



# Conversational Data Intelligence

by Kevin Rohling

Unlocking insights and competitive gains from vast data sets is a long-standing business challenge. Traditional methods of sifting through unstructured data, particularly human language, are time-consuming and labor-intensive. Off-the-shelf analytics tools are often expensive and still stumble when faced with diverse data formats - such as text, documents, or natural language. Even when data is well-structured, a data scientist is often needed to extract actionable insights.

Conversational AI changes the game. Conversational AI does more than decode the intricacies of human language; it fundamentally reshapes how we interact with data. Large Language Models can help us draw useful business and product insights using simple conversational questions. Now, posing questions about complex datasets and receiving answers is as natural as having a conversation. At Presence, we call this new AI-enabled method of querying data, Conversational Data Intelligence (CDI).

In this report, we'll examine both the core technologies behind CDI and its real-world business applications, to help you uncover new opportunities for driving business efficiency and innovation in processes, workflows, and products. We will explore:

- 01

Real-world applications of CDI for product innovation and business workflows across industries
- 02

Core technologies that enable the translation of conversational human queries into structured, actionable insights
- 03

Security risks and challenges of adopting and implementing CDI
- 04

What the future might look like as CDI evolves and becomes widely integrated

# What is Conversational Data Intelligence?

Conversational Data Intelligence (CDI) is a fusion of advanced AI technologies and data analytics that allows users to conversationally interact with vast amounts of data. Rather than relying on traditional, complex interfaces, this novel application of technologies offers a more intuitive and dynamic way of extracting insights and making data-driven decisions. What exactly sets CDI apart and how does it work?

## A Dialogue with Data

The beauty of CDI is its ability to simplify the interaction between humans and large datasets. Imagine being able to ask a database a question just like you would ask a colleague, and receiving not only an answer but also meaningful insights, recommendations, or even visual interpretations. This is the promise of CDI — a dialogue with data.

## Beyond Surface-level Queries

CDI delivers more than straightforward answers to queries, it can understand context and infer user intent to provide deeper insights. Going beyond simple tasks like retrieving specific data points, CDI can pull together comprehensive narratives, trends, and interpretations from seemingly disparate sources.

## Bridging Structured and Unstructured Data

While it's relatively simple to query and interpret structured data (e.g., databases, spreadsheets, and machine-readable file formats), it's significantly more challenging to gather insights from unstructured data, such as emails, documents, or social media chatter. CDI shines in its ability to unify insights across both structured and unstructured data.

## Dynamic and On-the-Fly Analysis

CDI is much more than asking and retrieving; its analysis evolves over time. As a conversation progresses, so does the depth of analysis. A user might start with a general inquiry and as CDI provides answers, the user can request clarifications, drill down for more specificity, or pivot to a related topic, all without starting from scratch. Find out how in the “The Building Blocks” section below.

## Democratizing Data Access

One of the most important and transformative aspects of CDI is its potential to democratize access to data insights. Instead of data scientists and analysts being the gatekeepers of data analysis and interpretation, professionals across departments will be able to mine datasets for insights and unique value. The only technical proficiencies required to achieve useful results will be the ability to ask questions in natural human language and an understanding of the business domain.

In essence, CDI is a paradigm shift in our relationship with data. It ushers in a future where our ability to gain insights from data are not defined by the limitations of our tools or technical training, but by the depth of our curiosity. Next, let's dive into the building blocks that make this revolution possible and explore how businesses can harness its potential.

# The Building Blocks

What we call CDI is not one technology, but a series of innovative components working together. Each component provides unique capabilities that, when combined, yield a holistic approach to data interaction and analysis. In this section, we'll unpack these foundational components: Large Language Models, Semantic Search, Natural Language to Query, Hybrid Search, and Autonomous AI Agents.

## 01

### Large Language Models (LLMs)

Large Language Models, such as GPT-4 and its peers, are intricate AI architectures trained on vast amounts of text data. Their design allows them to generate human-like text, comprehend context, and extract semantic meaning from input.

LLMs form the backbone of the conversational interface in CDI. Whether it's understanding context, making inferences, or handling complex queries, LLMs are the first point of contact, translating human curiosity into data-driven inquiry.

## 02

### Semantic Search

Semantic Search examines the meaning and relationships within data, ensuring relevant, context-aware results. Rather than relying on keyword matches, Semantic Search harnesses the power of AI generated embeddings to understand context and intent. At a high level, embeddings are numeric representations of words or phrases in a dense vector space. By positioning similar words or concepts closer together, embeddings capture semantic relationships and nuances. For instance, 'king' and 'queen' might be closer in this space than 'king' and 'apple.' These embeddings, usually obtained through deep learning models, enable Semantic Search to transcend keyword matching and instead return results based on context and meaning.

## 03

### Natural Language to Query

The 'Natural Language to Query' mechanism translates human-language queries into structured queries that can be executed against specific datasets. Natural Language to Query converts our everyday language into a structured query via code, be it SQL, Python, or any other querying language.



By understanding the intent and specifics of a question, thanks to LLMs, structured queries can be formulated to dive into databases or other analytical platforms. Whether you're asking for "the total revenue of Q3 2022" or curious about the "correlation between marketing spend and sales," this technology structures your query in a way that the system can comprehend and then fetches or calculates the precise answer.

# 04

## Hybrid Search & Query

Hybrid Search is the practice of amalgamating Semantic Search's contextual comprehension power with the precision of structured dataset queries. Hybrid Search blends the best of both worlds; contextual comprehension and deterministic filtering and sorting. Hybrid Search can, for instance, combine the semantic meaning of a product review with structured data such as sales figures, giving businesses richer, more context-aware insights.

# 05

## Autonomous AI Agents

Autonomous AI Agents can be thought of as digital assistants, capable of turning complex queries into actionable steps and strategies that can then be executed without intervention. Trained across a multitude of data environments, they can easily navigate diverse data sources and formats.

For example, imagine a query that requires insights from a legal document, sales data from a CRM, and social media sentiments. The AI agent can deconstruct this complex request, devise a strategy to extract necessary information from each source, reconcile discrepancies, handle potential errors, and finally, fuse this information into a coherent, insightful response. In a world of numerous varied data types this context-aware strategic execution is invaluable.

# Use Cases

CDI is more than a technological marvel; it catalyzes change across diverse sectors. By bridging the gap between datasets and human queries, CDI is redefining how industries access, interpret, and utilize information. In this section we spotlight the transformative potential and breadth of business applications across multiple verticals.

## Healthcare and Life Sciences

**Research Data Analysis:** In medical research data is paramount. With CDI researchers can parse intricate patterns within large datasets effortlessly. For example, professionals can ask, "Show me patterns of symptom X across patients aged 20-30 in the past year", and receive insights that could have taken weeks or months with traditional data analysis methods. CDI's accessibility will be a breakthrough for medical research, resulting in accelerated diagnoses, pharmaceutical treatments, and other healthcare innovations.

**Drug Discovery and Trials:** The path to drug discovery is riddled with intricate analyses and balancing efficacy with side effects. CDI offers pharmacologists a means to quickly synthesize qualitative and quantitative insights from large datasets, including patient data, research publications, and clinical trial results. By querying potential interactions or side effects, researchers can ensure safer, more efficient drug development cycles. For one of our healthcare research clients, we implemented a CDI system to query pharmaceutical research data, significantly reducing the time and effort required to interpret clinical trial results. Insights into drug efficacy, adverse events, biomarker trends, and patient demographics can be generated quickly with CDI instead of waiting days or weeks for a custom report.

## Legal

**Case Precedent Analysis:** The legal realm is characterized by its reliance on precedents. CDI can reduce hours of sifting through legal databases to mere seconds. Lawyers can ask the system about past cases with parallel circumstances to quickly amass valuable insights that will inform their legal strategies or bolster court arguments.

**Contract Review:** Contractual documentation, often dense and verbose, can be a source of inefficiency. With CDI, legal professionals can bypass the tedium of manual review. For one of our partner projects, we're developing a CDI system to pinpoint clauses related to specific conditions or terms, to streamline contract analysis, and ensure more efficient, comprehensive reviews.

## Finance

**Investment Analysis:** In the dynamic world of finance staying informed is paramount. Investment analysts equipped with CDI can query the performance trajectories of specific sectors, stocks, or commodities. With real-time responses analysts can make more informed decisions, capitalize on trends, and mitigate risks.

## Education

**Research Paper Analysis:** The academic sphere is flooded with new publications every day. CDI can be used by academics to navigate this flood of information. By requesting papers on niche topics, comparing findings, and spotlighting research gaps, academics can enhance their research and stay on top of emerging trends.

**Autonomous Research:** Revolutionizing traditional research paradigms, CDI offers the possibility of autonomous research. The system can be directed to review a multitude of papers and online resources, and synthesize all the information to craft novel research proposals or draft preliminary papers. This could expedite research cycles and open doors to previously unexplored intersections of knowledge.

These use cases underscore just a fraction of the transformative potential of CDI across sectors. These applications reveal how enhanced data efficiency, accessibility, and comprehension can lead to better decisions, superior industry innovation, and increasing business value.

## Government and Public Services

**Public Feedback Analysis:** In democratic settings, public sentiment is invaluable. Government bodies, with the aid of CDI, can swiftly assess public sentiment concerning specific policies or events. By analyzing data from social media and other public platforms, governments can maintain a pulse on their constituents, guiding policy decisions.

**Resource Allocation:** Governments are expansive entities handling myriad projects and managing expansive budgets. CDI can help ensure optimal resource allocation. Queries about the utilization of public funds, the progress status of projects, or staff deployments can offer officials a comprehensive overview ensuring public resources are used judiciously and efficiently.

# Risks & Challenges

As with any groundbreaking technology, CDI presents a new frontier of opportunities as well as new risks and challenges. To harness the potential of CDI responsibly it's imperative to be aware of them. Below we explore a few of the risks and challenges in CDI.

## Critical CDI Considerations

- ☐ **Data Privacy and Security:** Critical defense against data leaks in advanced data systems
- ☐ **Bias and Fairness:** Address data bias to ensure fairness
- ☐ **Dependence on Quality Data:** Accurate insights rely on top-quality data
- ☐ **Over-reliance:** Maintain balance between tech reliance and human judgment
- ☐ **Accuracy and Validity:** Accuracy in high-stakes data outputs is essential
- ☐ **Potential Misinterpretations:** Careful interpretation needed for complex data language
- ☐ **Complexity/Cost of Deployment:** Challenging and costly to start and integrate advanced data systems

**Data Privacy and Security:** In an age where data breaches are increasingly commonplace, the emphasis on data privacy and security cannot be overstated. CDI systems, given their access to and processing of enormous datasets, raise concerns about the storage, management, and protection of this data. Robust encryption, access controls, and regular security audits are essential to prevent unauthorized access and safeguard sensitive information.

**Bias and Fairness:** CDI's intelligence, primarily driven by the data it's trained on, is susceptible to inheriting biases present in that data. These biases can inadvertently perpetuate or even amplify existing stereotypes or prejudices leading to skewed or discriminatory results. This is particularly concerning in applications with societal impacts, like healthcare, government, or legal advisories. Continuous efforts in refining algorithms and consciously curating unbiased training datasets are essential to uphold the principle of fairness.

**Dependence on Quality Data:** The adage "garbage in, garbage out" is especially apt for CDI. The quality of insights is directly proportional to the quality of data it processes. Inaccurate, outdated, or incomplete data can lead the system astray, yielding misleading or even erroneous conclusions. Vigilance in data curation, validation, and continuous updating is critically important.

**Over-reliance:** The ease and efficiency of CDI might tempt professionals to over index on it, sidelining their own expertise and judgment. Such over reliance could stifle human intuition and expertise potentially leading to oversights. It's important to view CDI as a complement, not a replacement, to human intellect.

**Accuracy and Validity:** As CDI technologies become more integral across industries the stakes for ensuring output accuracy and validity rise. The quality of insights generated by CDI is highly dependent on its training data – flawed data will lead to faulty or misleading output. The nuanced and context-rich complexity of human language adds another layer of interpretation challenges, and reliance on external data or multi-step queries can introduce additional points of deviation or error.

To mitigate these challenges human oversight is crucial. Expert validation of CDI insights is indispensable, particularly in high-stakes domains like healthcare and finance. Continuous system recalibration, guided by real-world feedback, can refine its interpretative accuracy. Employing rigorous data vetting processes, especially for external data sources, can safeguard against information decay or bias. And finally, fostering a collaborative, interactive ecosystem where users can highlight errors or provide contextual cues can help refine and enhance CDI's responsiveness and precision.

**Potential Misinterpretations:** The richness of human language, laden with nuances and contextual cues, poses a challenge for even the most advanced CDI systems. Ambiguities or complexities in queries can lead to misinterpretations, and consequently inaccurate results. While advancements are ongoing, users should remain vigilant and cross-check results when uncertainty arises.

# The Future of Conversational Data Intelligence

CDI stands at the precipice of a transformation in how we interact with data. This future will redefine our engagement with information, evolving from static storage to interactive dialogues, heralding industry innovation and change. Let's explore some of these exciting horizons.

## Virtual Team Member

Beyond serving as a mere tool, CDI is on track to become an integral part of organizational teams. It will be able to join in collaborative brainstorming, provide real-time insights, and even foresee team queries based on ongoing discussions. This virtual team member will bring data insights to the table and infuse team decisions with actionable, empirical evidence.

## Enhanced Personalization

The future CDI systems will not just respond to queries, but will intimately understand the user's unique preferences, past interactions, and specific needs. This advanced level of personalization will ensure that responses are tailored, not just to a generic user, but to the individual, making data interactions feel intuitive and bespoke.

## Multimodal Interactions

The conversational paradigm will extend beyond text-based interactions. We can expect a future where CDI systems will seamlessly integrate voice, video, and even gestures. Users might vocalize a complex query, see a visual representation of the data, and interact using hand gestures, all within a unified CDI environment.

## Domain-specific Expertise

The broad knowledge base of current CDI systems will be complemented by models possessing deep specialized expertise in specific sectors. Be it law, medicine, finance, or any niche domain, these expert CDI systems will cater to professionals offering insights and data interpretations aligned with the nuances of each field.

## Proactive Insights

Instead of merely reacting to posed questions, future CDI systems will be proactive. They will anticipate potential inquiries, provide unsolicited yet relevant insights, and highlight intriguing patterns. Such a proactive approach will ensure users stay a step ahead, capitalizing on opportunities or mitigating risks even before they realize the need to do so.

## Autonomous Actions from Insights

A truly groundbreaking frontier for CDI is the integration of actionable intelligence. As CDI interprets and understands data it can also be paired with systems that autonomously act upon those insights. Imagine a CDI system that after analyzing a company's financial data not only provides insights, but also autonomously makes investment recommendations or even executes low-risk transactions. This synthesis of understanding and action promises to make CDI an active participant in data-driven decision-making processes instead of just a passive advisor.

Overall, the horizon of CDI holds abundant promise for innovation, efficiency, and efficacy across industries. From personalized interactions to proactively driving action CDI is set to redefine our relationship with data ensuring that our conversations with datasets are as dynamic, insightful, and actionable as those with our most trusted advisors.

Extracting actionable insights and business value from extensive unstructured data has long been fraught with challenges such as costly tools, specialized skill sets, and resource-intensive processes. Conversational Data Intelligence revolutionizes this landscape by making data interaction as simple as having a conversation. By fusing cutting-edge conversational AI with data analytics, CDI makes data-driven decision-making accessible to everyone, and eliminates the need for specialized data science skills. As we harness this powerful technology it's critical to be vigilant about the challenges and risks to ensure that implementation is secure and that the insights generated are reliable and actionable. The more we develop custom CDI solutions for our partners, the more we are amazed at the transformative potential of this technology. Contact the Presence team if you are curious about integrating CDI driven insights into your operations.

### About Presence

Presence is a respected digital product consultancy and end-to-end development firm known for creating market-leading technology solutions.

Reach out [here](#), or email [contact@presencepg.com](mailto:contact@presencepg.com)