verato®

Verato Referential Matching®

An inside look at what makes Verato® the industry-leading choice for identity matching in healthcare



Executive summary

Verato® represents a quantum leap forward in the ability to match person records together across disparate sources. Verato has pioneered a new matching technology, Verato Referential Matching®, that provides the following benefits:

- **More complete and accurate data:** Verato makes more matches automatically, by overcoming minor data discrepancies (nicknames, typos, phonetic similarities, etc.) and major data discrepancies (missing data, out-of-date data).
- **Multi-domain support:** Verato can match patient, provider, and consumer data with unprecedented accuracy thanks to reference databases that include person and provider data.
- **50 to 75 percent less data stewardship:** Verato requires less manual review and resolution of "potential matches" by matching more records automatically that other technologies would only be able to flag for review.
- **Faster time-to-value:** Verato has a low total cost of ownership and does not require extensive algorithm tuning like other solutions — saving time, complexity, and resources during implementation.

Contents

Executive summary 2

The challenges with identity matching 3

How Verato matches 5

- Using sophisticated probabilistic and fuzzy matching algorithms
- Using reference data

The Verato reference database 7

Features of Verato Referential Matching® 9

- Twin detection / familial detection
- Overlay detection
- Dummy / default value detection
- Uniqueness detection
- Matching pediatric records

How Verato differs from the competition 11

- Reference data is native to the solution
- Reference data is integral to the matching process

Benefits of Verato Referential Matching® 13

Verato Referential Matching® powers Verato MDM Cloud™ 14

Proof of Verato Referential Matching accuracy 15

Verato's matching is independently verified to be 24% more accurate than probabilistic matching (JAMIA 2022). Verato Referential Matching® is trusted by 100+ healthcare organizations, including some of the largest providers, payers, HIEs, technology vendors, and governmental organizations in the country. We deliver highly accurate identity matching that powers customers' enterprise initiatives.

This document will examine the challenges with identity matching, how Verato overcomes those challenges with Verato Referential Matching® technology, and the benefits of this approach.

The challenges with identity matching

The fundamental challenge with matching is the quality and completeness of the underlying data. In fact, over 30% of demographic data used to identify and match records is out-of-date, incorrect, or incomplete:

18%
of data is out-of-date



People move and change their addresses, get married and change their last names, get new phone numbers, and periodically may use different email addresses.

25%
of the population has ambiguity in their demographic data



People with similar data such as twin siblings, families with generational names (junior/senior suffix), or different cultural naming conventions (two last names that may be mistaken for a middle and last name).

6%
of data has errors



Spelling errors, transcription errors, homonym errors — for example, names sounding alike but having different spellings like Aaron and Erin — or simply hitting the wrong key on the keyboard.

5%
of data is missing



Records are often missing key identifying fields like birth date or address. Some data sources, like labs, are particularly notorious for only having a few demographic attributes to identify a record.

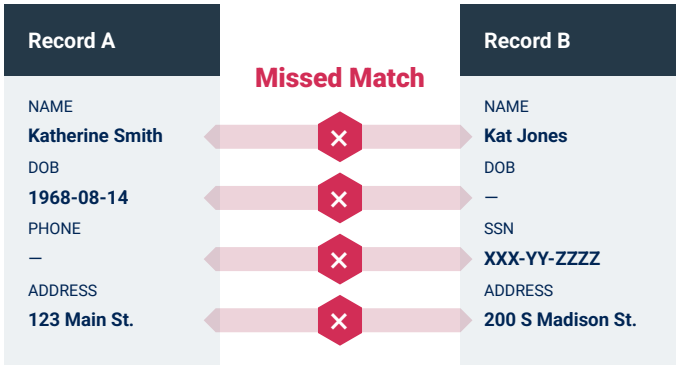
These data challenges are getting worse as data is collected from more and more disparate sources, with organizations having less and less governance over that data.

As organizations integrate with new data systems, bring in third-party datasets for analytics, allow more self-service options, and take advantage of new CRM and consumer-focused digital systems, the underlying data they're relying on to identify a person and match their records is drastically less reliable.

What does this mean for identity-matching technologies? Simply put, the algorithms that legacy products use to match records together have reached their mathematical limits. These algorithms are only as good as the underlying data — and the underlying data is inevitably getting harder and harder to match.

For example, no legacy matching algorithm can match two records when one is a person's old record containing a maiden name, old address, and birth date, and another is a newer record with the person's married name, new address, and SSN.

Legacy matching engine





How Verato matches

To match person records, Verato uses a powerful combination of sophisticated probabilistic (or “fuzzy”) matching algorithms and additional reference data. This combination of algorithms and data allows Verato to automatically make more matches than other approaches, while generating significantly fewer “potential matches” that need to be manually reviewed and resolved.

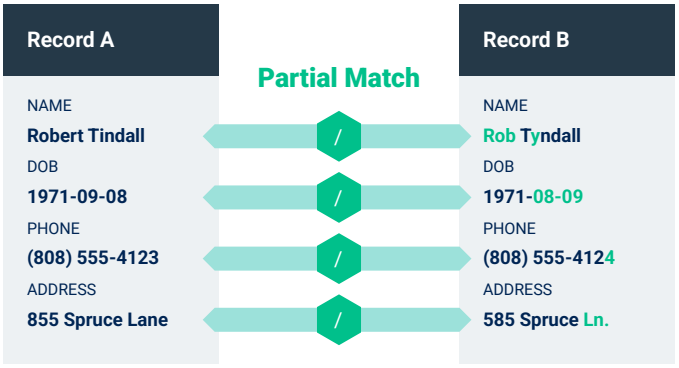
Using sophisticated probabilistic and fuzzy matching algorithms

Advanced probabilistic and fuzzy matching algorithms allow Verato to overcome minor data discrepancies between records.

For example, Verato can overcome things like:

- Nicknames
- Phonetic similarities
- Name transpositions (like transposing the First Name with the Last Name, or including the Middle Name in the Last Name field)
- Number transpositions (like in an address or in a phone number)
- Month/day transpositions in the birth date
- Typos (like Levenshtein edit distances of 1 or 2)

Examples of probabilistic approaches



- First name: common nickname
- Last name: phonetic match
- Street number: digits transposed
- Street name: abbreviation of Lane
- Birth date: month and day transposed
- Phone number: Levenshtein edit distance = 1

Using reference data

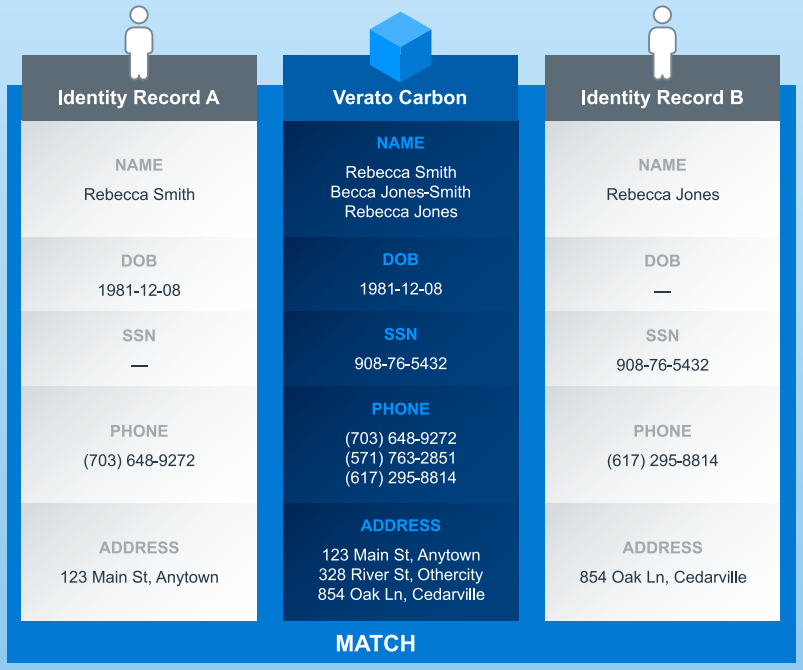
Verato also incorporates its native reference data as an integral part of every matching determination to overcome major data discrepancies between records. For example, Verato can handle things like:

- **Old data vs. new data:** For instance, a record with a person's maiden name and old address being compared to a record with that person's married name and new address. Verato collects and maintains both historic and current attribute instances for all individual identities within its reference data.
- **Missing data:** For instance, a record with only a person's name and birth date being compared to a record with just that person's name, SSN, and address. The Verato reference database includes all of these different attributes tied to an individual reference identity.
- **Provider data:** For instance, a record with an incomplete or missing NPI number. Verato contains other identifiers, such as state license IDs, DEA ID, and other Medicare/Medicaid IDs to match practitioner records.

Verato Referential Matching®

During matching, Verato not only compares records directly to each other to make a match decision, but Verato also compares each record to the Verato reference database. If records match to the same identity in the Verato reference database, then Verato knows that they match to each other.

Because each identity in the Verato reference database contains a complete profile of demographic data for that individual, Verato is using a larger target during the matching process — this helps Verato match records even if they look very different from each other, because those records both match to the same reference identity.



Verato understands how and when to use the reference data during matching — including when not to use reference identities as part of the matching process.

For example, if an incoming record matches to more than one reference identity, Verato understands that this ambiguity means that the reference data should not be used as part of the matching process. Or, in some very rare cases, two incoming records may match to the same reference identity, but fit a troublesome data pattern when compared directly to each other — for example, they may fit a common data pattern for twins. In such cases, Verato understands that these records should not be matched despite both records matching to the same reference identity.



The Verato reference database

Verato Carbon®, the Verato reference database, is a comprehensive and continuously updated database managed by Verato and containing over 300 million identities spanning the U.S. adult population as well as 7 million practitioner identities and provider organization records. Each identity contains a complete profile of demographic data attributes assembled from across the different reference data sources. These demographic data attributes (like names, addresses, birth dates, and phone numbers) span a 30-year history, and Verato receives over 60 million updates per month.

Verato sources the reference data from multiple commercially available, trusted, non-healthcare data sources. Importantly, the Verato reference database is sourced from industries where a person has a vested interest in providing their correct information — like the credit and banking industry, the utilities and telecommunications industry, and the government and legal industry. The reference data sources are also orthogonal — meaning they are independent from each other with only one source from each industry, to protect against data poisoning from intra-industry data sharing.

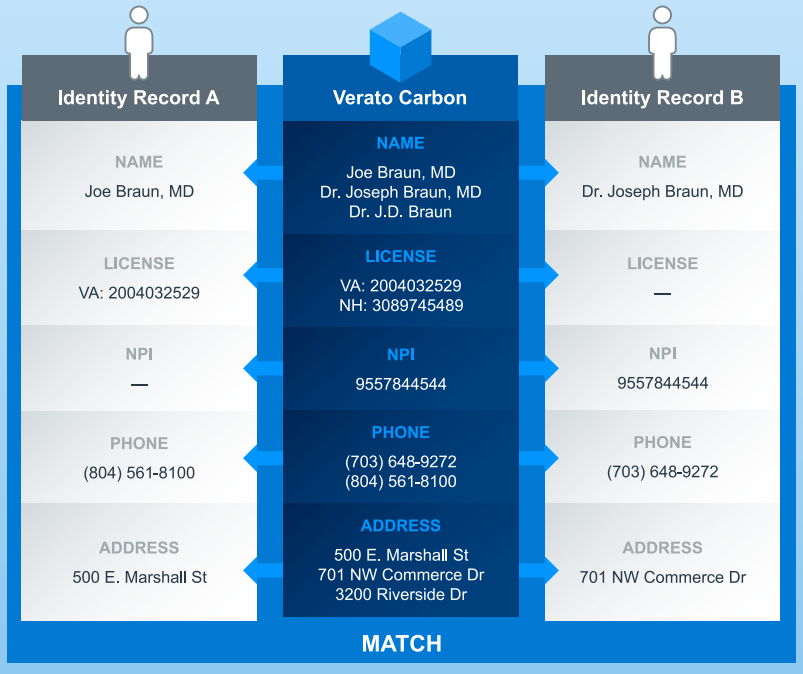
Verato manages and curates this reference data specifically for the purposes of identity matching. In other words, these billions of demographic attributes from across the different reference data sources have been poured into the Verato reference database, and the Verato data science team has invested hundreds of thousands of person-hours to managing, monitoring, and curating the data with the sole purpose of improving matching decisions.

Verato is always thoughtful about ingesting new reference data. Only data from trusted and vetted data sources is allowed in the Verato reference database. During the ingestion of new reference data, Verato ensures that the data is of high quality and accurately attaches to the correct reference identity. This is done on a record-by-record basis, and only individual records that meet strict quality standards are added to the reference database.

Importantly, for security and privacy reasons, Verato never ingests our customers' data into the Verato reference database. There is no sharing of one customer's data with another customer.

How does Verato match providers?

Referential Matching for providers works similarly to matching patients and consumers. Records are not only compared to each other but also to Verato Carbon® containing 7 million practitioner identities and organizational provider records. The data used to identify practitioners includes demographic information such as names and phone numbers, and credentialing information such as state license IDs, NPI numbers, DEA IDs, and other Medicare and Medicaid IDs.



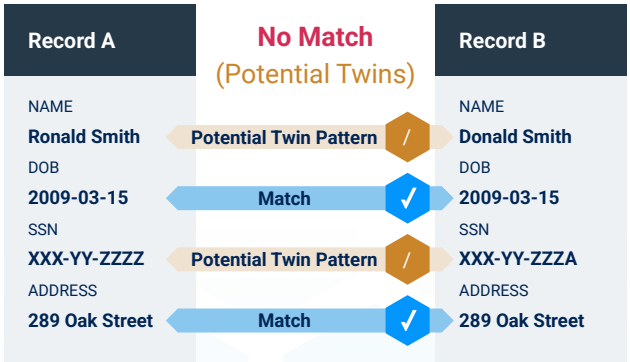
Features of Verato Referential Matching®

Powering our healthcare-focused patient identity solutions, Verato Referential Matching® was specifically designed to overcome challenges that are common and unique to healthcare, including mitigating the scenarios that most often cause "false positives" (incorrect matches) in other matching technologies.

Twin detection / familial detection

Verato looks for common data patterns for twins or other familial similarities. For example, even if records have the same last name, address, and birth date, and nearly identical first names and SSNs, Verato knows that this is a common data pattern for twins and will not match those records to each other. Similarly, if two records have the same first name, last name, and address, but the birth date is different by a generation, Verato knows that this is a common data pattern for a father/son pair and will not match those records to each other.

Verato Referential Matching®



In fact, Verato can even have further confidence that records belong to twins or to a father/son pair using its reference data — Verato may see that two similar-looking records match to two different identities in the reference database, and therefore know that those records belong to different people.

Overlay detection

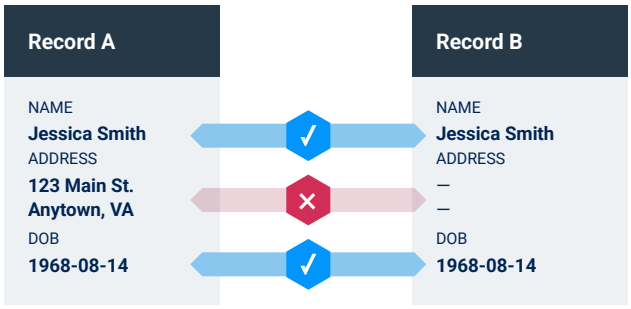
A common scenario in healthcare is an "overlay" — where one person's data is erroneously entered on a different person's record. For example, imagine that person A is in line to register at the hospital — the registration personnel creates a patient record in their EHR application for person A. Then the next person in line, person B, steps up to the registration desk. Instead of creating a new record, the registration personnel accidentally enters person B's information onto the person A record that is still on their screen. Verato can identify such an erroneous update in real-time and prevent it from occurring.

Dummy / default value detection

Another common data quality challenge in healthcare is the use of "dummy" or "default" values for certain attributes. For example, using "Baby Boy" or "Trauma Patient" as a name, using a placeholder birth date of 01/01/1900, or using a clinical site's own address by default instead of leaving the address field blank. Verato automatically understands that these types of dummy or default values should not be taken into account during the matching process. In fact, Verato has an extensive list of common dummy and default values, and customers can customize their own configuration to account for even more dummy/default values that are unique to that customer's datasets.

Uniqueness Detection

Do you think a person-matching solution should match the following two records together?



Probably not. The data is very "thin" — there are not a lot of attributes to base a decision on. And without the benefit of reference data, it would not be industry best practice to match these records with such limited information.

But what if you knew that there was only one person in the entire country named Jessica Smith born on August 14, 1968? Would that change your mind?

Verato can actually ask this type of question during the matching process — by using the Verato reference database. Verato is the only solution to utilize this powerful capability called uniqueness detection, which allows us to match records together that other solutions could never match. For example, other person-matching technologies are often especially challenged by records that contain a sparse amount of data, which is very common for sources like labs. But Verato can match these records with the help of its uniqueness detection.

Matching Pediatric Records

A common question is: can Verato match records for minors, even though there are no minors in the Verato reference database?

The answer is yes. Verato still uses sophisticated probabilistic and fuzzy matching algorithms to make matches, even if there is no reference data available for a person. Verato compares the records to each other and leverages its advanced algorithms to make matching decisions. In other words, Verato will match pediatric records through other proven methodologies — by using its probabilistic and fuzzy matching algorithms.

In addition, Verato has features that are especially helpful with pediatric records.

- **Dummy/default value detection:** Verato understands that names like "Newborn" or "Baby Boy" are dummy or default values that should not be considered during the matching process, helping prevent similar-looking records for newborns from being erroneously matched together.
- **Overlay detection:** Verato has special overlay handling specific to newborns, whose data is often updated from dummy or default values to their real information.
- **Twin detection:** Verato looks for data patterns specific to twins and prevents records from matching when that pattern is detected.

How Verato differs from the competition

Verato pioneered the concept of referential matching in healthcare, and naturally other vendors quickly took notice. A few of these other vendors partnered with third-party data providers so they could claim to do "referential matching." However, what these other vendors have accomplished is not on par with Verato for two reasons: **reference data is a native and integral part of our matching.** Other vendors' use of third-party reference data as an add-on means it is neither native nor integral to their matching processes.

The difference is stark. It's the difference between building a brand-new fighter jet from the ground up, versus pouring jet fuel into a 10-year-old sedan and claiming the sedan has "fighter jet technology."

The fact that reference data is both native and integral to the Verato matching process are key differentiators with important benefits.

Reference data is native to the solution

Verato manages the Verato reference database in house and curates it specifically for the purposes of healthcare person matching. Verato sources the data from trustworthy datasets where people have a vested interest in providing their correct data, then applies its own internal data science team to the database to optimize it for improving person-matching accuracy.

> Other vendors partner with third-party data suppliers, whose data is often collected for the purposes of selling that data or marketing to individuals — not for the critical initiatives of healthcare where a person's wellness or life can depend on the accuracy of the matching solution.

Verato has built its algorithms in tandem with the Verato reference database, so the algorithms know how best to use the data and how much to trust it for each matching decision. This allows Verato to make not just the most accurate but also the safest matching decisions.

> Since other vendors use third-party data sources, they may have to more blindly trust the quality and completeness of the third-party data, and they haven't built their algorithms to leverage the reference data intelligently.
>
> Other vendors may even require lengthy exercises from the customer to tune, configure, and adjust the tool's algorithms to effectively leverage the third-party data — a recurring challenge and resource burden that organizations don't want to invest in.

Because Verato Referential Matching® is native to the solution, Verato even exposes relevant reference data natively in its stewardship user interface, allowing data stewards to make more efficient and more accurate decisions when manually adjudicating matches — by having this additional data right at their finger tips. And, with native reference data, Verato makes more matches automatically than other solutions, meaning there are fewer matches that actually need manual adjudication.

Reference data is integral to the matching process

Verato uses the reference data for every match decision. In other words, every customer record is compared to the Verato reference database to see if there is a matching reference identity that can improve our matching decisions.

> Other vendors often use their third-party reference data for only a subset of matches — just a few that the vendor's solution has already identified as being close, but not so close it can make an automated decision. In other words, customers are not getting the benefit of the reference data every time.
>
> In fact, because other vendors rely on third-party partnerships, using reference data for every match would be cost-prohibitive, placing the burden on their customers to figure out when to use reference data.

Verato is also a cloud-native solution, which means Verato Referential Matching® can scale to millions of records and maintain performance while using the reference data for every match decision.

> Since other solutions are not cloud-native, performance and scalability with reference data becomes a large challenge. Using reference data for every matching decision would be nearly impossible for these legacy solutions because they are based on outdated technology.

Verato also uses the reference data for every step of the matching process:

- **Candidate selection:** This is where the matching solution identifies a subset of records to try matching an incoming record to. Verato uses its reference data as part of this process — so even if a record contains an old address and maiden name, Verato will still be able to find the correct candidate to match it to.
- **Matching:** This is where match comparisons are made to decide whether records belong to the same person. Verato uses its reference data as part of the match comparisons to overcome out-of-date and incomplete data.
- **Post-matching adjustments:** This is where Verato makes additional adjustments, like using uniqueness detection, or determining whether two records may be twins or other familial scenarios that shouldn't be matched.

> Other vendors only use the reference data to assist in a subset of matching judgments — so (a) they miss matches they could find if the reference data was part of their candidate selection, (b) they may make twins/ familial matches they shouldn't, and (c) they don't have the benefit of using uniqueness detection to help make matching decisions.

Benefits of Verato Referential Matching®

Verato Referential Matching® gives customers direct and significant benefits to their businesses, their ability to offer exceptional services and care, and their resource requirements for implementing, maintaining, and stewarding our solutions. Benefits include:

More matches made automatically



Verato is able to make more matching decisions automatically than other vendors by leveraging its native and integral referential matching approach.

Less data stewardship



Verato requires 50-75% less data stewardship (the manual review and resolution of matches) because Verato is able to automatically match records that other solutions would only ever be able to flag as “potential matches.”

More efficient data stewardship



Verato natively exposes relevant reference data in its stewardship user interface, allowing for more efficient and more accurate manual decisions.

No extensive tuning and configuration



Verato allows for faster implementations and easier integrations of new data sources. Other solutions require lengthy and complex tuning and configuration cycles during implementation and often require re-tuning and reconfiguration as new data sources are integrated into the solution.

Verato Referential Matching® powers Verato MDM Cloud™

Healthcare organizations have specific data management needs that generic master data management (MDM) solutions can only meet with extensive tuning. High standards for clinical data, compliance requirements, and disparate data sources make record matching challenging for generic identity solutions. That's why we built Verato MDM Cloud™. This healthcare master data management solution is purpose-built to meet the industry's complex needs without requiring extensive and repeated algorithm tuning. Powered by Verato Referential Matching and including Verato Provider Data Management™ along with patient and consumer identity, the platform manages patient, provider, and consumer data with one solution for a 360-degree view of every person across your enterprise.



Proof of Verato Referential Matching® accuracy

Verato is proven in the industry at being more accurate and resolving matches that other solutions miss. Verato has 100+ healthcare customers spanning some of the largest providers, payers, HIEs, healthcare technology vendors, and governmental organizations in the country — all of whom trust Verato to solve their matching challenges after having rigorously vetted the technology.

Consider the number of "potential matches" that some customers had before and after Verato — these customers had hundreds of thousands or millions of records that would have otherwise required manual review and resolution because their legacy matching technologies just could not overcome the data challenges so common in healthcare.

Required manual review and resolution



verato

Verato®, the identity experts for healthcare, solves the problem that drives everything else: Knowing Who Is Who®. More than 100 of healthcare's most respected brands rely on Verato for a complete, trusted 360-degree view of the people they serve. Our HITRUST-certified cloud identity platform delivers unmatched accuracy and time-to-value, so you get identity right from the start.

© 2026 Verato, Inc. All rights reserved.