

GopherCon 2021 Lightning Talk

Northwestern's Data Science Program Features Go

Tom Miller

Faculty Director, Data Science

tom.miller@northwestern.edu

Northwestern

SCHOOL OF
PROFESSIONAL STUDIES

My story begins in an applied statistics class

I don't care what language you use . . .

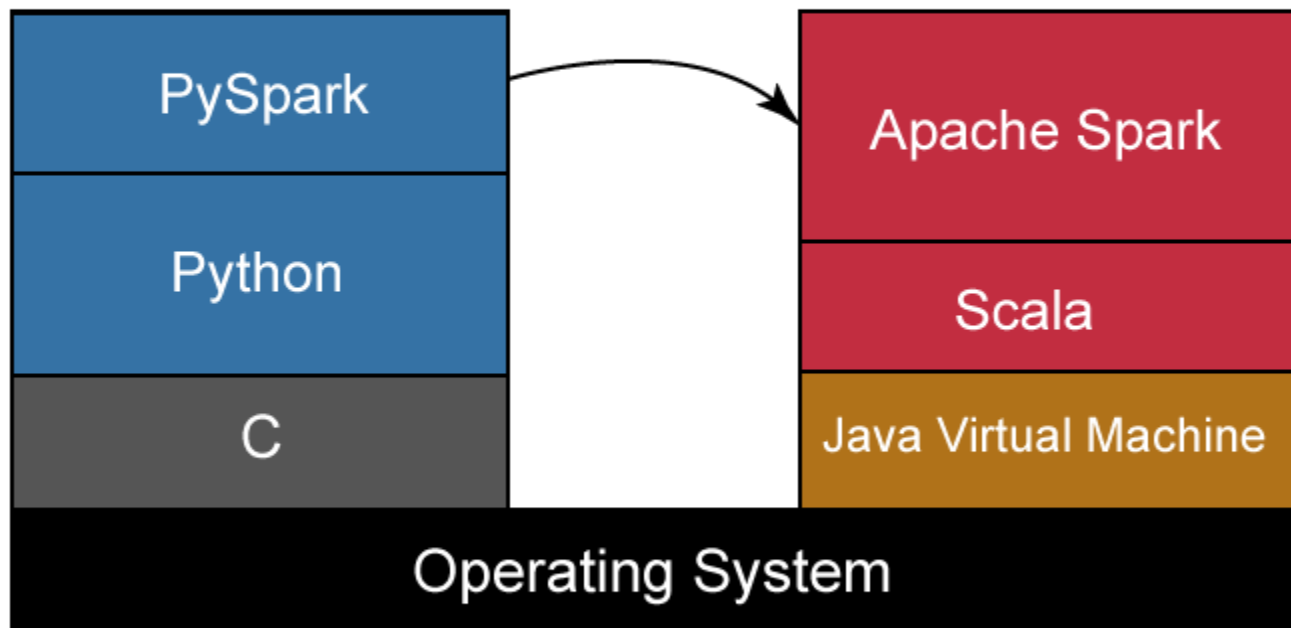


Organizations need data science, and data science needs Go.

Advantages of Go:

- Go is machine efficient
- Go is programmer efficient
- Go simplifies the software stack

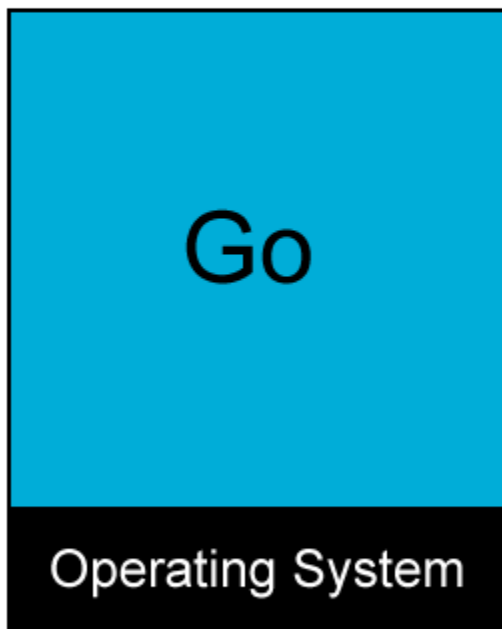
Infrastructure Alternative: Spark for Scalability and Performance



Rather than working directly with Scala or Apache Spark, data scientists may prefer a Python client, PySpark

Infrastructure Alternative:

Go for Simplicity, Scalability, and Performance



November 2009

Go introduced
written in C

February 2015

version 1.5

Go becomes self-hosting

Go is written in Go

Data Science Needs to be Multilingual



Explore
Data



Develop
Models



Implement
Systems

Three Languages for Data Science

Used to explore data, develop models, and implement systems:

- R, with many packages for analytics and modeling, is well-regarded by statisticians. It is an excellent choice for scientific programming and applied research.
- Python is the most popular language in data science. It is especially strong in natural language processing and serves as the primary client to deep learning platforms.
- Go is a systems programming language designed for today's multi-processor computers. It is well-suited for implementing scalable, high-performance systems for data science.

Students gain experience with the three languages for data science and can tailor their studies to one language or another.

Three Languages for Data Science

IEEE Top Programming Languages 2021

<i>Rank</i>	<i>Language</i>	<i>Introduced</i>	<i>Trending</i>	<i>Jobs</i>
1	Python	1991	1	1
2	Java	1995	2	3
3	C	1972	3	2
4	C++	1985	5	5
5	JavaScript	1995	4	4
6	C#	2000	11	9
7	R	1993	12	7
8	Go	2009	6	6

<https://spectrum.ieee.org/top-programming-languages/>

Data Science at Northwestern University

Courses with Go as the Primary Language

- MSDS 431-DL Data Engineering with Go
- MSDS 432-DL Foundations of Data Engineering
- MSDS 436-DL Analytics Systems Engineering
- MSDS 459-DL Knowledge Engineering

Many data science faculty accept assignments in any of the three languages for data science.

My story ends

I don't care if you use Go, Python, or R, as long as you do your own work.



Organizations need data science, and data science needs Go.

Master of Science in Data Science for Individuals

<https://sps.northwestern.edu/masters/data-science/>

- datascience@northwestern.edu
- 312-503-2579

Professional Data Science Training for Corporations

<https://sps.northwestern.edu/information/data-science-for-professionals.php>

- corporatedatascience@northwestern.edu
- 312-503-4925

Tom Miller

tom.miller@northwestern.edu

Northwestern
SCHOOL OF
PROFESSIONAL STUDIES