healthverity®

Synchronize the Science℠

Move beyond legacy tokenization and advance the science with synchronization

Real-world data is out of sync

The pharmaceutical industry has been experiencing major breakthroughs in science, with treatments for diseases once thought insurmountable, mRNA vaccinations that saw us through the COVID pandemic and hold promise for future endeavors, and advances in gene therapy, precision medicine and more. To continue this progress; however, requires a continual source of regulatory-quality data to better understand disease progression and the real-life impact on patients.

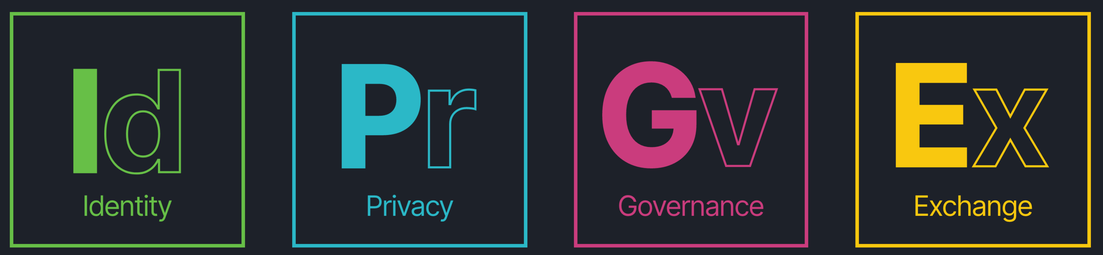
This data can be found with each step in a patient's healthcare journey that leaves behind a digital footprint. When this real-world data (RWD) is accurately woven together, it tells a story that can solve complex diagnosis and treatment challenges. And while more and more entities are making valuable RWD available, there are still many hurdles:

- Disparate and fragmented RWD makes discovery and exchange overly difficult and inefficient
- Complex and evolving regulatory environments require greater data governance
- Data aggregators tie the data to their own analytics applications, restricting transparency

Without the means to properly manage these challenges, you are left with only a partial view of the patient journey, susceptible to privacy and regulatory violations, and limited in the analyses you can conduct.

It’s time to synchronize the science

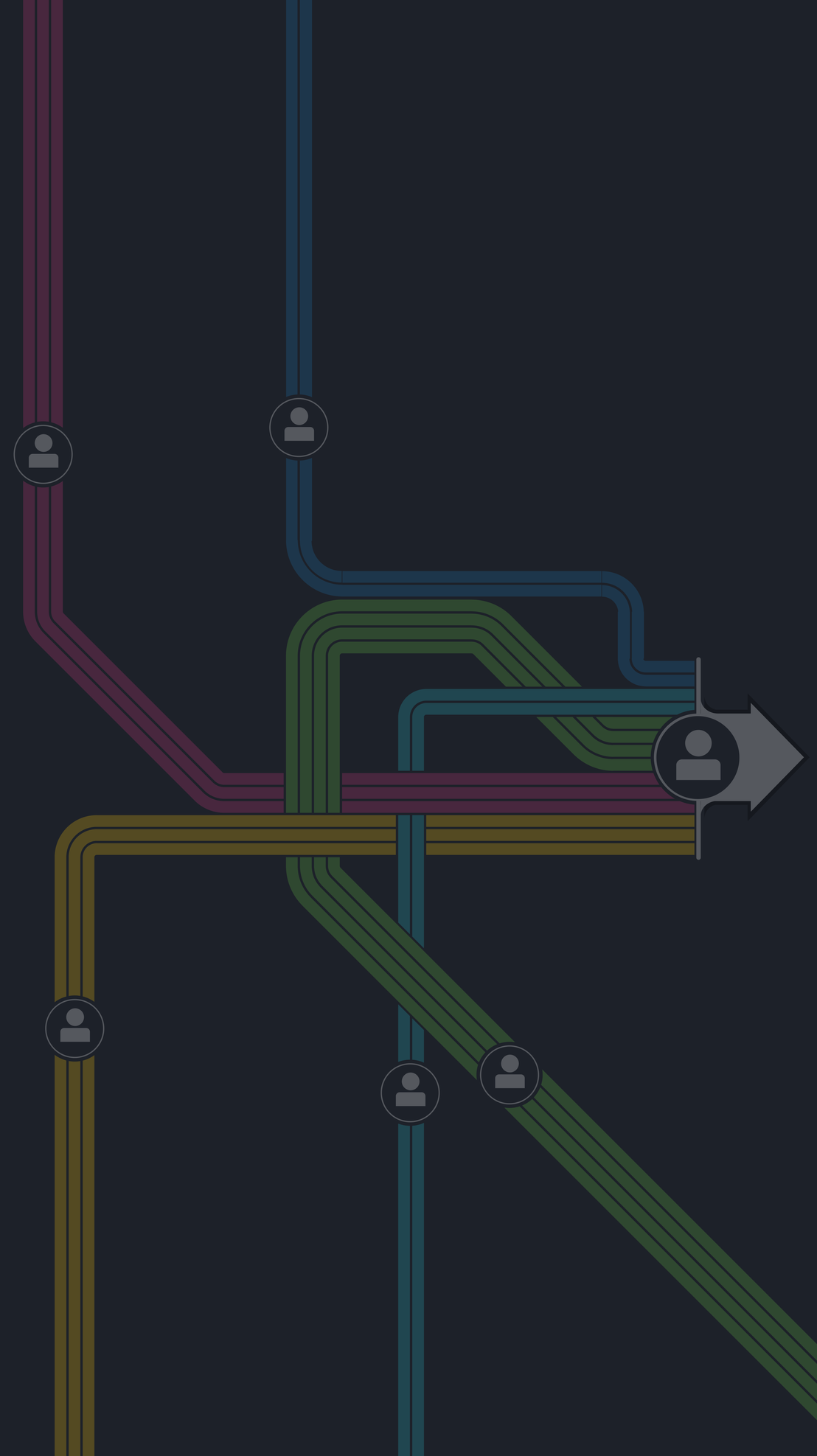
HealthVerity believed that an entirely different approach was the best way forward for the thousands of enterprises seeking to address the nation’s most critical healthcare challenges. In today’s massively fragmented data environment, patient data interoperability can no longer rely on legacy tokenization technologies, rife with high errors and inconsistent patient identities. To truly advance the science, whether it be physical science, data science or the science of your choice, requires a solution that synchronizes the foundational elements of Identity, Privacy, Governance and Exchange (IPGE). Doing so provides the ultimate source of truth in patient **Identity**, built-in data **Privacy** and HIPAA compliance, and **Governance** of data usage rights and permissions, ultimately enabling the efficient discovery and **Exchange** of fully interoperable, research-ready healthcare and consumer data.



HealthVerity pioneered this as our IPGE approach. By synchronizing these elements with our transformational technologies and the nation’s largest healthcare and consumer data ecosystem, we can weave together the most comprehensive and accurate patient journeys while governing data rights and usage permissions, providing the broadscale exchange of healthcare and consumer data that is delivered directly to your environment of choice, when, where and how you want it so you can apply analytics and applications fit for your purpose.

In this ebook, we will discuss the benefits of synchronizing each of the elements of IPGE and how this synchronized solution allows you to advance the science.

Contents



IDENTITY

The ultimate source of truth

In order to properly analyze RWD, a patient's identity must be accurately resolved across time and disparate sources. The inherent noise in RWD that comes from inadvertent errors, typos, missing fields, nicknames such as Jennifer, Jenny and Jen, changes in values (address, phone number, email or name changes), and other issues, can make matching patient records difficult. And if there are errors in matching patient records, you can make incorrect correlations or have a fragmented view of the patient journey.

There are two primary types of errors that can occur when matching patient identities: false positives and false negatives.

False negatives

With false negatives, the same person is incorrectly counted as separate individuals. This leads to a fragmented view of the patient journey, causing potential missed opportunities, overestimates in cohort populations or blind spots to key indicators.

An example would be if Elizabeth Wilson and Liz Wilson were tokenized as two separate individuals, but they're actually the same person. If this patient had a rare disease, it could lead you to think that two individuals were diagnosed with the condition when only one actually has it. Additionally, you might think that one patient is taking a prescription for the condition and the other isn't because you have a fragmented view of the patient journey for Elizabeth Wilson.

According to an internal study, legacy tokenization techniques have a false negative rate between 9% and 42%, depending on how noisy the data is, with an average error rate of 15%.

False positives

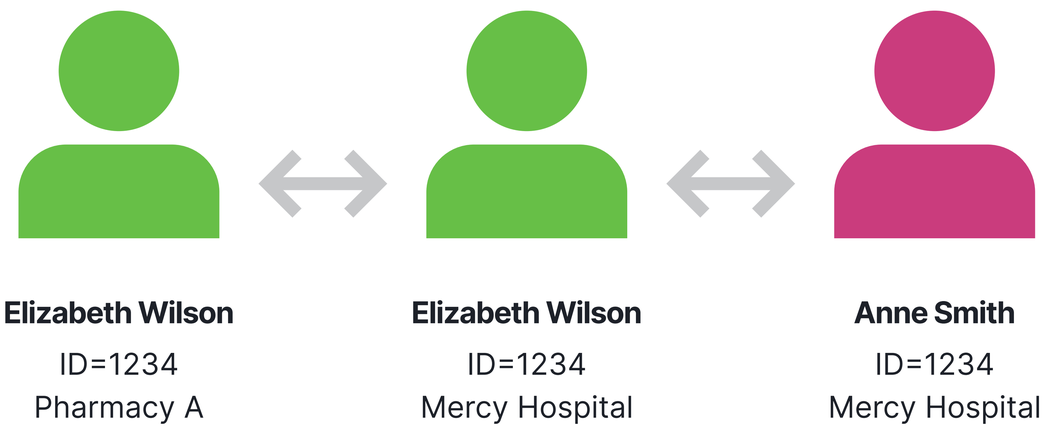
False positives are errors that occur when one or more records for a patient are incorrectly assigned to a different person. This takes away from the benefits of connecting RWD sources, providing misleading information that can cause inaccurate correlations or erroneous conclusions.

For example, if two different individuals are assigned the same ID, you could think the patient you're studying is taking a certain prescription and later develops chronic migraines. From this data, you could incorrectly conclude that the migraines are a potential side effect of the prescription; however, the person actually taking the prescription doesn't have migraines. The migraines are being experienced by a different individual, not taking the prescription.

According to an internal study, legacy tokenization techniques have an average false positive rate of 2%.

False Negatives

False Positives



A synchronized solution

To gain the accuracy that is critical in clinical research, a synchronized solution is needed that moves beyond the high error rates found in legacy tokenization. Traditional tokenization simply links patients between datasets, creating a new token every time an individual appears, and using deterministic matching, which requires an exact match, to create a link.

HealthVerity has developed advanced identity resolution techniques that synchronize unlimited patient records over time and across data sources with 10x more accuracy than a traditional token by generating a unique but persistent HealthVerity ID (HVID) that serves as a single source of truth for patient identity.

Synced beats linked℠

HealthVerity Synchronization

- Patients matched to a continuously updated referential database of over 200 billion healthcare and consumer transactions
- Probabilistic matching, which takes into account the inherent noise in RWD
- A universal HVID is applied, synchronizing patient identities over time and across data sources

Legacy Tokenization

- Patients matched dataset to dataset
- Deterministic matching, which requires an exact match
- A new token is created for each individual in a dataset, even if that person was previously identified in a different dataset

How synchronization is 10x more accurate

More than simply replacing personally identifiable information (PII) with an alphanumeric string, or token, HealthVerity synchronizes patient identities with a continuously updated referential database of over 200 billion healthcare and consumer transactions and leverages machine learning techniques to ensure the highest accuracy rate when assigning an HVID from our master patient index. This allows us to utilize probabilistic matching, which can handle the inherent noise in RWD, as opposed to legacy technologies that require an exact match.

By synchronizing identity resolution in this way, we are able to learn from the billions of transactions that we see and assign an HVID with greater confidence than legacy tokenization techniques that do not have a reference when making decisions to link patient identities or not. This knowledge, coupled with our advanced algorithms, allows us to recognize and respond to common issues seen in RWD, such as the following examples:

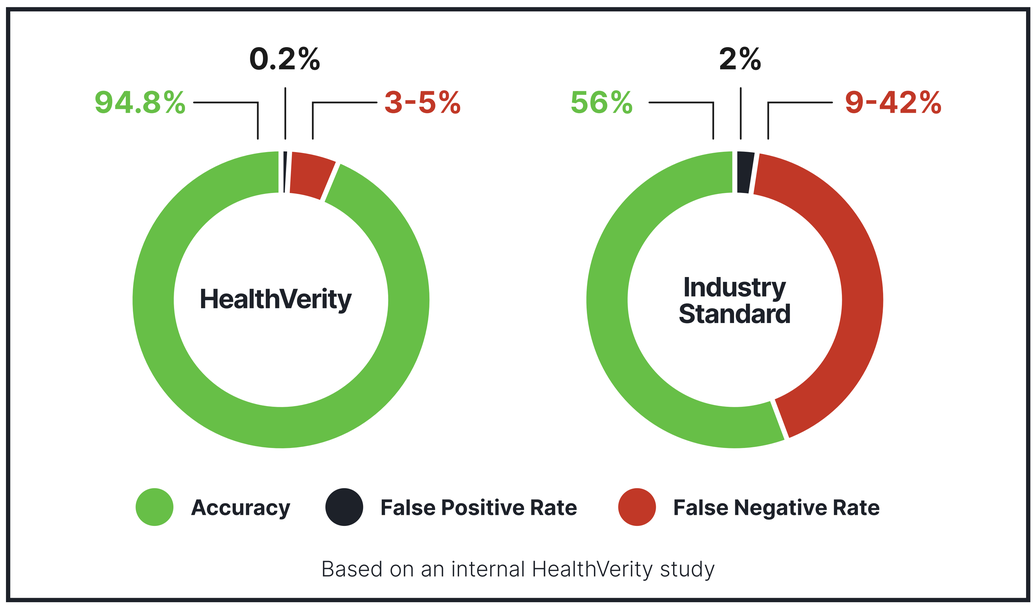
Nicknames - By tracking patient identities over time, we learn preferred nicknames, such as John Smith sometimes going by Johnny, but never Jonathan.

Juniors and seniors - Even without a date of birth and with father and son living at the same address, our referential database enables us to distinguish between the two.

Twins - With similar names, such as Jimmy and Timmy, at the same address and with the same date of birth, legacy matching techniques might see this as a typo, but our state-of-the-art technology is able to recognize these are two distinct individuals.

Messiness of data - Our system takes into account potential errors in data, such as the first name and last name in the wrong columns, swapping the birth day and month, etc.

This synchronized solution has led to an average false positive rate of just 0.2% and a false negative rate of 3% to 5%. Additionally, it allows us to seamlessly synchronize unlimited data sources, whether your first-party data or data from another third-party source.



Based on an internal HealthVerity study

PRIVACY

Built-in HIPAA compliance

Today's regulatory environment requires hypervigilance to ensure that data-driven strategies remain compliant with HIPAA and other evolving regulatory frameworks. Simply replacing PII with a token is not enough to meet these strict standards. Even if an individual dataset is deemed HIPAA compliant, issues can arise when combining data sources, such as different fields or privacy approaches, opening the potential for re-identification and HIPAA violations.

The HIPAA Privacy Rule established two standards to deem that data is appropriately de-identified and, thus, privacy protected: Safe Harbor and Expert Determination.

Safe Harbor

With Safe Harbor, a data provider must remove 18 PII and protected health information (PHI) related data fields, which limits the ability to link patient identity across disparate datasets and eliminates certain data elements that are critical for analysis.

Expert Determination

Expert Determination is where a person with appropriate knowledge and experience with generally accepted statistical and scientific principles and methods conducts a HIPAA review to determine that the risk of re-identification is very small, documenting the methods and results that justify such a determination.

A synchronized solution

A synchronized solution is needed to ensure privacy laws are adhered to and maintain timeliness with data research. HealthVerity has synchronized HIPAA compliance across our entire data ecosystem by achieving Expert Determination from an independent, third-party examiner, Dr. Brad Malin, accenture professor of Biomedical Informatics, Biostatistics and Computer Science at Vanderbilt University, as opposed to using an internal, company-controlled reviewer. This means that any data licensed from HealthVerity is research ready from day one. Additionally, this HIPAA certification is applied to our clients' proprietary data when we synchronize it with data sources from the nation's largest healthcare and consumer data ecosystem.

Our unique expertise in privacy techniques enables HealthVerity to deliver novel combinations of data and manage complex patient workflows in a privacy-protected way. By synchronizing HIPAA compliance across our data ecosystem, consisting of more than 200 billion healthcare and consumer transactions, we are able to provide our clients fully interoperable, HIPAA-compliant and research-ready data without any burden or risk being placed on them.

GOVERNANCE

Managing data rights, usage and permissions

There are even more hurdles to working with RWD than ensuring accuracy and privacy. You also have to manage the multitude of data contracts, determining usage rights and who has the rights to view the data, particularly when working with both identifiable and de-identified data. Additionally, you have to manage patient permissions at each stage of the product life cycle.

Interacting with patients requires tracking and adhering to a wide variety of permissions that are often collected and stored in silos, and sometimes need to be shared with and verified by partners. If all of this is not properly managed, violations of these permissions can have significant repercussions at the federal and state level, and damage the relationship with the patient.

When working with electronic patient data, life sciences organizations must comply with 21 CFR Part 11, which was established by the FDA to ensure the authenticity, integrity and confidentiality of electronic records and signatures.

A synchronized solution

Managing all of these data rights, usages and permissions can be overwhelming and without a synchronized solution, that burden falls on you. HealthVerity partners with the more than 75 data providers in our ecosystem, allowing us to synchronize all aspects of contracting, licensing and procurement for you, providing a single contract for multiple data sources.

Additionally, our leading technologies efficiently consolidate, track and manage the wide array of patient permissions, including opt-ins, opt-outs and expirations, across individual clinical trials and various brands, providing a single source of truth enterprise wide that is 21 CFR Part 11 compliant. We synchronize multiple consent touchpoints, such as websites, e-Consent forms and external partners, working seamlessly with your critical workflows to enable different data strategies across the organization while ensuring the right data flows to the right teams at the right time.

EXCHANGE

Discover fully interoperable, research-ready RWD

All of this effort is to ultimately source RWD for your enterprise-wide research projects, but just finding the data sources you need and having them delivered where, when and how you want them can be difficult. Following are some of the obstacles to sourcing RWD:

- Obtaining a variety of data sources and types
- Attaining interoperability amongst multiple data sources, including your proprietary data
- Normalizing the data to a common model for analysis
- Maintaining privacy while combining data sources
- Determining if patients in the dataset meet your specific criteria
- Ascertaining that RWD exists for patients in your proprietary data
- Establishing data provenance for regulatory submissions
- Receiving transaction-level data to apply fit-for-purpose applications and analytics
- Having the data delivered directly to you or a partner in a timely manner

A synchronized solution

A synchronized solution enables the efficient discovery and exchange of RWD. By synchronizing identity resolution and privacy across the nation's largest healthcare and consumer data ecosystem, all data from HealthVerity is already fully interoperable, normalized and research ready.

Our data partners trust the synchronized IPGE approach and, thus, trust us to ingest and host their data. This enables full transparency and the ability to share the provenance of the data, as well as provide the HealthVerity Marketplace™ platform where you can build custom cohorts based on expansive healthcare criteria, such as diagnosis, drug, procedure, lab test or biomarker, and see where there is overlap between datasets in real time. You can also overlay your proprietary data across our universe of RWD to determine gaps and overlaps in the patient journey.

Once you've discovered patients of interest, our synchronized solution allows you to easily license only the data you need from multiple sources with a single contract and have the transaction-level, fully interoperable, research-ready data delivered directly to the environment of your choice for immediate use, accelerating your research.

10 Benefits of synchronization

Synchronizing Identity, Privacy, Governance and Exchange with transformational technologies and the nation's largest healthcare and data ecosystem provides:

01.

The most accurate identity resolution

02.

Assurance of certified HIPAA compliance

03.

Fully integrated privacy-compliant data, including your primary data

04.

Efficient governance of data rights and permissions

05.

Data contracting, licensing and procurement handled for you

06.

The ability to exchange broadscale healthcare and consumer data with a single contract

07.

The option to license only the data you need

08.

Data provenance, transparency and normalization

09.

Data delivered to your environment of choice when, where and how you want it

10.

Delivery of research-ready data in a fraction of the time as legacy technologies

This synchronization allows you to advance the science.

HealthVerity synchronizes transformational technologies with the nation's largest healthcare and consumer data ecosystem to power previously unattainable outcomes and fundamentally advance the science. We offer a comprehensive, yet flexible approach, based on the foundational elements of Identity, Privacy, Governance and Exchange (IPGE), that synchronizes unparalleled Identity management with built-in Privacy compliance and Governance, providing the ability to discover and Exchange a near limitless combination of data at a record pace. Together with our partners in life sciences, government and insurance, we are Synchronizing the Science.

TO LEARN MORE, VISIT

healthverity.com