

REPORT REPRINT

RStudio Inc becomes RStudio PBC, expands data science product focus with Python support

FEBRUARY 4 2020

By Krishna Roy

The company has changed its corporate status from a C corporation to a benefit corporation, and has shifted its product focus beyond the R open source programming language to support Python open-source-based data science.

THIS REPORT, LICENSED TO RSTUDIO , DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, WAS PUBLISHED AS PART OF OUR SYNDICATED MARKET INSIGHT SUBSCRIPTION SERVICE. IT SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR RE-POSTED, IN WHOLE OR IN PART, BY THE RECIPIENT WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



Introduction

RStudio has operated as a C corporation since it was founded as an open-source-based data science and analytics vendor by CEO J.J. Allaire, the creator of the ColdFusion programming language and web application server, in 2009. The vendor has now changed status to a B Corp – or public benefit corporation – in order to formalize the way it has been run for some time, as a force for good. RStudio is also undergoing an expansion in its product focus. Having concentrated on supporting the R open source programming language for data science and analytics since inception, RStudio has now embraced the Python open source programming language, too.

451 TAKE

The vast majority of data science and machine learning platforms support R- and Python-based data science, as well as a number of other open source tools and frameworks employed for advanced analytical use cases – in recognition that data scientists, as well as other types of users, have their preferred languages and packages that they don't want to ditch in order to learn something else. RStudio's platform is therefore a stronger offering as a result of embracing Python. However, the vendor could do better in addressing the needs of non-data scientists, who could be assisted through the introduction of AutoML functionality. After all, 18.1% of organizations surveyed in 451Research's latest Voice of the Enterprise: AI & Machine Learning survey cited a skills shortage as the most significant barrier to using machine learning, which ranked as the second-biggest obstacle behind budgetary limitations, as cited by 19.1% of respondents.

Context

Now formally a B Corp, rather than informally being run as one (see below), RStudio has also expanded its product strategy since our last check-in with management in early 2019, when R-based data science was the company's raison d'être. RStudio's roots in supporting data science using the R open source programming language date back to the vendor's eponymous IDE employed by data scientists to develop statistical programs using R. Open source and proprietary tools that enable R to be more effectively utilized for data science and analytics remain a core part of the company's focus – as is active engagement with the open source community, in general.

Furthermore, RStudio expanded its strategy to support enterprise data science via the delivery of three proprietary offerings – RStudio Server Pro, RStudio Connect and RStudio Package Manager – in 2018.

In 2019 RStudio kicked off a strategy to support another popular language for data science and analytics: Python. The company is adopting an incremental approach to Python support by infusing it into its existing portfolio, in recognition that many organizations struggle to use R and Python together, but need to do so because of a corporate mandate or other reasons (see below).

Public benefit corporation status

RStudio started out as a C Corp, which is the standard (or default) incorporation of a US company under IRS rules. However, about a year ago Allaire and president Tareef Kawaf started investigating changing it to a B Corp – otherwise known as a public benefits corporation. Management’s thinking was that the company was run in accordance with B Corp principles anyway, so making a legislative change to validate that was the next logical step.

For those that don’t know the difference, B Corps essentially run on metrics other than profitability, involving social and environmental objectives, which are not critical priorities for C Corp organizations. Even though some do operate with B Corp ideals, such as placing employee welfare as a top priority, C Corps are focused on profit.

RStudio intends to be a profitable B Corp, noting that it already has a history of profitable growth, and will aim to keep in the black – even though it has changed its corporate status and obtained the certification required to legally operate as a certified B Corp.

The B Corp certification process involves evaluating the impact of a company’s operations and business model on workers, community, environment and customers. It includes looking at an organization’s supply chain and input materials, charitable giving, and employee benefits to ensure they meet certain conditions. Moreover, it is designed to send out a message that the company has its employees’ welfare at heart, considers something other than profit as a motivator, and is in it for the long haul – selling a B Corp requires taking its legal organization as a public benefit into consideration.

Furthermore, it is worth noting that not all investors have a funding policy to support B Corps – even though the company’s existing backers note that they are supportive of the move. RStudio’s investors are General Catalyst Ventures and Baseline Ventures. They invested an undisclosed sum in the company in May 2013, which is the only financing it has received, although management notes that RStudio receives financing offers regularly. These offers are turned down because they are not needed; the company can grow without them.

As evidence of RStudio’s recent growth, the company cites 1,200 paying customers at the end of 2019, noting that its modular platform is used by thousands of data scientists and hundreds of thousands of business users within these accounts. RStudio introduced new packaging and pricing in 2019 to give its data science platform further luster. The vendor also notes that it has a headcount of 160 employees across the US, Canada and Western Europe; it intends to have more than 200 staff by the end of 2020, in accordance with its strategy to continue to grow profitability.

Product strategy

While data scientists remain a core audience and the type of user for which RStudio is best known, the vendor is also seeking to address the needs of other users involved in data science projects and scientific research. RStudio’s corporate mission is to create free and open source software for data science that can be used by IT administrators, DevOps personnel and business users – a strategy the vendor has addressed by supporting only R in the past.

RStudio’s move to embrace Python as well as R is gradual and ongoing. The company is keen to get away from the idea that it is merely serving R devotees and is an R advocate by infusing Python support into core offerings within its data science platform. RStudio is also getting involved with the Python open source community, as well as the R one.

REPORT REPRINT

Integrating Python support into RStudio Server Pro is an early deliverable from the company's fresh product strategy. RStudio Server Pro is the proprietary version of the RStudio Server open source platform, which the company also supports. RStudio Server Pro is all about providing organizations with a back end to the RStudio IDE for collaboration, centralized management, metrics and security.

RStudio Pro Server provides similar capabilities for users of Jupyter Notebooks and Jupyter Labs, which are popular Python-based tools in use by data scientists. Jupyter Notebooks enable the creation and sharing of documents that contain live code, equations, visualizations and explanatory text, while Jupyter Labs essentially enables users to work with content including Jupyter Notebooks, as well as text editors and custom components, in a flexible, integrated and extensible way. In short, integration with Jupyter Notebooks and Labs enables data scientists to develop R and Python applications and code using a standard editor.

RStudio Connect's ability to host Python content, including Python Notebooks, is another fruit of the firm's strategy to support Python-based data science. RStudio Connect is the company's publishing environment. It used to be for publishing all work a company's data team carried out in R, but now the offering will support data science and analytics performed in Python, too.

Python support comes via integration with Jupyter Notebooks, so organizations get a flexible way to deploy Jupyter Notebooks – as well as reticulated R content – against a variety of Python versions. Reticulation is underpinned by a reticulate package that was designed to provide a comprehensive set of tools for R and Python interoperability, including translation between R and Python objects and the ability to call Python from R in a variety of ways.

Finally, it is worth noting that RStudio also has a cloud strategy underway. It revolves around the free RStudio Cloud, which the vendor created for professionals, hobbyists, teachers and students in order for them to share, teach and learn data science using R – which is why the RStudio IDE is available on the RStudio Cloud. However, the game plan is to flesh out the RStudio Cloud and the company's cloud deliverables in order to attract SMBs.

Competition

RStudio has definitely elevated its competitive credentials, given that the vast majority of data science platforms already support R- and Python-based data science – albeit using a somewhat different modus operandi that doesn't involve deep community participation or the support of free open source tools. Furthermore, RStudio is different from other data science platform providers targeting enterprises because the firm doesn't have an outbound sales team, which means it is unlikely to face off against vendors that do, including Dataiku, Domino Data Labs, Alteryx, BigML and Ayasdi.

RStudio's focus on the individual data scientist means it is likely to elicit comparisons with Gigantum, Mode Analytics, ZEPL and RapidMiner Studio, which also target this audience. Its focus on providing data science as well as scientific research for teams reminds us of SAS Institute, Mathworks and Wolfram Research.

We think RStudio's evolving cloud strategy, while further cementing the vendor's data science credentials, will also take it into more competitive bake-offs with Amazon Web Services, Google and Microsoft, which are also gunning hard for the same types of users RStudio is courting.

Finally, it is worth pointing out that RStudio plays a complementary and competitive role in certain situations. The RStudio open source IDE is supported in a variety of data science platforms because data scientists like to use it to create R-based analysis, which potentially gives the company a foot in the door when they are seeking more than a free development environment to use.

SWOT Analysis

STRENGTHS

RStudio's B Corp status singles it out from the rest of the data science platform and analytics crowd – we've never come across another vendor in these sectors with these credentials. RStudio's expanded focus on R and Python makes it more attractive and relevant to today's data science needs.

WEAKNESSES

RStudio is light on automated machine learning (AutoML), which is becoming table stakes in order to address non-data scientists, who are a faster-growing audience than data scientists. The company could also do a better job of addressing MLOps (data science operationalization).

OPPORTUNITIES

RStudio is still arguably best known as the purveyor of the RStudio open source IDE, which means the firm can offer organizations free open source data science, as well as charged-for additional functionality that is well suited to enterprise deployment.

THREATS

Any data science platform vendor that offers free or low-cost tools presents the biggest rivalry to RStudio, since the company has a highly cost-conscious pricing structure for its entire portfolio.