EVERSTRING

The Account-Matching Innovation That Is Transforming Data For Business

Experts Have Developed The Ultimate Machine-Learning Reference Source For B2B Information

TABLE OF CONTENTS

INTRODUCTION

Unreliable data is the biggest challenge facing go-to-market and operations teams. According to SiriusDecisions, account-based marketing (ABM) departments cited 'lack of account and contact insights', and 'data usability issues' as their top hurdles, second only to 'insufficient budget'. Most notably, data matching causes significant challenges when entities must merge together to avoid duplicates while guaranteeing accuracy and coverage.

In data science, a 'match' is only as good as the reference set being used to render it. **If a match gets recorded against an inaccurate reference set, then major gaps in business insights arise.**

Data vendors have historically matched accounts using a traditional approach called 'fuzzy matching'. However, there is a new innovation powered by machine-learning and natural language processing technology that is completely revolutionizing the way businesses organize and leverage data across the company.

Below are details about this critical data management topic, including definitions, answers to frequently asked questions, expert commentary on the particularly unique challenges of B2B account matching, and the innovation changing the world of B2B data.

AN OVERVIEW OF DATA MATCHING

Data matching refers to the effort involved in identifying, mapping and updating two or more sources of information together, resulting in a single entity with an updated view of information.

In an organization's CRM, typically an account record will display a variety of information that is leveraged across all business units, to make strategic and tactical decisions. Internal data sources include transactional points like finance, credit, web forms, a marketing automation system, and Salesforce, whereas external data sources include EverString, ZoomInfo, D&B, Circle Back, and others.

To understand account matching, here is a common example:

1. The database has an existing account record titled "ACME, Corporation".
2. An incoming data stream has new information relevant to this account, so it attempts to locate and update the record.
3. If the new data source is labeled slightly different, (such as The ACME Corp instead of ACME, Corporation), matching ensures those entities are still linked together to avoid duplication.
4. If it does not find an existing match, it will create a new account record and fill in the fields as much as possible.

> "A match is only as good as the reference set being used to render it. If a match gets recorded against a flawed reference set, significant gaps in business insights arise."

THE CRITICALLY-IMPORTANT REFERENCE SET

Any matching process requires a 'reference set'. The reference set is like the DNA blueprint for a database. It represents the master source that any incoming data is verified against.

Consider the difference between fingerprint matching technology and DNA testing. The former method is useful at helping identify someone, but the latter provides an unsurpassed, submicroscopic level of granularity. Just like DNA contains trace bits of code that make each human unique, the internet has trace bits of data that tell a unique story about each business, each industry, and each market.

When most data vendors discuss matching, they are using a fairly static reference set. This reference set is likely being maintained with human and machine inputs, to varying degrees of accuracy and coverage. In today's fast-paced world, business data is constantly evolving. **The best, most real-time source of information today is the world wide web.**

Traditional Approach: A no-win compromise between accuracy and coverage

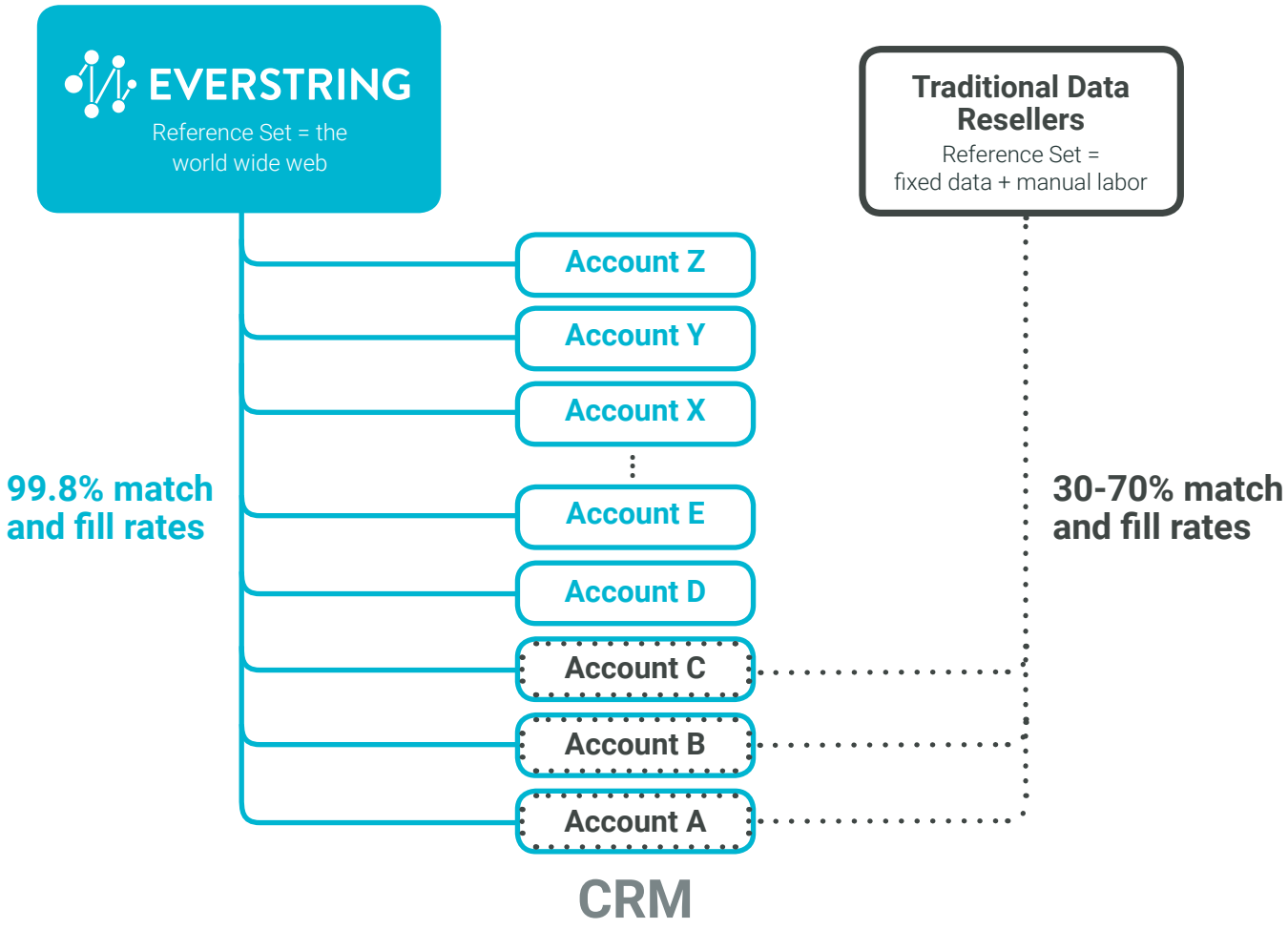
Under a traditional approach, the reference set being used is significantly limited. The source is often form fills, webinar attendance, or surveys. Although it may include millions of records, an outdated reference set could also include numerous dead or shell companies resulting in suboptimal accuracy and coverage rates.

Modern Method: The World Wide Web renders a 99.8% match-rate

Instead of relying on traditional methods, smart matching uses the ultimate reference set: the World Wide Web. By crawling and extracting bits of trace data from search engine results pages, these information bits are gathered and synthesized to produce the most accurate, up-to-date reference set from which the best data matching can occur.

> "Traditional data providers that are using an outdated reference set generally miss between 30-70% of the potential addressable market."

Figure 1: The Difference In Matching Coverage Based on Reference Set



For example, imagine a business focuses on selling to enterprise-level accounts so they have a filter set in their marketing automation system for 2,000 employees or more. Then, this week Forbes Magazine reports a massive surge in hiring at ACME, Inc., doubling its workforce from 1,500 to 3,000 employees.

- With traditional data matching, the organization will MISS this opportunity to engage ACME, Inc. because the account data is not representing an accurate employee count, reflective of present-day marketplace news.
- With modern matching techniques, the database will be enriched with this new target account, along with a multitude of other relevant business information based on your target industry, trends, news, and advanced machine-learning algorithms trained to think like executive business leaders.

FUZZY VS. SMART MATCHING

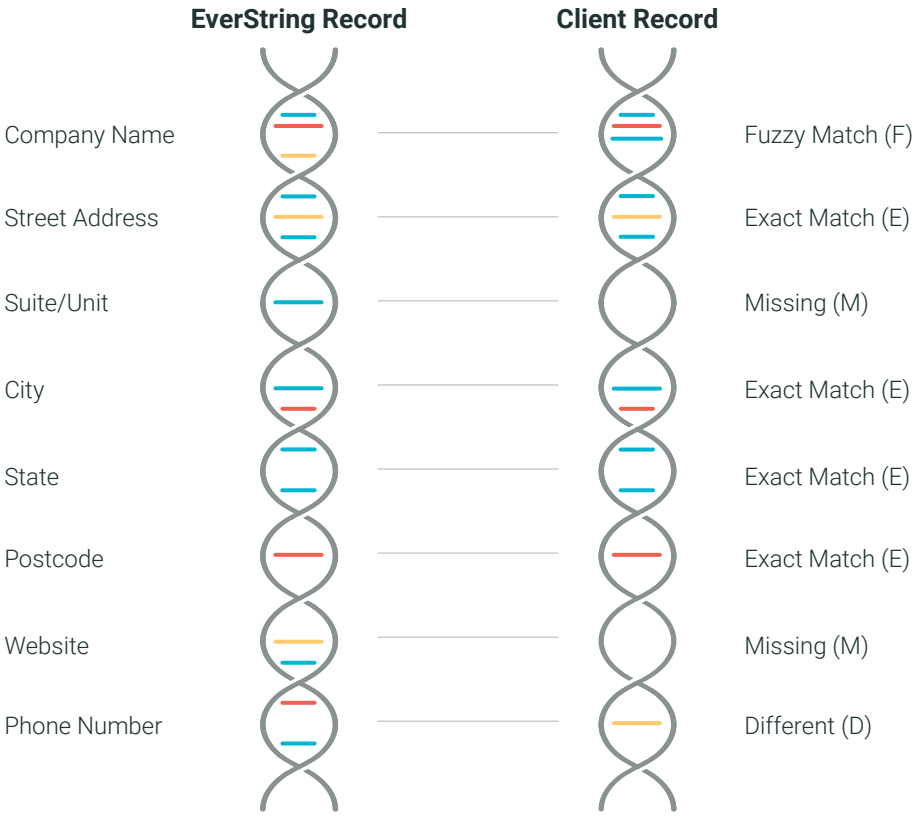
The Current State of Matching

Fuzzy matching is a computer-assisted translation technique employed to produce a data set match, when an identical pair cannot be found. In place of an exact match, fuzzy matching applies a percentage to find the closest possible entity. Fuzzy matching is not ideal, since it has an estimation factor that is inferior to what data scientists can now achieve with machine-learning advancements.

What's Coming Next With Smart Matching

The World Wide Web is full of information, both business and non-business related. With AI and machine-learning, smart matching filters what is relevant for a business and what is not and then finds patterns from the bits of data from all pages indexed on the web. Those patterns are then linked to the company entities through advanced algorithms, with a level of accuracy never seen before.

Figure 2: Illustration of Various Match Types (Fuzzy, Different, Exact, Missing)



Smart matching happens when you can teach a machine to answer the question 'What would a human being do to reach the conclusion that this data belongs to a particular company?'

THE VITAL DIFFERENCE BETWEEN DATA PROVIDERS

Traditional Data Resellers

Value-added data resellers purchase bulk data stores from various sources, enrich, and then resell the data in different segments at high fees. Along with the data, they typically offer a customer data platform (CDP) or predictive marketing tool to add value.

Although these resellers offer professional services to help teams operationalize the information, the underlying data quality problem is not solved. When data is verified using a flawed reference set, it may still register a match even though it is pairing up to previously incorrect information.

Data point triangulation does not address the ultimate problem either. Although triangulation can improve accuracy, if the reference set is erroneous, then the match will be flawed, even though it is recorded as 'complete'.

Examples of value-added data resellers include: D&B, DiscoverOrg, ZoomInfo, Experian, and Equifax.

Modern Data Sourcing

The optimal source for a reference set is one that is used and continuously refreshed by the entire world. Modern data providers leverage this up-to-date information from the World Wide Web, along with **machine-learning technology that has been taught how to model the evaluation process of leading business executives.**

Using business-oriented natural language processing (NLP), models are developed and machines are deployed to crawl, extract, and synthesize bits of data left on the web.

If a business does not have a website yet, they still have a digital presence. Smart matching technology leverages any indexed page on the world wide web to extract data even if the business has no actual company website or domain.

By the numbers, EverString covers:

- 10M companies with a website in the U.S.
- 5M companies with some digital presence in the U.S.
- 12M companies without a website, but still some digital presence in the U.S.
- 60M dead-to-shell companies in the U.S.
- 10M companies across the rest of the world

RESULTS OF SMART ACCOUNT-MATCHING

Uncover New Opportunities

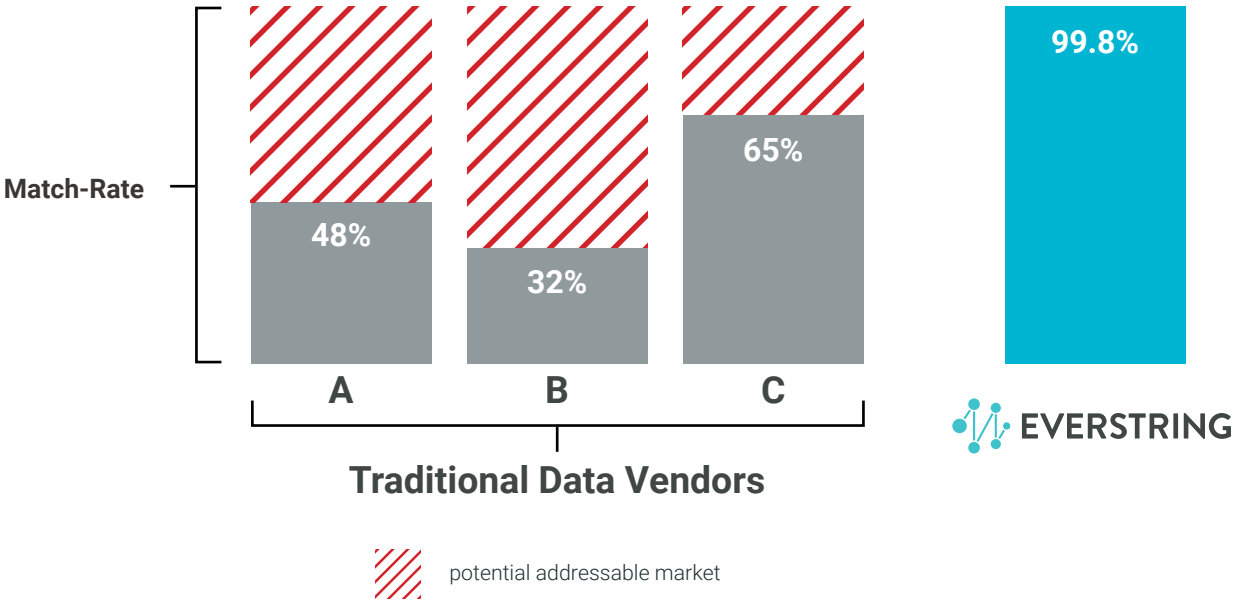
Fuzzy matching excludes a significant portion (between 30-70%) of potential high-fit accounts, meaning those potential customers will never exist in the company's ecosystem. Even if the entity does exist, no data enrichment will occur since an outdated reference set is used. In other words, current information is excluded and therefore never triggers the account into a marketing nurture or sales activation campaign.

Smart matching leverages the most current reference set, the World Wide Web to generate a **99.8% match-rate, helping businesses stay on top of new sales opportunities that otherwise would have gone unexplored.**

Unified, 360-Degree Customer View

Smart account-matching creates a comprehensive and united view of an account, including both internal and external data at any level of granularity needed. The result is a robust data foundation that includes tens of thousands of different account attributes and buyer behavior patterns, ranging the full span of the organization.

Figure 3: Traditional vs. Smart Matching Rates



IN CLOSING

Data matching plays a critical role in the health of a company database. When matching errors occur, duplicate or inaccurate data is funneled into the business, causing leaders to miss significant insights and revenue opportunities.

Smart matching technology transforms how data matching happens by using the trillions of digital traces consumers and businesses leave behind on the World Wide Web to ensure a near perfect match across a company's digital ecosystem. The result is a significantly more comprehensive and robust data foundation, full of continuously enriched attributes that drive business insights from a centralized source of truth.

FREQUENTLY ASKED QUESTIONS (FAQ)

Q: What is EverString's source of data?

A: Most data providers are reselling information from other flawed sources, carrying the faulty information onward and perpetuating the problem. These traditional data resellers use surveys, web forms, call centers and other time-consuming, error-prone human labor techniques that do not scale. The result is a trade-off between coverage and accuracy.

EverString has trained advanced, business-specialized machine-learning algorithms to mimic industry professionals. These systems crawl cached data available from the entire World Wide Web. EverString owns and maintains this database, which is continuously and automatically kept up-to-date.

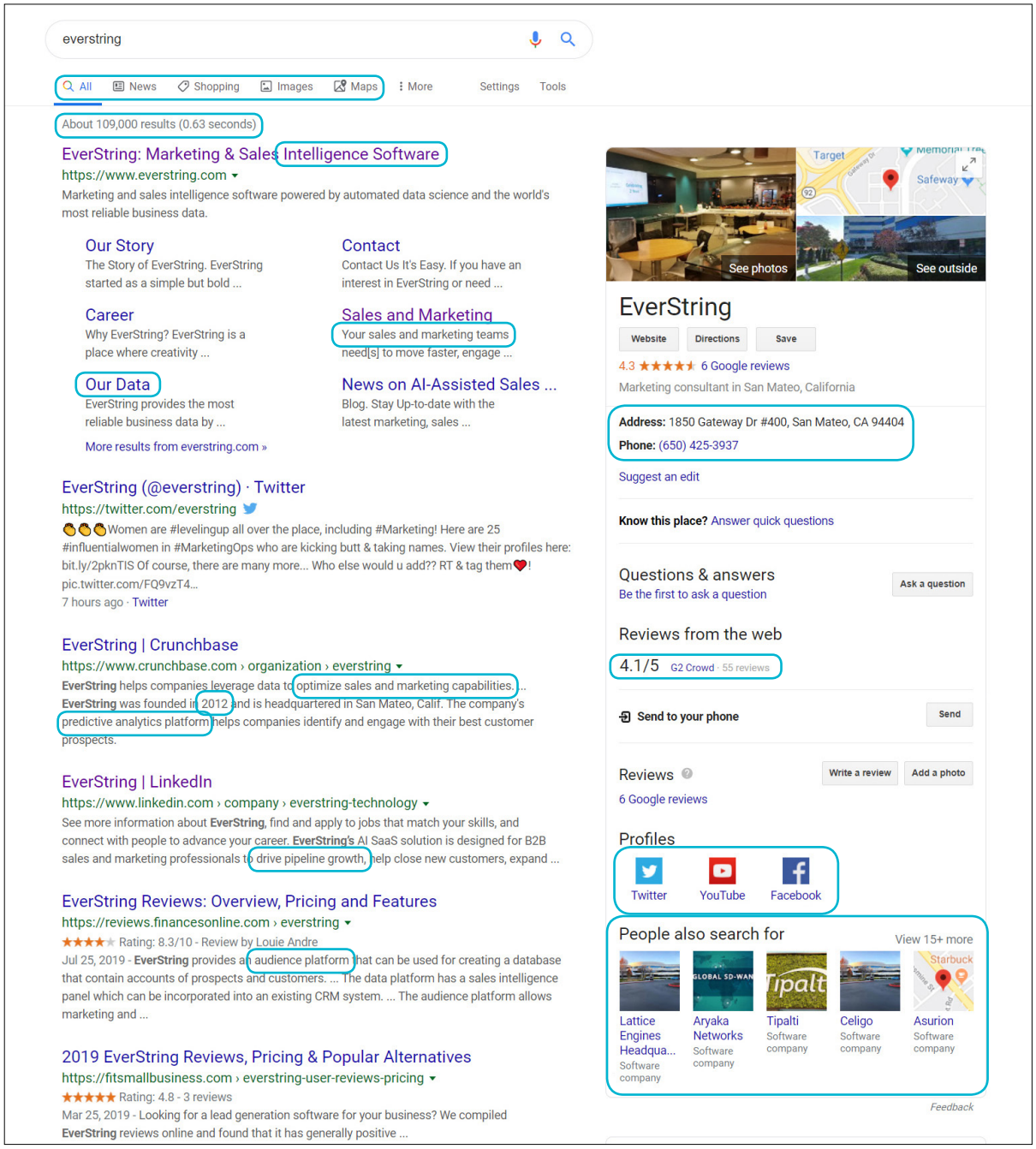
Q: What if a company does not have a website? How does information get collected?

A: Companies that do not have a website still have some digital presence available, whether it's through sites like YellowPages.com, Yelp, Google Business Listings, Google Maps, LinkedIn and others. These digital traces are indexed on the World Wide Web.

Q: Does EverString crawl directory pages like Yellow Pages, LinkedIn, Facebook?

A: No. EverString extracts information about businesses indexed by the World Wide Web. EverString uses proprietary business natural language processing techniques, a process that mimics how a business professional would think if tasked with finding information about a company or a solution. This includes but is not limited to executive-level business logic, reasoning, context, nuance, jargon and other relevant semantics.

Figure 4: A search engine results page, highlighting some digital bits of account information that are crawled and extracted.



Q: Does EverString include international data?

A: Yes, EverString serves all major English speaking countries, including the U.S., Europe, Canada and Australia.

Q: How many companies are in EverString's database? What is the coverage area?

A: With the World Wide Web as the reference set, EverString is able to provide data on **as many companies that can be searched on the internet**. Trillions and trillions of digital footprints are available on multiple locations, sites, pages, and properties across the world wide web.

Q: What is EverString's data accuracy rate?

A: Since EverString uses the broadest and most up-to-date reference set, our data provides the following accuracy metrics:

1. Nearly 100% match-rate (3x better than traditional vendors)
2. 3x better match quality than traditional vendors
3. 100% fill-rate (3x better than traditional vendors)
4. 70-90% fill accuracy on NAICS or SIC (2x better than traditional vendors)

Q: How does EverString measure data accuracy?

A: Data accuracy is measured in 4 parts: Match-rate, match accuracy, fill-rate, fill accuracy. Periodically and regularly, EverString Data Teams take a random, stratified sample and then vet that sample against known golden datasets.

Q: Does EverString use human labor to clean data?

A: No. EverString creates pristine data sets and uses those to train machines to mimic the logic, thought process, valuation measures and other nuances of top business professionals.

Q: How frequently is EverString's data refreshed or updated?

A: EverString systems are monitoring the web 24x7, and updating data monthly. Extraction happens in real-time. Behavioral data is refreshed weekly.

Data accuracy has two classifications:

1. Factual Attributes, such as phone, CEO, Industry classification. These are fairly straight-forward to capture, record and update.
2. Non-Deterministic Attributes, such as revenue and spend, require more finesse.

   EverString's smart matching algorithms have been trained to think like a top-performing business executive, and can mimic how a human would think about a company's revenue, then assign an accurate value to it, based on other factors unique to that industry.

Q: How does EverString's matching process differ from others?

A: EverString is the only data source that uses business-oriented natural language processing (NLP) technology to mimic the knowledge, logic, reasoning, and analysis of industry experts.

When assessing coverage of a data vendor, the traditional approach is to ask that vendor how many companies they can match in a given zip code area. Then, the vendor tries to produce a list of companies it has in the database, that can match the criteria. But this is a flawed method for one critically-important reason: The matching process is referencing a potentially outdated database as the master file.

Instead of zip codes, smart matching synthesizes trillions of digital signatures together, to develop a comprehensive view of an organization including the logic-based calculation of useful information such as a company's position in the marketplace compared to competitors, propensity to purchase a certain software, and many other attributes.

Consider the difference between fingerprint matching technology and DNA testing. The former method is useful at helping identify someone, but the latter provides an unsurpassed, submicroscopic level of granularity. Just like DNA contains trace bits of code that make each human unique, the internet has trace bits of data that tell a unique story about each business, each industry, and each market.

Q: How does EverString achieve a 99.8% match-rate?

A: EverString's smart matching technology produces a 99.8% match-rate because the World Wide Web is used as the reference set, ensuring the most comprehensive list of accounts are matched and further scored. Machine algorithms trained to think like business leaders, continuously synthesize the information with expert semantics, nuances, logic, reasoning, education and experience.

Other data providers base a match on a static database that requires constant, manual scrubbing and limited machine resources, resulting in up to 70% of potential addressable accounts being excluded.

ADDITIONAL RESOURCES

The Enterprise Guide To The B2B Data Revolution
www.everstring.com/resources/the-enterprise-guide-to-the-b2b-data-revolution/

Buyers Guide - Questions To Ask Data Provider
www.everstring.com/resources/buyers-guide-questions-to-ask-data-providers/

Hear From EverString Customers:
www.everstring.com/customers/