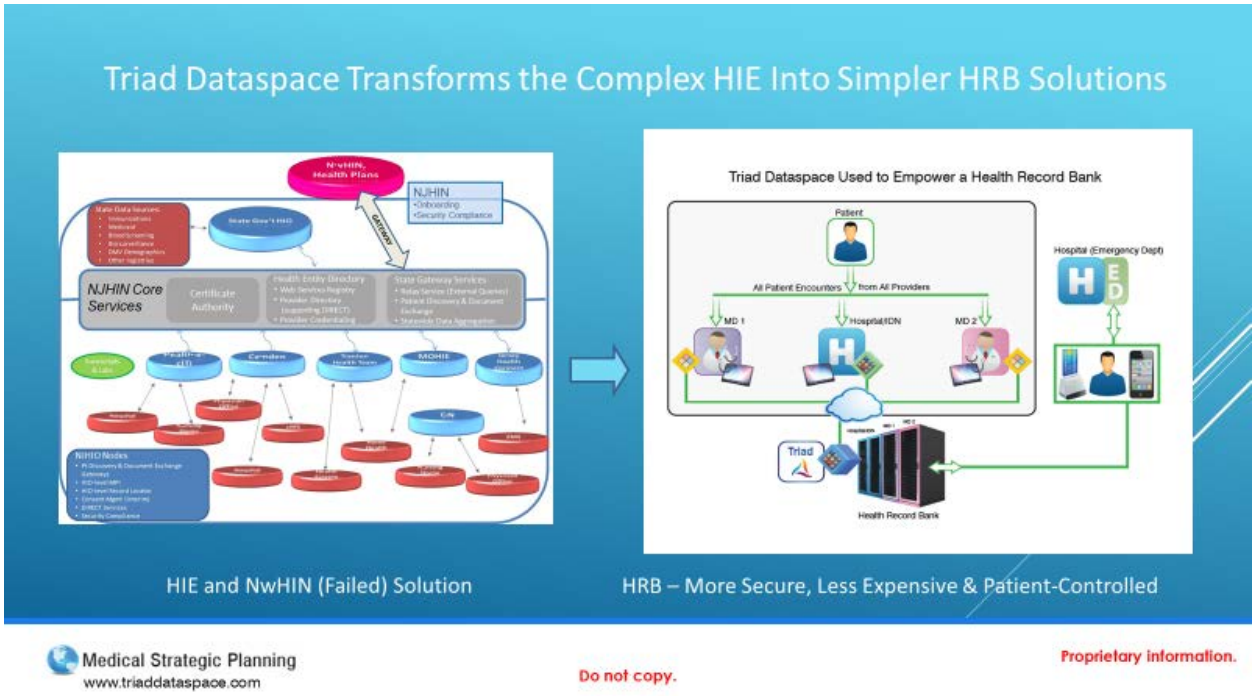
SUN3Law LEGAL ALERT - November 2018

TOPIC – CLOUD BASED STORAGE AND SECURITY ADVANCES

Two breakthrough advances in information technology have been introduced in the last few years. One is the Health Record Bank (HRB) architecture promoted by the Health Record Banking Alliance, a consortium of IT vendors who propose it as an alternative architecture to the federally-supported Regional Health Information Organization (RHIO) / Nationwide Healthcare Information Network (NwHIN) architecture that evolved from the original HITECH Act in 2009.

The Advantages of the HRB Architecture over the RHIO+NwHIN Architecture

The HRB architecture has several advantages over the RHIO+NwHIN architecture, specifically it allows retrieval of all a patient's records without having to access to the NwHIN or any nationwide Master Patient Index of patient; and secondly, if used as a primary storage point for an encrypted lifetime Precision Medical record, it would be potentially far more secure than the current government approach. A study reported by Gemalto on data breaches make clear that some solution to record security needs must emerge before anyone can have trust storing their lifetime, Precision Medical record anywhere in the cloud. The government RHIA and the alternative Health Record Bank (HRB) architectures are depicted below.



Looking at the diagram above, the simplicity of the HRB approach is apparent even to the non-technical reader. In the HRB architecture scenario, the patient would direct all of their medical providers to store the patient's medical encounter information (or at first at least a copy of it) in one

HRB of the patient's choice. This simplifies the process of later retrieving all of a patient's medical records from one location when needed later for routine or emergency care access.

The advantages of the HRB is that no nationwide health information network (NwHIN) or nationwide master patient index (MPI) is required to "find" a patient's electronic health record (EHR) files (or Lifetime Precision Medicine records) that are currently scattered all over the Internet in a variety of storage repositories, the location of which is unknown to any provider who needs to access them. The HRB is the only architecture in which a patient and all of their providers knows in advance where all of a patient's records are stored and has immediate access to them without a nationwide search of hundreds of RHIO sites and millions of provider submissions.

Only when a doctor knows where the records are, can they be rapidly and securely retrieved can care be safely provided, something that the $40 billion dollars expended on the RHIO + NwHIN has not yet achieved.

In fact, several government reports reveal that less than a third of the states have fully-functional RHIOs connected to the NwHIN, making any notion that the system is ready for "life-critical" use simply insane!

The straight-forward logic of the HRB architecture greatly reduces any dependence upon the NwHIN and eliminates any need for a national Master Person Identifier.Index maintained by some government entity; both of which are quite beneficial because they enhance security and circumvent further government invasion of privacy. The HRB approach puts the patient back at the center of controlling many aspects of their own medical records and replaces a super-expensive and inefficient government system, which a private-enterprise approach that is immediately effective and far less expensive.

Each patient/provider is free to store all of their patient's records in any one HRB, and multiple HRBs can compete with each other as cloud data stores for patient records based on storage cost, edata record security, ability to perform analytics and other factors important to the patient and their physicians. There could be two, twenty, two hundred or two thousand HRBs, but no matter how many there were, all the records for any one patient would be stored in only one of them – and be immediately accessible to the patient and their doctor's for care.

What About Secure Storage and Record Interoperability

The HRB architecture assures record accessibility and offers modest advances in e-data health record security. However, the security of all the EHRs of any patient is only as good as the security of the entire HRB in which their records are stored. The HRB architecture leaves unaddressed population health analytics, patient data encryption, EHR record interoperability, EHR record aggregation and iron-clad record security. That is precisely why the second breakthrough in information storage technology (Triad Dataspace™) is so important, since it provides all of the missing items when implemented using the HRB architecture, providing the first comprehensive and secure patient record storage, analytics and retrieval solution. Only when Triad Dataspace™ is combined with an

architecture like an HRB, is both e-data security and record interoperability achieved. These are the two missing elements which $40B in government expenditures on EHR systems has so far failed to delivery.

Issues With "Big-Data" Cloud Repositories

In the ideal case a big data repositories would have 3 attributes.

- The first attribute would be a small storage footprint.
- The second attribute would be full-time encryption of stored data.
- The third attribute would be very short record retrieval time to each "query" issues against the stored information structure.

The problem is that no real cloud data storage repositories were exhibited all 3 attributes; until now.

Implications of e-Data Reducing Storage Footprint

For example, if one uses a "zip" type data compression program to obtain reduced storage footprint, rapid e-data retrieval time is sacrificed. So is the ability to perform analytics on the "zipped" e-data mass. Worse yet, if the reduction of storage footprint reduction is significant, the amount of "empty" cloud storage space that must reserved to later "inflate" the zipped e-data mass becomes impractical. This is shown in the table below.

	Number of Attributes Per Record	Number of Records	Original File Size (bytes)	7Z Final File Size	7Z Reduction Ratio
File 1	4	30	810	472	1.7
File 2	4	42	1,141	757	1.5
File 3	4	250	6,784	1,515	4.5
File 4	4	1,000	103,000	31,000	3.3
File 5	19	1,200,000	45,225,000	8,484,000	5.3

In the above example, database did NOT include medical images, only s

Note that it you used a commercial ZIP program (7Z) for compression, and it achieved 5.3 times footprint reduction for File 5 (bottom line). That means the repository itself must have 45.2 Mbytes minus 8.4 Mbytes = 36.8 Mbytes of empty space in order to ever "reinflate" the 8.4 Mbytes of "compressed" e-data stored. Someone has to "pay" for the empty storage space that is required to ever reinflate the data – which is 4 times larger than the e-data itself. How cost effective is that? What if the storage footprint reduction was 100 times, not 5.3 or 1,000 times? The amount of "empty" space required become astronomical!

The only way e-data storage footprint reduction could ever become practical is if one can perform query and analytics operations on the "compressed" e-data footprint and the "compressed" footprint never needed to be reinflated. Only in that case would storage footprint reduction actually become cost effective (because no "empty" storage space would have to be maintained for compressed e-data reinflation). In that one case, storage footprint reduction would become a huge cost "game changer!"

Since the Accountable Care Act (ACA) requires performance of "population-level" analytics on records stored and conventional storage footprint reduction schemes like 7Z or PKZip don't allow analytics, footprint reduction has been considered impractical in cloud storage repositories used for EHR and Precision Medical record storage (whether they use the RHIO + NwHIN or the better HRB architectures). Thus, no ideal cloud storage repository that provides: 1) practical footprint reduction, 2) encryption for enhanced security and 3) expedited query and analytics processing – can be achieved, because criteria (1) doesn't appear to be achievable.

Because of what appears to be mutually-exclusive problems, cloud e-data storage vendors gave up on trying to achieve the Ideal "Big Data" cloud storage repository some time ago.

The Triad Dataspace™ Breakthrough

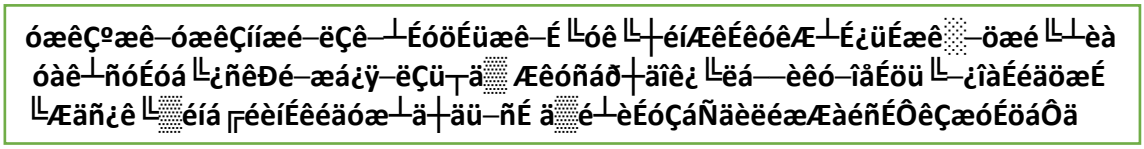
But there is a breakthrough, in fact a disruptive one. Suppose that one could in fact achieve the ideal combination of reduced footprint, and encryption (with a reduction in storage footprint) and efficient query and analytics on the encrypted, reduced footprint e-data *without* either decrypting or inflating it. What would that mean? Look at the missing part of the table above, shown below.

	Number of Attributes Per Record	Number of Records	Original File Size (bytes)	7Z Final File Size	7Z Reduction Ratio	Triad™ Reduction Ratio	Triad™ Final File Size	Triad™ Improvement over 7Z
File 1	4	30	810	472	1.7	188	4.3	2.5 times
File 2	4	42	1,141	757	1.5	260	4.4	2.9 times
File 3	4	250	6,784	1,515	4.5	1,090	6.2	1.3 times
File 4	4	1,000	103,000	31,000	3.3	3,265	31.5	**9.5 times!**
File 5	19	1,200,000	45,225,000	8,484,000	5.3	23,000	1,966	**370 times!**

In the above example, database did NOT include medical images, only structured medical text information

Note that the Triad representation achieves not 5.3 times storage footprint reduction but 1,966 times stored footprint reduction (on File 5), which is 370 times better than 7Z. That in itself is quite an achievement, and Triad™ itself isn't a storage footprint reduction technology. Triad™ is a new and more universal way of expressing information itself. Moreover, Triad generates encryption because the new e-data expression does not contain the original data, but instead a means to creating it.

The figure below shows the Triad representation of File 1 in the above Table. It is 188 bytes in length. The original file was 810 bytes in length. By looking at the Triad representation there is no way to determine the number of records or data fields, nor the meaning of any of the data fields. Therefore, possessing the Triad™ representation doesn't "expose" any of the data used to create it.
Triad™ is therefore both: 1) sparse and 2) an encrypted expression, meeting the *first two* (of the three) characteristics of an Ideal cloud storage repository!



Even with less than 1,000 bytes, Triad™ creates a footprint that is 250% smaller than created by 7Z (see row 1 and the table above). Note that as the data file sizes become larger, the reduction in stored footprint advantage for Triad™ over 7Z become much greater!

The above two figures reveal that 1) encryption and 2) practical reduced stored footprint are simultaneously possible – meeting 2 of the 3 criteria for an Ideal Cloud Storage Repository.

Triad™ Caveat – With Triad™ the original data type impacts the storage footprint reduction achieved. The examples shown are for text and all types of numeric data. However, for image data the reduction is not as impressive, but is still 2 to 3 times smaller than using JPG 2000 lossless compression. Like JPG2000, Triad™ is a bit-perfect, lossless, fully reversible (isomorphic and bijunctive) expression of the original image, but unlike JPG 2000, Triad carries additional image content information that JPG 2000 does not contain.

Queries & Analytics On Triad™ Representation

We saved the best for last, Triad performs query and analytics directly on its sparse, encrypted footprint and achieves data aggregation and information interoperability? Triad™ is both HIPAA and GDRP compliant, even when printed above, the source information is still protected. If you doubt this, decrypt the source information and send it to us. That means it is "hacker proof" and secure.

Triad™ is the first information representation technology that meets all three criteria for an Ideal Cloud Storage Repository, and it's the only technology that does!

How?

This leads naturally to the question, how can one query this encrypted information without first unencrypting and inflating it? It's a really good question and it is not simple to explain (nor will we do so since it's a trade secret), but here is a framework.

Aggregation and Interoperability

Triad™ is a dimensionless, binary information representation mapped into a unit interval. Data is no longer associated with metric parameters; rather it is universally represented.

In Triad™ body temperature is the exact same binary representation, whether the C-CDA expression temperature was original in degrees C, F or Kelvin! That is why it's inherently interoperable and data from multiple EHRs can for the first time be aggregated, checked for consistency, and used in a "dashboard" created from multiple EHR encounter summaries originally collected on different EHR systems with differing metric systems.

The Triad™ Distinction and Sustainable Competitive Advantage

In every other database management system use in cloud repositories, "temperature" as a parameter would be metricized to one system (say Celsius) and all data with other metricizations would have to

be converted by formula to the Celsius expression. *Triad™ overcomes this limitation* not only for temperature, but every other metricized parameter. A novel Triad™ Composite Clinical Data Dictionary (T-C2D2) is used to "remember" what metricization is expected for a particular supplier of incoming data, so it can be expressed properly when it is retrieved by that specific end user.

Processing Queries/Analytics

Query processing is the heart of all analytics. All information models need a means of efficiency processing queries.

Conventional information management system that are glued to data metrics, have: 1) data FIELDS (which may be *indexed* or *non-indexed*), 2) TABLES of multiple FIELDS, and express RELATIONSHIPS between FIELDS and the tables that contain them.

The Catch-22 of this conventional approach is that any query performed that contains a *non-indexed field* requires a record-by-record (read) of the entire repository to determine the validity of that query.

The larger the repository grows and the more records it contains, the longer the Search Operation takes, and that in turns greatly limits the speed with which analytics can be performed, since most analytics contain thousands of queries.

With Triad™ the validity of the query is determined without either decrypting or inflating the stored Triad™ representation, achieving the final (3rd) requirement for an IDEAL 'big data" cloud repository!

Because Triad™ can express each database record as a single binary expression, queries are performed by calculation, not by physical storage media "reads." The concept of "Search" in Triad™ representation is obsolete. In its place Triad presumes a "query" is a conversion request, and the query is converted to a Triad™ binary expression which has a numeric value. When sorted, the query "fits in" to the e-data records in one of two ways. It either matches the value of an existing record exactly (in which case the query is true) or it represents a "new" binary value that does not already exist in the representation. In this case, the "query" is false. In both cases the query value is calculated by the repository, which is a very fast calculation – making query response time a near-real time operation even for the largest repositories. "Search" is gone, Triad™ CIA transformation replaces it.

Triad's TWO Crucial Implications.

First, *analytics* can be done in the Triad™ Ideal cloud storage repository *without moving stored e-data* to some separate "analytics" repository. Moving data is always risky, and Triad™ avoids that exposure to data breach.

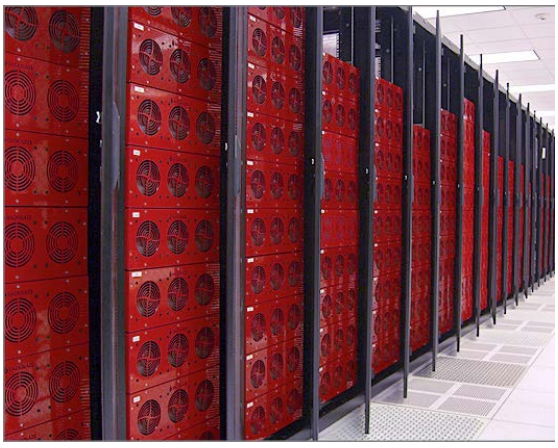
Second, analytics efficiency is vastly increased because each query is so much faster than in conventional databases. We are talking minutes, hours, days or even weeks faster for truly "big data" repositories with billions of records and Triad™ is scalable to those record volumes.

For some applications, Triad™ eliminates the need (and expense) of doing analytics for really "BIG" data on quantum computers or using HADOOP Hives for analytics.

Goodbye, Amazon Web Services, HADOOP Hives, Apache! No need to pay for them any longer.

Triad™ Data Storage Scalability Using COTS Technology

This has a profound implication for storage economy, security and analytics.



Start with an inexpensive storage appliance rack. The one pictured above left has 3 rows of 15 hard drives, or a total of 45 hard drives. Suppose it were filled out with a Western Digital 10 TB (terabyte) hard drive (say WD Red 10TB NAS) drive that costs $315 each. The cost to fill the appliance would be $14,175 for 450 TB (terabytes) of storage.

Now suppose that ONE vertical cabinet/bay (pictured above right) could hold 10 sstorage appliances. At 450 TB per rack, that would be 4,500 TB per cabinet, or 4.5 PB (Petabytes). The one cabinet/bay would cost $141,750 for 4.5 PB of storage. Any vendor could do this.

But now suppose the drives were written in the Triad dimensionless, binary representation that got 2,000 times smaller storage footprint. Suddenly, one $141.750 cabinet with the Triad drives could hold 9 EB (Exabytes) of stored data compared to the same cabinets storing only 4.5 Petabytes. The same number of drives are required in both cases, but the Triad™ configuration store 2,000 times more data.

Let's suppose we needed to store 9 Exabytes of data to start commercial operations of a small cloud storage repository. In the Triad™ case, two cabinets occupying about 30 square feet (SF) are required. In the case of a conventional implementation, 4,000 racks are required, costing $283.5 Million USD. Which would be the least expensive to build, power, cool and operate? Which is more cost-effective

to upscale for really BIG data? Scalability, if 110 racks with Triad data were put in a datacenter, it would only have to be about 1,700 SF to hold 1 Zettabyte of data. That seems like a significant breakthrough for cloud-based data storage repositories economics.

A Practical, Cost-Effective Triad Health-Lock™ HRB

Medical Strategic Planning, one of our clients is seeking funding to build an upstart Health Record Bank built out with only two Triad™ Cabinets for about $2 million USD. It would start with up to 18 Exabytes of storage and could license storage operations to any American company or healthcare provider interested in driving the cost of e-data storage and analytics down by 3 orders of magnitude.

Don't like 10 TB hard drives as storage media? It's OK, Triad™ does not impose any storage media constraints. Use conventional volatile DRAM or non-volatile (EMP survivable) bubble memory. Build it in a properly grounded, 10' x 40' Faraday container, to which only fiber optic cable enters.

If interested in funding and owning 49% of such an endeavor, give our client MSP a call and speak with Art Gasch, CEO or browse www.medsp.com on the Internet. The build time is about two years. Total project funding is about $8M. Ask for a Private Placement Memorandum (PPM) or just call MSP for a demo. Read the PPM. Yes, it's a risky endeavor, but one associated with a 5.2 P/E over the first 5 years.

Self Disclosure: SUN3Law is MSP's Outside Counsel, and an MSP shareholder and was engaged for the purposes of PPM review and other legal services.