The rise of open source

Once restricted to Silicon Valley, open source solutions such as Hadoop are finding their way into the enterprise and being used by mainstream firms around the world as data storage and processing engines. And it’s just one of many open source solutions that are finding their way into the enterprise. Others include Mahout for machine learning, Spark for complex event-processing, and specialized tools that are being adopted alongside commercial software. And, of course, there’s R, the open source language and environment for statistical computing and graphics.

The key to an open source initiative is finding the distinct value that can come from adopting the solution. Open source can have a distinct role, but it generally has to be part of a broader overall strategy. For example, Hadoop can be effective when you have “real” big data that is multistructured, volume heavy, and slow to process. It’s a case of finding the right tool for the job.

Risk management must also be part of the equation when an open source tool is used. What happens if the army of volunteer open-source developers moves on to the “next big thing”—or simply wants to be paid? What if the quality of the solutions declines along with the quality of talent working on them? It’s easier to calculate your risk exposure if you have a clear picture of the portion of your infrastructure that relies on, or is built on, open source solutions.

Open source solutions come with unique benefits, not the least of which is economic value. That said, companies have to keep in mind the cost and availability of people who can work with these emerging technologies. Those people are getting harder and harder to find.

#AnalyticsTrends2015