

CluedIn to Data Management

The Company

CluedIn was founded in 2015. It is aimed at tackling problems in data management associated with data preparation and data integration, using graph technology for flexible data modelling and AI-driven data quality. The company has built up a portfolio of customers across the globe, including Sega, Nykredit, Saxo Bank, Evercore and Plains Midstream Canada. The company has around 100 staff. It is privately held, with investors including Dawn Capital and Nordic Makers.

CluedIn is headquartered in Copenhagen, Denmark.

Visit their website: www.cluedin.com

Company Name Global Presence & Financial Strength

>\$10M	>10	>50+	>150
Annual revenue	Countries with customers	Employees	Corporate customers

The Market

CluedIn aims at business rather than technical customers. It has capabilities in master data management, data quality and even data governance and metadata management. It has extensive customer traction in the insurance and finance industries, as well as in consumer-packaged goods and retail.

Overview

CluedIn is a modern data management platform that is based on graph database technology and makes heavy use of AI for automation. It has an AI agent that suggests data quality rules, and machine learning is used for record merging and matching.

Key Features:

Feature 01	Feature 02	Feature 03	Feature 04
The underlying graph database used by CluedIn makes it easy to analyse data relationships.	Continuous data monitoring by AI agents improves data quality continuously instead of relying on static rules.	Handles unstructured data like documents and images as well as traditional structured data.	Native integration with the Microsoft stack, such as with Purview.

The Technology

CluedIn is a cloud native data management platform based on a graph database, an OEM of Neo4J. It is able to ingest data from a wide variety of systems including ERP and CRM systems, but also unstructured data, either in batch or by streaming real time. It has integration with Azure Data Factory. It tags and store this data in a graph model rather than a traditional fixed schema. It is able to infer entities like customers and products, and attributes of these, from the ingested data, and work out the relationships between these entities. This flexible schema avoids the need for up-front modelling, and can easily adapt to change. CluedIn has built in data quality capabilities including data cleansing and entity matching. This means that it can

deduplicate data, produce validation rules, match records using machine learning and enrich that data where necessary. The technology builds survivorship rules to handle situations where duplicate records exist, prioritising more trustworthy or fresher records over older records or less trusted source systems. The product has data governance features including data lineage, data steward workflows, approval processes and audit processes. Natural language queries are provided along with AI assisted rule creation and data cleansing suggestions. It can derive entities even from unstructured documents, and has a knowledge graph feature to display a visual map of data relationships.

Competitors

CluedIn operates in both the master data management (MDM) market and the data quality market. In the former, it competes with products like Profisee, Reltio, Ataccama, Informatica and SAP MDG. In the latter, it competes with general data management platforms with data quality elements such as Informatica and Ataccama, as well as with more specialist data quality tools such as Precisely and SAS. Although it has some data observability features, it does not really compete with specialist tools in that area, such as Monte Carlo, Anomalo and Soda.

Summary

CluedIn is a modern master data management and data quality platform, with a wide range of functionality that extends into data governance. Its graph database underpinning means that it is strong at handling complex data relationships, and it uses AI and machine learning to infer relationships and suggest data quality rules rather than needing hard-coded schemas. The product has a good track record with some demanding customers. It should be considered as an alternative to more traditional MDM and data quality platforms.

Market Differentiation:

- Graph-native database
- Agentic automation model
- Tight Microsoft integration
- Can handle unstructured as well as structured data.

“We chose CluedIn for its powerful AI capabilities and seamless Azure-native integrations.”

Felix Baker Head of Data Services Sega

About The Information Difference

At the Information Difference www.informationdifference.com, we offer in-depth analysis of the data management industry, analysis of the main vendors, assessments of the marketplace and white papers discussing key issues and best practices. If you are contemplating a data management project, we can advise you on strategy, vendor selection and best practice.