DataFrameworks:
Data that Matters

Today, the amount of data to deal with is overwhelming. When it comes to unstructured data, stakeholders are not in a position to leverage it. They do not have the tools to evaluate and leverage the data which would otherwise contribute to making businesses better.

California-based DataFrameworks makes it to the list of companies that such stakeholders need to collaborate with to overcome this problem. DataFrameworks provides an innovative file management software—aptly named ClarityNow—that standardizes the organization of data across heterogeneous storage systems. The idea behind this is to enable enterprise users to more efficiently find, use, and organize files; and for IT administrators to manage storage infrastructure, efficiently provision storage, monitor and report on storage use, and effectively make decisions about new infrastructure purchases.

In an interview with CIO Applications Will Hall, CEO, and Paul Honrud, Founder, share insights on the company's solutions that make a difference in the life sciences space.

Challenges

"There are many challenges in this field, but the biggest of all is the huge amount of data generated by advanced technology and the ways to capture them. Businesses are not "data-enabled" and often lack the toolset to evaluate them for generating valuable business insights.

With the legacy setup, businesses lack the analytic research tools associated to bring the real business value of data. As a result, they spend an enormous amount of time and money understanding the assets, especially large MNCs, and managing the day-to-day orchestration

of data. Besides this, companies also suffer because of cost-inefficiency from an operational and capital expenditure standpoint—they have to spend millions of dollars in storing the physical data.

The Solution

We have developed ClarityNow to enable businesses to harvest all the unstructured data by providing a scalable neural time index and catalog all rich data with an OPED view of the digital aspect. It starts with the capability to capture the base index, and it's been a challenge to get a real catalog in the search file system. Customers sometimes take months to get a real understanding of data, but with our–near real-time index, they can scan file systems in a short period, once the initial database index is built. Once the baseline is found, businesses can quickly find the relevant data with the search feature we have created for them.

We can find files in real time. This ability to quickly find and act on the data is one of the core features of DataFrameworks. The visibility we have to organize data in a logical business fashion has helped businesses deal with information systematically. We also have a chargeback feature that allows companies to report on the true cost of the digital aspect quickly. Our primary aim is to help businesses make better decisions around data on a daily basis. Besides, we leverage the existing tools and frameworks associated with large systems for security.

The Methodology

We have one of the fastest scanning algorithms, and the purpose built object-oriented database can ingest over 1.5 million files per second. The resulting impact for a customer through our method is that they have an

PAUL HONRUD,
FOUNDER

accurate dataset in real time to take a historical look at data. It is paramount first to give the customer an understanding of what is happening in their environment and that, I believe, solves 80 percent of problems. Getting the core index right is essential to start the systematic process. Leveraging the features on how to organize and understand the data and allowing them to view the data in a logical business fashion also are quintessential in this equation.

> **The ability to quickly find and act on the data is one of the core features of DataFrameworks**
>
> WILL HALL,
> CEO

Success Stories

A reasonably large-scale genomic research company in the West Coast was attempting to build a scientific archive. They identified three to four core aspects and had a look at the entire environment, and they had a disparate file system. However, what they needed was a tool to give a universal picture of the whole environment, and it had to be something that the end users could use namely research scientist and primary investigators.

Therefore, we built for them a system with the help of system vendors and a local systems integrator back in 2013. In the first thirteen months, the client with our system, had archived 4.5 petabytes of data, and in the course of three years, the company saved 2.3 million dollars. **CA**