

Go for Data Science

Daniel Whitenack, Ardan Labs

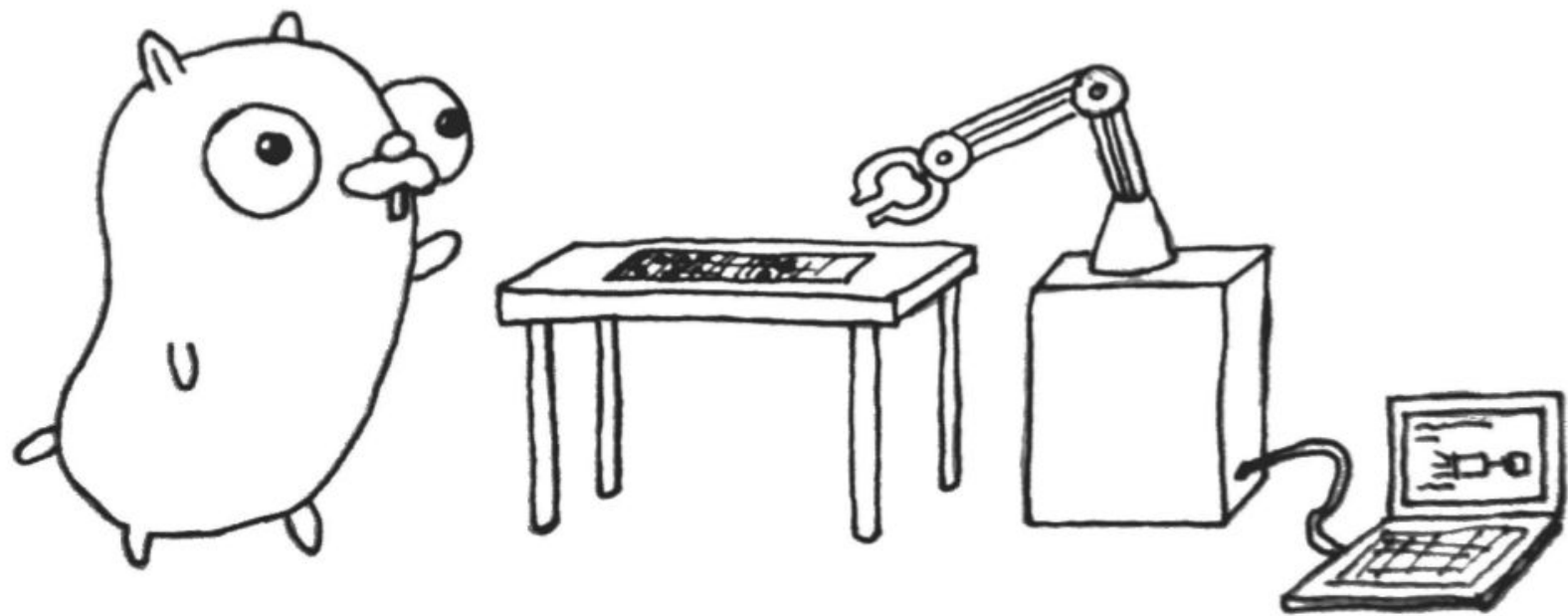
@dwhitena, @ChicagoGolang, #datascience

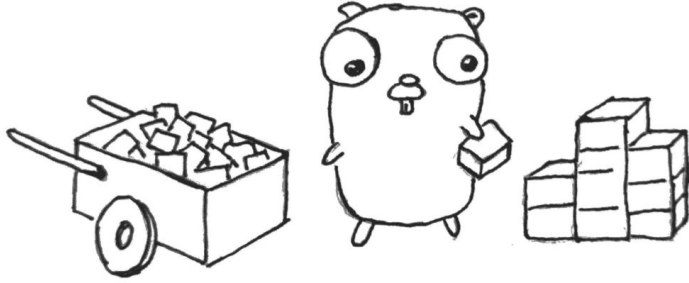
Outline

- Intro
- Data science.
- Data science with Go.
- Data science Gophers.
- Data science demo.

Intro

Data Science

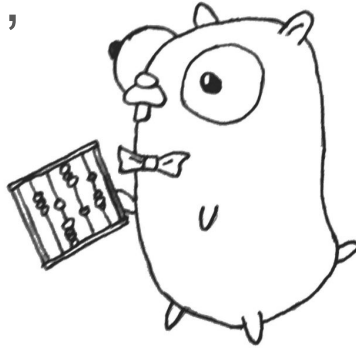




ETL, Data Cleaning,
Organization



Parsing, Extraction of
Patterns



Arithmetic

Data Science with Go

Oh yeah...

Data Science Gophers

live here ...



The Coral Project



SANTA FE
INSTITUTE



ŷhat

influxdata

and interact here ...

#data-science, gophers slack

https://gophers.slack.com/messages/data-science/

Gophers ▾
● Daniel Whitenack

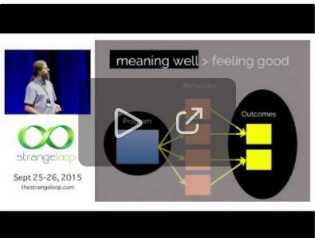
CHANNELS (417)
cayley
chicago
data-science
general
go-interpret
golang-jobs
gophercon-denver
gopherconband
gotimefm
mgo

DIRECT MESSAGES (9721)
♥ slackbot
● Daniel Whitenack (you)
○ Alex Sánchez
○ Andrea Masi
○ Brendan Tracey
● Carlisa Campos
○ Fabrizio Mito
○ Kevin Keuning
○ Leo Kalneus
● Sebastien Binet
● Varun Kumar

#data-science
197 members · Golang for Data Science, ML, Statistics, NLP, etc. Resources (under dev): <<https://github.com/gopherds/resources>>, Discussion list: <<https://groups.google.com/forum/#!forum...>>

🔍 Search

"I See What You Mean" by Peter Alvaro ▾



Daniel Whitenack 7:35 AM
Thanks @sam!

Samuel Lampa 7:35 AM
It outlines how datalog queries are resulting into dataflow program on-the-fly, and how to make these dataflow programs distributed, and a bit more

The nice thing with this is how well suited Go is for data flow programming (I more or less see Go as a dataflow programming language... and with something like [goflow](https://github.com/golang/goflow) or [flowbase.org](https://github.com/golang/flowbase) you even come a bit closer to the next step, flow-based programming!) (edited)

So, we now just need a flexible query language that generates the Go code for us 🤖

Daniel Whitenack 7:37 AM
I really need to look more into this and do some reading/experimenting. I love the thoughts around this.

Daniel Whitenack 7:48 AM ☆
I think I just need to experiment and read to help things sink in.

@sam! do you have a good reference for data flow programming in general?

Along with the video above.

Samuel Lampa 7:52 AM
@dwhitena Apart from the (pretty fragmented) plethora of blog posts, there are two nice books. One really short and concise one, outlining the differences between dataflow, reactive and flow based programming: <https://amzn.com/1497422442> ... that I think should be a good start.

Then, for going deeper into FBP, John Paul Morrisons FBP book is the way to go: <https://amzn.com/1451542321> ... it is much longer, and much less focussed though, although it contains a lot of interesting insight. (edited)



Otherwise, where I learned through doing, was by playing around with Slibrov's GoFlow library (<https://github.com/trustmaster/goflow>).

Also, the FBP mailing list is shock-full of super-interesting discussions, and very very insightful people (mostly IT veterans from industry): <https://groups.google.com/forum/#!forum/flow-based-programming>

Otherwise, wikipedia should of course not be forgotten. I know the FBP guys have tried to keep the FBP article up to date lately. https://en.wikipedia.org/wiki/Dataflow_programming

+ Message #data-science

[gopherds Google group](#)



Groups

[NEW TOPIC](#)[↻](#)[Mark all as read](#)[Actions ▾](#)[Filters ▾](#)

My groups

Home

Starred

▼ Favorites

Click on a group's star icon to add it to your favorites

▼ Recently viewed

[gopherds](#)

[golang-codereviews](#)

[mgo-users](#)


gopherds

Shared privately

3 of 3 topics ☆

Welcome! This is a place to discuss all thing Go and data science.
[Edit welcome message](#) [Clear welcome message](#)


☐



Curated Go data science resources (1)

By me - 1 post - 0 views


☐



Introductions?

By Samuel Lampa - 2 posts - 8 views


☐



Video: Go for Data Science by Daniel Whitenack




By Samuel Lampa - 3 posts - 4 views


[gopherds on github](#)



This organization |

[Pull requests](#) [Issues](#) [Gist](#) [ToDo](#)


  





GopherDS (Go Data Science)


Tools/training for Go-based data science.

[#data-science on gophers slack](#)

 **Repositories**

 **People** 3

 **Teams** 0

 **Settings**

Filters ▾

New repository

resources

Resources for Go-based data science.

Updated 6 hours ago

Go

★ 57

🔗 5

gophernotes

It's time to use Go in notebooks.

Updated 7 hours ago


Go

★ 494


🔗 31

People


3 >



dwhitena
Daniel Whitenack



gerred
Gerred Dillon



sbinet
Sebastian Binet

Invite someone

Data Science Demo



github.com/gopherds/gophernotes

github.com/gonum

github.com/kniren/gota

github.com/sajari/regression



Thank you!

Daniel Whitenack, [Ardan Labs](#)
[datadan.io](#), [@dwhitena](#), dwhitena on gophers slack

The Go Gopher, the basis of the drawings herein, was created by [Renée French](#).