RStudio looks to the cloud, sheds light on open-source-based data science business

OCTOBER 29 2020

By Krishna Roy

The company is delivering its data science portfolio in the cloud as it continues to build an open source and commercial business around R and, increasingly, Python open source programming languages used to code data science.

451 Research®
Now a Part of

S&P Global Market Intelligence

REPORT REPRINT

Introduction

When we checked in with RStudio earlier in 2020, we noted that the open-source-based data science vendor had a game plan for the cloud underway. RStudio is now elucidating its cloud strategy as part of an overarching objective to continue to develop free and open source tools, as well as commercial offerings for coding data science in R and Python, and to make these projects scalable, shareable and enterprise-ready. RStudio has also provided visibility into its business, which became a public benefit corporation (or B Corp) earlier this year.

451 TAKE

RStudio is an intriguing and somewhat unique company. The vendor is seeking to balance a mission of creating free and open source software for data science and scientific research with the pragmatic reality of operating a commercial business that is profitable but also socially responsible as a B Corp. RStudio's cloud strategy should increase its appeal to existing customers, given customer adoption trends (see below), and bolster its addressable audience. The public cloud is the most popular venue for each stage of the AI process, according to 451 Research's latest Voice of the Enterprise: AI & Machine Learning, Infrastructure survey. However, any data science platform that is free or low-cost presents a competitive threat. RStudio's free open source tools compete with other free open source data science tools, while the vendor's commercial data science platform was designed to be a low-cost alternative to other commercial enterprise data science environments.

Details

RStudio continues to invest in open source R and Python communities, noting that it devotes 50% of its engineering resources to open source data science. That said, RStudio continues to have commercial software ambitions, which are another aspect of the firm's fresh positioning for 'serious' data science. To RStudio, 'serious' data science means it needs to be open source and code-oriented (because coding is the most powerful approach), as well as centralized.

Furthermore, RStudio is extending its 'serious' data science message to the cloud. The first fruit of this strategy is RStudio Cloud, which the vendor released into general availability in August. RStudio Cloud was designed to be a lightweight, cloud-based environment for coders to carry out, share, teach and learn data science online in a web browser. It uses the RStudio IDE – for which the company is still arguably best known – as a linchpin and is available at no cost, as well as through charged-for pricing plans, starting at $15 per month, plus tax, for instructors to teach with it.

At the same time, the company notes that many of its customers want to move from an on-premises deployment to the cloud. In order to address cloud rollouts from an IT and DevOps perspective, RStudio plans to expand cloud-native integrations for authentication, logging and scaling purposes in 2021. The endgame behind unleashing these capabilities is to make three core offerings: RStudio Server Pro, Package Manager and Connect (explored here) cloud-ready.

While the cloud is one development focus, it's not the only one. RStudio will continue to embrace Python as well as R within its portfolio. For example, Python support in RStudio Package Manager is currently in beta release. Making it easier for data science teams to move apps, reports and models from development into production by unleashing project management capabilities is planned for 2021.

Additionally, RStudio is hatching a partner strategy to address arenas the vendor doesn't want to serve in its portfolio, such as AutoML and drag-and-drop features. RStudio also wants to complement ModelOps platforms and BI tools through tie-ups with partners.

On the business front, RStudio is expanding its industry focus from a sales perspective by initially alighting on insurance and pharma, with a view to focusing on additional verticals in future. The firm has also changed the way it reports customer numbers. RStudio now reports every sale. As a result, the company cites more than 5,100 active commercial software accounts. The company also cites a headcount of 185 employees. RStudio reported 1,200 customers during our previous check-in in January, when it also cited 160 employees.