ORACLE®
Adaptive Intelligence

Oracle DataFox Cloud Service:
How Is Data Sourced?

Oracle DataFox Cloud Service is a company intelligence platform with a modern data engine that uses AI to automate data collection. It leverages a combination of natural language processing (NLP), machine learning (ML), and human-in-the-loop techniques to scan the web and create trusted B2B company data and signals.

WHY DATA SOURCES ARE IMPORTANT

Throughout the history of databases—and Oracle has been there since the beginning—many methods for sourcing data have been discovered and employed. Understanding how data is sourced is important for ensuring accuracy, compliance with legal requirements, and ethical best practices. With Oracle DataFox Cloud Service, Oracle is committed to providing the highest quality data, sourced the right way, to ensure the success of its customers' data-driven initiatives.

MULTI-LEVEL DATA COLLECTION METHODOLOGY

This brief explores the various methods Oracle DataFox Cloud Service employs to source the highest quality data for its customers. It addresses the problem of scale versus accuracy and explains how Oracle DataFox Cloud Service has combined multiple sourcing methods to deliver both elements.
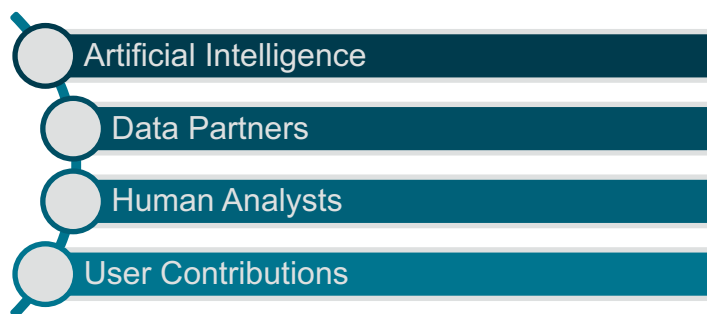
Data Snapshot

- 2.2 million new, verified companies added per year
- 21 million signals processed every year
- More than 240 million data points structured by NLP
- More than 100 human analysts to verify critical data points

- Artificial Intelligence
- Data Partners
- Human Analysts
- User Contributions

Figure 1. Using AI, data partners, human analysts, and user contributions ensures high-quality data.

ARTIFICIAL INTELLIGENCE

Oracle DataFox Cloud Service uses multiple AI techniques in its data sourcing.

Natural Language Processing-Based Data Creation

Oracle DataFox Cloud Service uses NLP to parse the public web for relevant information. It processes free-form text from unstructured content—such as news articles, press releases, government filings, social media, job listing, blog posts, and more—and incorporates it into the database.

ORACLE®

Machine Learning-Based Anomaly Detection

ML surfaces anomalies and then sends the issues to a human analyst team to verify accuracy. Oracle DataFox Cloud Service has used human auditors for years, both to verify data accuracy and provide more training data to feed back into the ML models.

Machine Learning-Based Signal Detection

With proprietary trained models, a signal is narrowed down into 68 pre-defined categories such as an executive leadership change, IPO, funding round, acquisition, new product launch, and more. This enables the system to deliver signals in real time. The human analyst team also reviews many critical signals to ensure accuracy and continued improvement of model results.

Artificial Intelligence-Based Matching and Deduplication Engine

Trained models pick out identifiable companies from articles or snippets of text across the web and determine the correct company entity to match that data to, even if company names are ambiguous. All data points are deduplicated and matched to a legitimate, verified company.
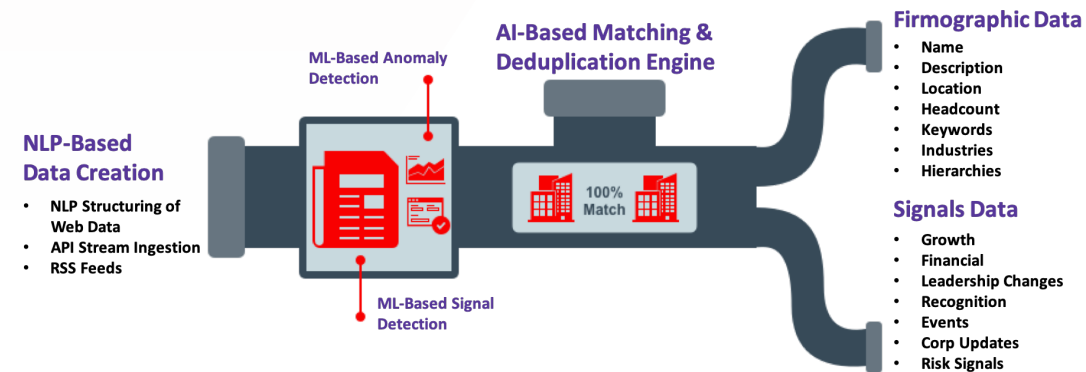

Figure 2. Oracle DataFox Cloud Service uses NLP, ML, and AI to create high-quality data for your CRM system.

DATA PARTNERS

If a data vendor specializes in a certain data type, Oracle often partners with that vendor rather than build another, redundant sourcing mechanism. These partnerships with data providers ensure the greatest possible quality of data. Following is some insight into a few of Oracle's data partners:

CrunchBase for Funding Data

Although it provides a wide range of data points, CrunchBase specializes in data on financial events such as funding rounds and investors. Oracle partners with CrunchBase because it has built a powerful crowdsourcing platform for funding data. Based on research from Oracle's data science team, CrunchBase funding data is highly accurate and is consistently updated as soon as funding rounds are announced, providing crucial data for customers that leverage this data for funding-related engagements.

Intrinio for Public Company Data

Intrinio is a critical partner for rich, dynamic data on public companies. Oracle partners with Intrinio because it has developed sophisticated mechanisms for sourcing data from public filings, maintaining more than 300 financial data feeds.

Key Benefits

- Gain a better understanding of your customers and enhance sales and marketing efficiency.

- Expand your target market by identifying new accounts to sell to and increase revenue with a larger prospect base.

- Create a faster path to revenue by focusing on the best target accounts and opportunities.

- Enable sellers to have meaningful conversations without wasting time on manual research.

- Optimize campaigns by prioritizing highly qualified prospects, not just leads that are the most active.

HUMAN ANALYSTS

Verified Data Workflows

Oracle's data verification team works on verifying and updating core data points on company profiles, particularly data points used for customers' key segmentation, territory assignment, and routing rules. Examples include verifying or updating core data points on a profile, including headcount values and location data as well as critical signals such as acquisitions and IPOs.

Data Anomaly Resolution

Data anomalies detected by AI workflows are routed to the analyst team for resolution. Examples include companies with conflicting data points, such as large headcount number and a small revenue number or a website URL ending in .au with location not in Australia.

User Contribution Implementation

The analyst team also resolves inbound customer-facing requests from the user contribution tools within the Oracle DataFox Cloud Service user interface.

USER CONTRIBUTIONS

Every single one of the millions of company profiles in the Oracle DataFox Cloud Service database contains an option to suggest an edit to a data point. These suggestions are sent to our human analyst team to verify and correct. Since Oracle DataFox Cloud Service is used daily by thousands of sales, marketing, operations, finance, and innovation leaders across more than 300 customers, user contributions create a powerful feedback loop that ensures the data is as accurate as possible.

CONNECT WITH US

Call +1.800.ORACLE1 or visit cloud.oracle.com.
Outside North America, find your local office at oracle.com/contact.

B blogs.oracle.com/oracle   f facebook.com/oracle   🐦 twitter.com/oracle

Integrated Cloud Applications & Platform Services

Oracle is committed to developing practices and products that help protect the environment

ORACLE®