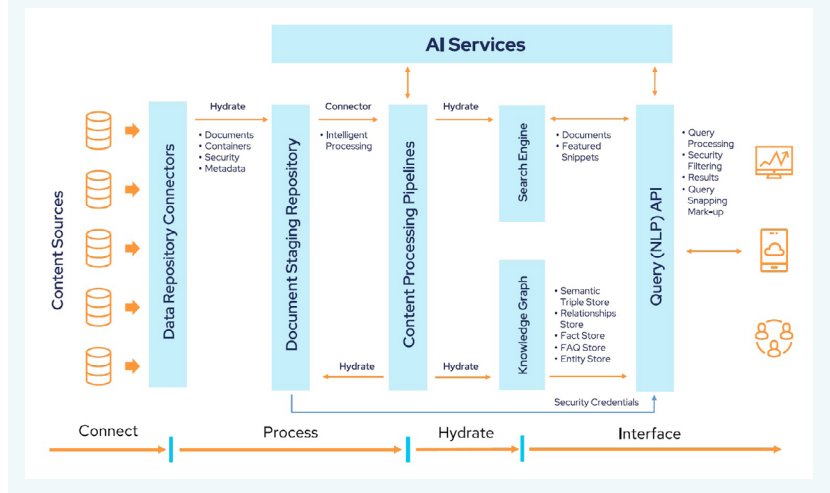
So, what PDP provides is a means to connect data assets, potentially from multiple databases, and – just as importantly – process

both the assets and any associated metadata. But first, PDP can enrich data assets by adding tags. Data assets such as documents can also be split into elements such as paragraphs and snippets. This is the starting point to vectorize the data, adding multiple values to data elements to provide a foundation for complex cognitive search, applied semantic understanding via natural language processing (NLP), machine learning, and populating the knowledge graph. There is some clever technology at work here, like leveraging semantic triples, BERT, and elements of Hugging Face, a reasonably new open-source set of NLP pre-trained modules to answer questions, recognize similarities, summarize, etc.

PDP has been designed to connect to pretty much any file system, database, or third-party application to do all of this. Each customer will be different and have quirks, so such complex integration, though accelerated through PDP connectors, may still require some configuration. But we did note that the PDP connectors are designed to monitor and update new data assets or, for that matter, deletions, at regular intervals. A fairly obvious requirement, but we have encountered federated search engines that have trouble keeping connections and assets updated.

In simple terms, PDP works through a four-step process (see Figure 1). The first is connecting the different data sources as mentioned above. Next, the data assets are processed. For larger files and data stores, the assets would typically be copied to a cloud staging environment where the data is analyzed and, where necessary, cleaned, normalized, and enhanced. This approach also allows for large volumes to be broken and batch processed in manageable chunks. After this stage, the data is published to the search engine and knowledge graph.

Figure 1
Pureinsights Discovery Platform – AI Services



The last step is developing a user interface via an API for the new search experience and the query-parsing to understand the user's intent when searching.

The goal of all this work is, to use Pureinsights' terminology, to make enterprise search "work like Google." That's a bold claim, and though that is something many end users want to see and experience, few vendors can provide it. The challenge for enterprise search vendors is that public search engines crawl across mountains of well-tagged assets that have been designed to be found easily. Enterprise search is a different kettle of fish in that it crawls across mountains of poorly labeled and hard-to-identify assets. Hence, the goal of PDP is to clean and process all of those "unfindable" assets to make them easily "findable." In short, PDP is not a search engine in and of itself; instead, it provides the means to process and query search engines more intelligently and efficiently. Therefore, rather than replacing a pre-existing search engine, it augments and improves it.

Our Opinion

Though not a unique proposition, the hybrid product and services business model of Pureinsights caught our attention. Enterprise search and AI, in general, are not technologies that can just be switched on and left to their own devices; they require careful curation, moderation, and management over time. The use of PDP, along with its associated services, could indeed potentially transform a traditional enterprise search experience to work more like Google.



Advice to Buyers

If you are using or plan to use Solr, Elastic, or OpenSearch as your foundational enterprise search technology then you may want to look at Pureinsights. Good enterprise search requires specialized skills that most firms do not have, so the managed service, business process outsourcing (BPO)-style approach offered here (which Pureinsights calls “SearchOps”), along with the add-on products, makes a lot of sense. Moreover, the PDP capabilities provide you with the tools to create a more intuitive and hopefully more relevant search experience for your users.

SOAR Analysis

Strengths

- Deep expertise in enterprise search
- Enhances and improves, rather than rips and replaces, existing search environments

Opportunities

- Build on existing large open-source search systems
- Deliver a full BPO-style service for enterprise

Aspirations

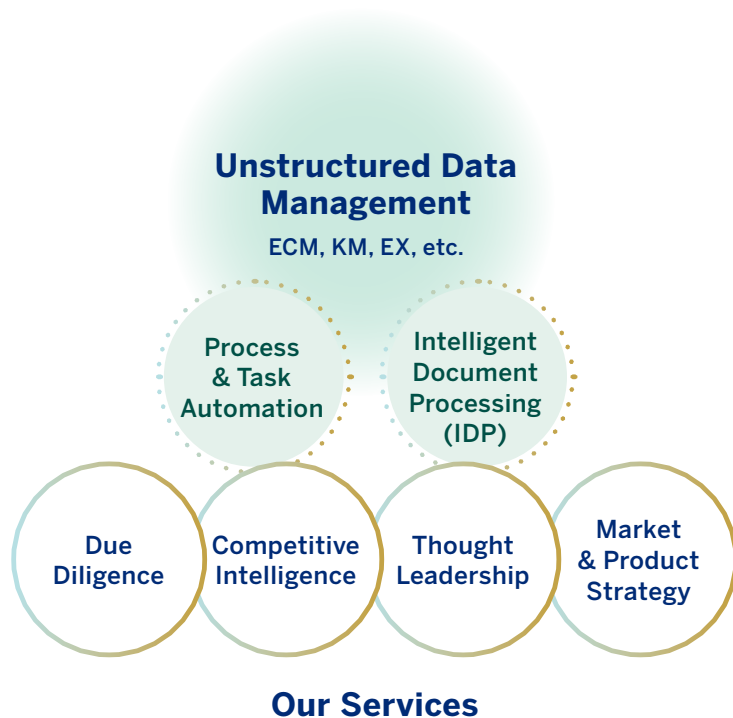
- Become the standard platform for enterprise search integrating knowledge graphs and AI
- Work with any enterprise search engine

Results

- Bootstrapped and profitable
- Already acquired notable customers

About Deep Analysis

We Research Innovation



Deep Analysis is an advisory firm that helps organizations understand and address the challenges of innovative and disruptive technologies in the enterprise software marketplace.

Its work is built on decades of experience in advising and consulting to global technology firms large and small, from SAP, Oracle, and HP to countless start-ups.

Led by Alan Pelz-Sharpe, the firm focuses on Information Management and the business application of Cloud, Artificial Intelligence, and Blockchain. Deep Analysis recently published the book "Practical Artificial Intelligence: An Enterprise Playbook," co-authored by Alan Pelz-Sharpe and Kashyap Kompella, outlining strategies for organizations to avoid pitfalls and successfully deploy AI.

Deep Analysis works with technology vendors to improve their understanding and provide actionable guidance on current and future market opportunities.

Yet, unlike traditional analyst firms, Deep Analysis takes a buyer-centric approach to its research and understands real-world buyer and market needs versus the "echo chamber" of the technology industry.

Contact us:

info@deep-analysis.net
+1 978 877 7915