



Should data scientists using R and Python switch over to Julia?



AP/WILFREDO LEE

Code switch?



By **Dan Kopf**

Data editor

Published August 17, 2018 • This article is more than 2 years old.

It might be time for data scientists to learn a new programming language. Particularly if they have a need for speed.

Last week, the lead developers behind the open source programming language Julia announced the 1.0 release of their project. This signals that the language, which is optimized for data analysis and machine learning, is no longer a work in progress. Julia code written in the 1.0 version will still work even when new versions are released—by contrast, code written in version 0.4 was not guaranteed to work under version 0.6.

Users of popular data science languages like R and Python may be perturbed by the notion that they learn something new. These programmers have likely spent years getting to know all the nooks

and crannies of their favored language and, for the most part, probably feel satisfied. Why should they consider using something else?

“If you are a mathematician, scientist, or engineer, you have historically had the choice to pick a language that was fast, like C++ or Java, or a language was easy to learn, like Matlab, R, or Python,” says Viral Shah, one of the creators of Julia. “In Julia, we created a language that was simultaneously fast and easy.”

Shah says that the key inspiration for developing Julia was seeing how many people had to write the same program twice. Data scientists would first use a tool like Python or R to develop an algorithm, because it was easy to explore the data and make charts in those languages. Then, when they were happy with the algorithm, they would rewrite the program in C++ or Java to get fast computer processing performance. Julia is faster than Python and R because it is specifically designed to quickly implement the basic mathematics that underlies most data science, like matrix expressions and linear algebra.

Julia is already widely used, with over 2 million people having downloaded it, but the community of users has bigger ambitions. It hopes that Julia will overtake Python and R as the central language for data science, and particularly for machine learning. Shah nows runs Julia Computing, a consultancy that helps other companies implement Julia. The New York Federal Reserve and the investment firm Blackrock are among its customers.

Most of the key developments in Julia now come from MIT’s Julia Lab, which is led by fellow Julia creator and MIT mathematics professor Alan Edelman. Julia’s other two creators are Jeff Bezanson and Stefan Karpinski (the name of the language came from an old project’s of Bezanson’s). These developers now are only a small part of Julia’s progression, with over 700 volunteers contributing to the 1.0 version of the software.

So, why shouldn’t every data scientist learn Julia? There are a couple of reasons.

One, if processing speed isn’t important to you, Julia is probably

inferior to whatever product you are using—at least for now. I am an R user, and most of the statistics work I do is on relatively small datasets, and involves simple calculations. The community of R developers, particularly the rockstar data scientist Hadley Wickham, have developed terrific tools, with thorough documentation, for doing simple data analysis tasks. I tried using Julia to complete some of the basic tasks I now do in R. Julia's tools did not seem as developed for these purposes.

Second, Julia is behind Python and R in terms of tools for debugging and identifying performance issues. Shah says that now that the basics of the language are completed, he hopes more of the community will turn to developing these tools, which make the language easier for new users.

Personally, I am not ready to take the plunge into learning Julia, but I can imagine that one day I will be. Once the kinks are worked out, Julia's speed advantages put it in prime position to become the key language for data scientists.

Kick off each morning with coffee and the Daily Brief (BYO coffee).

By providing your email, you agree to the Quartz Privacy Policy.