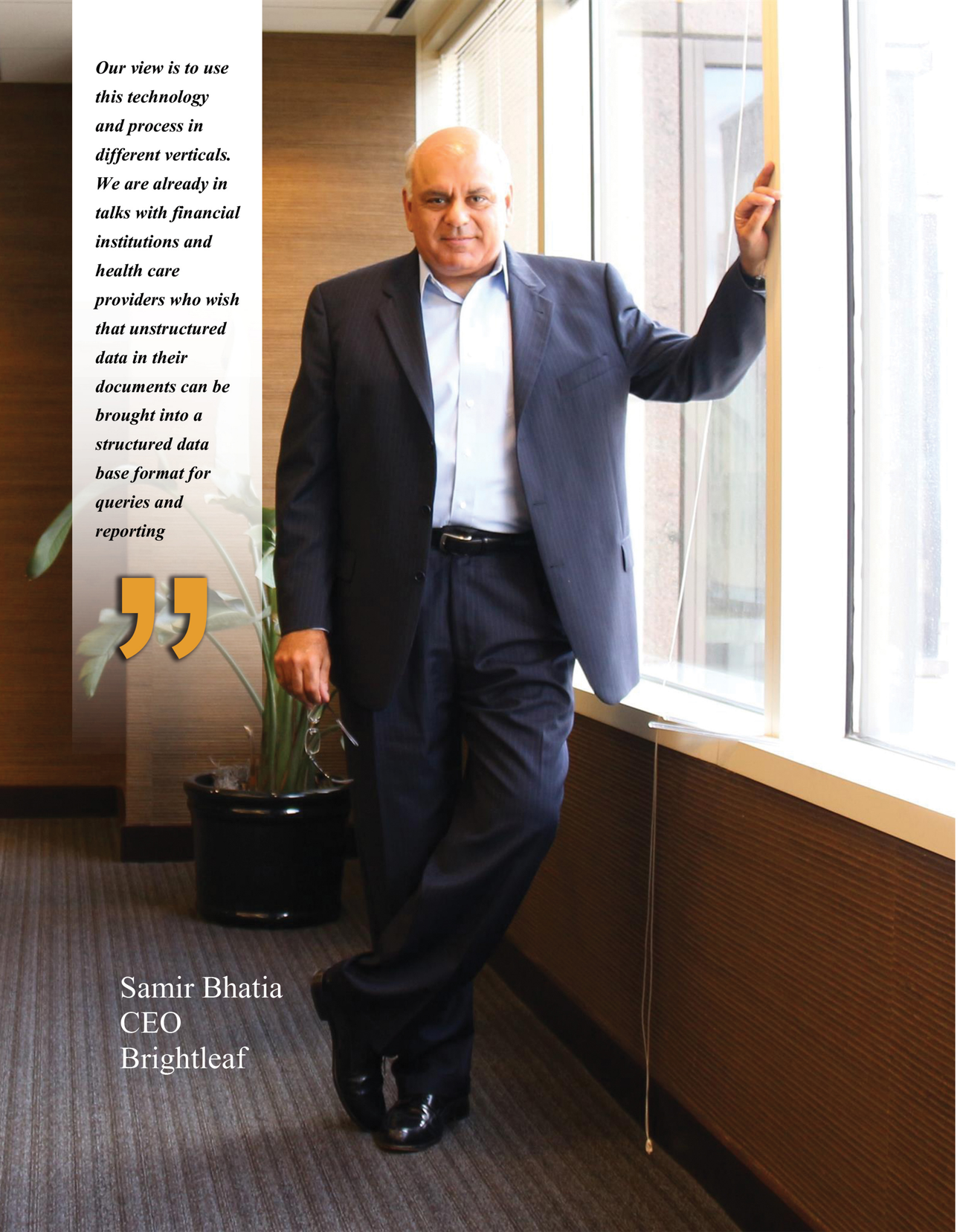

Our view is to use this technology and process in different verticals. We are already in talks with financial institutions and health care providers who wish that unstructured data in their documents can be brought into a structured data base format for queries and reporting

Samir Bhatia
CEO
Brightleaf

Simplifying Work Through Semantic Engine

In this highly competitive environment, organizations of all sizes need to access crucial information like terms of a contract or auto-renewal clauses from within a plethora of data. This often leads to error prone and slow information gathering which can set a company back by facing huge losses. Brightleaf is a highly focused company which solves the most arduous task of extraction of data-points from legacy contracts that can be populated into a Contract Lifecycle Management System (CLM). They do this using a combination of state-of-the-art software technology comprised of Natural Language Processing (NLP), Machine Learning (ML) and Artificial Intelligence (AI) technologies. While software does most of the work, a team of highly specialized lawyers ensure up to 100% accuracy in extracted data points.

This data, when populated into a CLM system, adds tremendous value to the system as the CLM can trigger when the contacts expire, run reports on the extracted data to keep the users in compliance with any regulatory changes, and even find hidden revenue by policing penalty clauses, CPI increases, and other parameters and information. “We extract each and every obligation from contracts, sometimes running into couple thousand paged contracts needing multi thousand data points!” says Samir Bhatia, CEO of Brightleaf. “We have signed alliances and work with most of the CLM vendors - Apttus, Conga, The Corridor Company, Determine, Symfact and others, as we solve a major pain point for their clients.”

The company has focused on having full control of all three dimensions to solve a real problem. Software, People and Process. To give an example, what if someone has mis-typed items in a contract? What if the scan from a paper document was poor that any OCR software fails? That is where the human element comes in. But the duo, software and people, is not enough unless there is a process following Six Sigma principles encompassing it. Brightleaf’s software works at multiple levels. At the highest level, it can be configured to find patterns of words. At the next level, it knows Natural Language – verbs, nouns, plurals, even lemmas. Encompassing that is an AI and ML layer, that can learn as it extracts more information and more documents are processed through the technology. The software, when configured, never misses anything. But there can be a lot of missing and wrong information in the original documents itself which no technology can find.

For instance, For a Class 1 railroad company, Brightleaf extracted data from their track maintenance agreements, including data-points such as the length of track, the mileage marker where the track starts and ends, and whose responsibility it is to maintain that track; the railroad company or the counter party whose property/land the track sits on. It turns out that the railroad company was maintaining the tracks, when in fact the other party should have, and they found millions of dollars of yearly billings that were not done. “We even processed contracts as far back as 1892!” adds Samir.

Brightleaf follows a very stringent process, in which they define in detail not only the data points that need to be extracted but also the format in which the client requires it to ingest the data into a CLM system (rules to follow). With this stringent process, and up-front work, there is nothing left to interpretations or challenges at the end. “Our view is to use this technology and process in different verticals. We are already in talks with financial institutions and health care providers who wish that unstructured data in their documents could be brought into a structured data base format for queries and reporting,” says Samir.IE