

Leveraging Data Intelligence for the Life Sciences Industry

DATAVID

Before we start: A primer on data intelligence

Before we delve into leveraging intelligent data insights, let's discuss what this term means. Data intelligence is the practice of using analytical tools and methods—such as artificial intelligence or machine learning—to analyse and transform large datasets into new insights.

The application of these tools helps enterprises **develop a better understanding of the data they collect**, and how it can be used to improve services, investments and processes now and in the future.

Often confused with business intelligence (BI), data intelligence is more concerned with *extracting additional value out of disparate (i.e. unstructured) data* which would otherwise go unused or unnoticed.

Instead, business intelligence may focus on things like business performance, customer behaviour and demographics, and predictive KPIs. Its primary concern is to help visualize structured, known data.

Table of contents

> Leveraging Data Intelligence for the Life Sciences Industry.....	1
> Before we start: A primer on data intelligence.....	2
» Why data silos are a big problem in life sciences.....	4
» How cognitive search removes data silos...	7
» Life sciences: How to implement search the right way.....	9
» Investing in data intelligence: 3 best practices.....	12
» A security-first approach to data integration.....	15
» Choosing a trusted partner in data intelligence.....	18
> Conclusion.....	20

Why data silos are a big problem in life sciences

Science has always benefited from collaboration.

But in life sciences, the biggest obstacle to using advanced data analysis isn't a lack of skilled data scientists or technology; *it's getting access to all of the data available.*

Analysis of large amounts of data can make a revolutionary difference in drug research and innovation. **Shared data benefits the life sciences industry through expanding research, faster drug development and lower costs**, while internal and external data can help an organisation gain a complete view of their data to generate accurate insights.

It makes sense that when data is isolated in silos, it causes big problems. Even with research and development increasing and genetic and immunological advances, limited access and fragmentation of data are slowing down progress.

Silos have real-life consequences including; undermining drug development, limiting collaboration, and increased costs.

(Continues on page 5)

Why data silos are a big problem in life sciences

Silos tend to arise in life sciences enterprises for a few reasons:



Non-standardisation of data: Collecting, consolidating and standardising data takes time. When data is not standardised, an organisation cannot combine or compare data sets easily across departments.



Multiple data sources: Without the right interface, and with loosely-structured data from multiple sources, collecting both internal and external data can be difficult, and it makes gaining accurate insights near impossible.




Infrastructure: Despite the benefits of integrated data analysis, some organisations see the cost as prohibitive, forcing them to use subpar, siloed data solutions.



Company mindset: Companies fear sharing data and losing their knowledge advantage, or they don't realise there is a need for shared knowledge within the company.



Sensitive information: Many data sets are not shared freely for good reason as organisations have a duty to keep control over sensitive information.



Life sciences should aim for rapid-delivery in research and development and bringing safe, innovative products quickly to market. But how do you do this while sifting through a growing amount of disparate data?

How cognitive search removes data silos

Cognitive search (sometimes referred to as intelligent search) is powered by artificial intelligence and **enables life scientists to make use of all data across their organisation**—including previously unsearchable text like Word files, spreadsheets, and presentations. *But what exactly is cognitive search?*

Where traditional search relies heavily on the user inputting the right keywords to find relevant information and matching the query exactly, cognitive search is more like a conversation between you and the information you are looking for.

When the information you need is buried within the text of a research paper and doesn't contain any specific keywords, traditional search is unable to surface it.

By using natural language interactions between you and your data, *cognitive search can find information that would previously be ignored* for being unreadable, buried within thousands of documents, or stored in underused data sources.

Cognitive search breaks down data silos, increases speed, and ensures decision making is made with access to all available information through:

1) Discoverability

Seamlessly integrating data from multiple sources automatically which reduces the time spent searching for data in multiple locations.

2) Relevance

Returning information in a relevant, entity-driven format that tells you everything you need to know about what you are searching for.

3) Redaction

Understanding the context of search and preventing unnecessary information from surfacing in search results (e.g specific patient information and sales).

4) Speed

Retrieving information quickly, allowing to act on it a lot faster and more effectively. This is great for compliance purposes as it allows quick auditing.

Life sciences: Implementing search the right way

We've discussed how cognitive search differs from traditional search products. With cognitive search, AI not only extracts meaning from the content being searched, it learns while doing so to **provide increasingly relevant and complete results to users.**

Since human language is so complex, it's very complicated to turn words and phrases into something computers understand.

Context can change the meanings of certain words and, with scientific terms, add an extra set of challenges.

To ensure cognitive search is successful in a scientific context, there are a few steps to take.

These include data sourcing, automated integration, natural language processing and redacting unnecessary information from the data source. Identifying the right data sources—*both internal and external*—is crucial to implementing cognitive search.

Good data should improve business and research decisions, leaving out potentially-harmful or inconclusive sources.

(Continues on page 10)

Life sciences: Implementing search the right way

Also, integrating data from various sources into one single hub is essential to accelerate innovation while keeping costs down.

An automated pipeline removes manual efforts and easily combines data from various sources into one place, structuring data based on the definitions assigned by your company.

Because human language is concerned with semantics, cognitive search uses Natural Language Processing (NLP) to help make sense of the content it is searching.

NLP makes connections between language blocks by *separating language into fragments*, and filtering out words that add no value to gain an understanding of what is being processed.

The successful implementation of cognitive search relies on completing these steps with input from the user and an understanding that each organisation is different.

Life sciences organisations need a strong foundation for cognitive search to successfully work which means taking care to include all relevant information for researchers to derive trusted insights.

DATAVID

Cognitive search in life sciences: a real-world case study

Learn the incredible benefits that a global agricultural enterprise experienced when they invested in data intelligence:

- **Improved discoverability** for 16M+ documents across 22 data sources
- **Reduced document research time** from days to just a few minutes
- **Increased capability to deliver** innovative products and reduced time to market

... and more.

With cognitive search, your life sciences organisation can benefit from similar results too. Read the full case case study below.

[READ CASE STUDY](#)

Data intelligence: 3 best practices for investment

As technologies like artificial intelligence and cognitive search get more advanced, there is a belief that human expertise is no longer needed. This is just not true. **Human expertise is the only way cognitive search can be implemented successfully.**

Cognitive search might be able to give us relevant answers, but human engineers are the only ones who can ask the questions and derive the information needed to succeed.

The technology relies on a high level of customisation for the specific use case: what data is needed, what context is this going to be applied in, what ontologies are available, and so on.

All of these aspects combined provide unique value to each individual life sciences organisation.

Since a successful deployment of data intelligence solutions is not just about having the right tech, enterprises need to focus on investing in a *culture that adopts an agile mindset*. It's an evolving practice which requires an iterative delivery approach.

(Continues on page 13)

Data intelligence: 3 best practices for investment

Adopting an agile mindset to both culture and technology will get businesses further and help overcome the shortcomings of rigid processes like traditional ETL (Extract, Transform, Load).

As teams get access to more information, discover more powerful uses for it, and collaborate to innovate faster, *an agile mindset becomes the "glue" that holds the entire effort together.*

That's when technology comes into play.

If your organisation is using point-to-point integrations to move data through your IT environment, then your team is missing the opportunities a modern data hub can provide.

While data warehouses and data lakes act as endpoints for data collection, **data hubs serve as conciliatory points and help facilitate data sharing**—eliminating data silos.

Data hubs can organise data efficiently, store it cost effectively, and push it towards key business functions. They are easily integrated, enable easy de-duplication and data standardisation, and are highly secure and scalable environments.

Follow these 3 best practices to invest in a successful implementation of your data intelligence project:

1) Focus on the experts

Investing in data intelligence means enhancing scientists' work, not putting another technology roadblock in front of them. Your project requirements should start with their needs in mind.

2) Adopt an agile mindset

Collaboration is key to making the most of the information provided by a cognitive search solution. Being agile means adding value to the researcher experience iteratively.

3) Leverage data hubs to their full extent

Once the requirements are clear, a well-implemented data hub will solve 90% of the foundational issues resulting from data silos. Data hubs are the fundamental infrastructure that allows to develop solutions like automated data integration, entity extraction, cognitive search, and more. It's important that you consider the technology carefully; MarkLogic is renowned for its easy integration coupled with great search.

A security-first approach to data integration

As cognitive search often needs access to large data lakes and various sources of data, these sources can be at risk of leaks, attacks, or manipulation.

A lack of security around these sources has the potential to enable attackers to gain access to information from patients as well as operational insights, damaging your intellectual property.

Therefore, the process of **integrating data for cognitive search must be followed with security in mind**. In life sciences, insecurely held data can result in more strict audits, regulatory fines and damage to the integrity of the organisation.

Pharma companies are particularly at risk of non-compliance if they cannot meet the challenges of increased regulation.

In 2019, the FDA publicly rebuked a large Swiss pharmaceutical company for including manipulated data in the approval of a gene therapy drug. Whether intentionally or not, they failed to report the falsified data before gaining agency approval.

(Continues on page 16)

A security-first approach to data integration

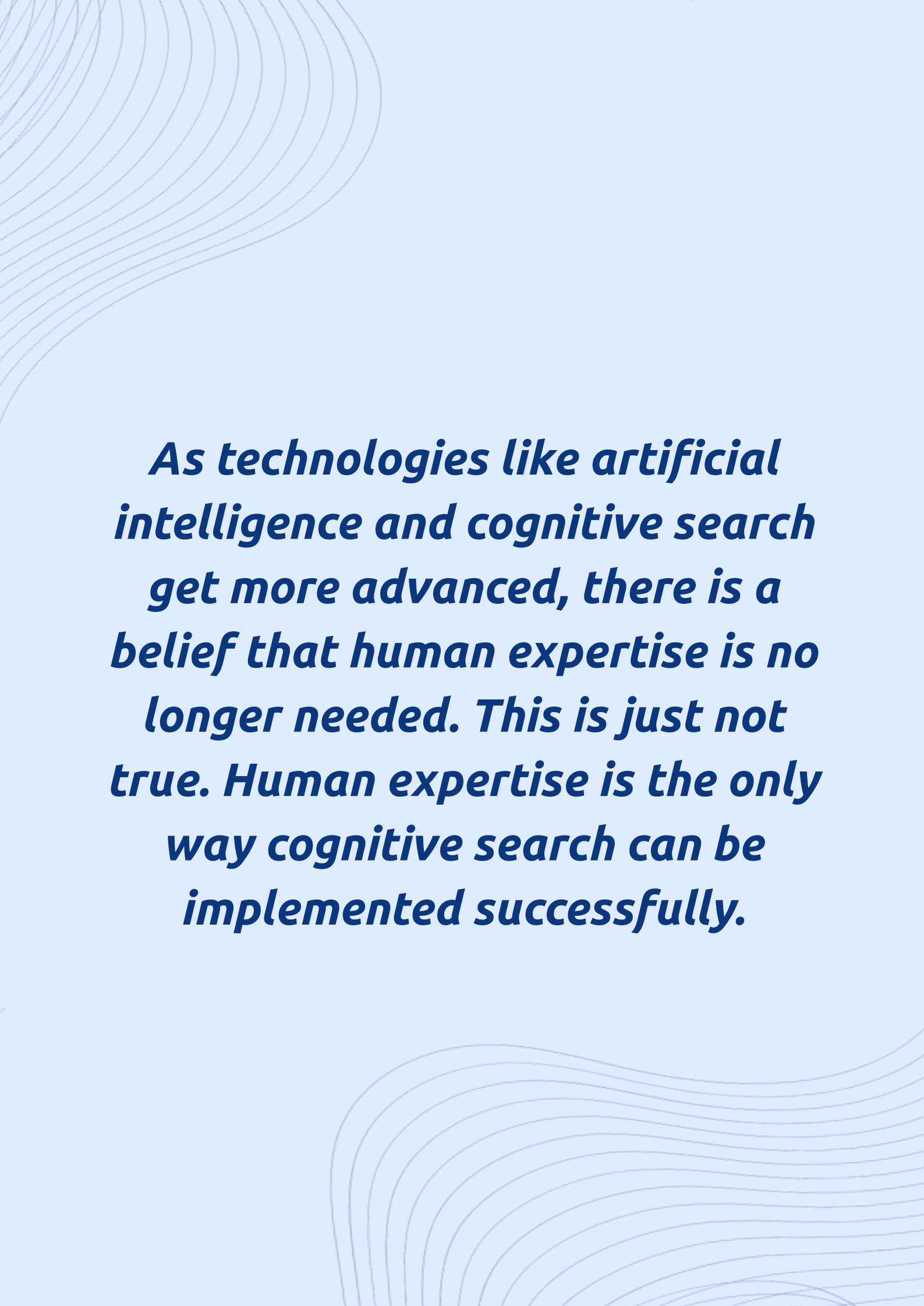
Transparency in the collection and analysis of data can only be a good thing and protects individuals and the organisation as a whole from situations where data can't be certified.

A **secure search environment** helps you achieve that.

Malware, ransomware, cyber and phishing attacks can all hit life sciences companies, grinding research to a halt and pushing pause on clinical trials. Sometimes, *data is even lost forever*.

In 2017, a German pharmaceutical giant was the victim of a devastating cyberattack which spread malware across the organisation, taking down around 30,000 computers across sales, manufacturing and research units, causing \$870M worth of damage and disrupting production of a new vaccine.

Data intelligence can have a positive effect on data security as compliance is not only addressed within the confines of each individual data silo but across the whole of the organisation. This gives a complete picture of whether the data is compliant or not.



As technologies like artificial intelligence and cognitive search get more advanced, there is a belief that human expertise is no longer needed. This is just not true. Human expertise is the only way cognitive search can be implemented successfully.

Choosing a trusted partner in data intelligence

The benefits of data intelligence (*and especially cognitive search*) for life sciences are clear: zero data silos, better discoverability, higher productivity, and reduced time to market.

But actually getting to experience those benefits is a risky journey. Whenever there is confidential data involved, you need to work with the right people, and hire the right skillset.

Some of the things to look for are:



A **team of data-driven, senior developers**. Data intelligence problems are hard to solve, and they require a deep understanding of the technical challenges ahead.



Years-long experience in areas such as Artificial Intelligence, Natural Language Processing, Machine Learning, and data-intensive systems architecture.



A **proven record of delivering value** to other life sciences enterprises without jeopardising data integrity.



Clear communication that sticks to achieving your organisation's goals and doesn't deviate from it.

(Continues on page 19)

Choosing a trusted partner in data intelligence

All of these elements are hard to find in an increasingly competitive business environment where the right talent and low time to market define the difference between failure and success.

Datavid works with multiple enterprises in the life sciences industry, all of them operating at a global scale.

We've completed a variety of projects successfully on their behalf, including powerful cognitive search platforms.

The company is committed to continue to deliver this value to our customers in the form of bespoke services and products.

Our latest development, **Datavid Rover**, is a complete data intelligence platform that provides all the benefits of our personalized services at a fraction of the delivery time.

This increases the likelihood of your project succeeding tenfold, and it provides a great user experience for scientists looking to gain valuable information to advance their research.

We're enthusiastic to see our customers succeed, and we have one last word for you as you've reached the end of this guide...

(Conclusion on page 20)

DATAVID

Thank you for trusting Datavid in your data journey

If you've learned something new in this guide, or if you're more aware of the benefits of data intelligence for life sciences organisations, our job is done.

Datavid is thrilled to continue educating business leaders of the value that data carries, and we're fully committed to helping you extract that value.

If your company is looking for a trusted partner to eliminate data silos and achieve new heights in productivity and innovation, consider booking a free consultation with one of our data experts.

[GET A FREE CONSULTATION](#)