

# Guilherme Vituri

vituri.vituri@gmail.com / <https://github.com/vituri>

## EDUCATION

### Doctor of Philosophy – Topological Data Analysis

Universidade Estadual Paulista (Unesp) / Ohio State University (OSU)

2016 – 2020

- I spent six months at the Ohio State University under [Facundo Mémoli](#) supervision.
- We generalized the Vietoris-Rips and Čech constructions using “motifs”: families of graphs that can be attached to metric spaces (more generally, networks) and generate simplicial complexes. As a corollary, we obtain several clustering methods on graphs.
- The resulting work is “Motivic constructions on graphs and networks with stability results” and can be found [here](#).

### Master Degree – Algebraic Topology

Universidade Estadual Paulista (Unesp)

2014 – 2015

- I studied a certain subgroup of the fundamental group, called “Gottlieb group”, and proved some results about its properties. The dissertation can be found [here](#).

### Bachelor in Mathematics

Universidade Estadual Paulista (Unesp)

2010 – 2013

## WORK EXPERIENCE

### R/SHINY DEVELOPER

#### Appsilon

Feb 2025 – today

I work as a L4 R/shiny engineer at [Appsilon](#), a company that provides data science services to the biggest pharma companies in the world. I contributed to [open source packages](#) and other internal tasks.

### HEAD OF INTELLIGENCE

#### Argus Solutions

Jan 2020 – Jan 2025

I founded the area of Data at Argus, and my role was crucial as the company went from a small-sized tech to a multinational. My team now has more than 12 people as diverse as developers, data scientists, engineers and mathematicians. I act as a manager, code reviewer, project manager, machine learning engineer, DevOps, R/Julia developer and database administrator. I can do a bit of everything.

- We went from hundreds of Excel files to a MariaDB database on AWS. I designed the database, its indexes and applied some normal forms.
- I wrote dozens of R/Julia scripts to read and write data from several sources and created a pipeline of information that ran smoothly

- I automated the generation and mailing of thousands of daily/weekly reports using RMarkdown and Sendgrid, which previously was done by 3 people.
- I created several analytics dashboards using the shiny framework, in place of a former Tableau dashboard. I also developed the main platform used by the team to analyze photos/videos.
- We created a machine learning model that [predicts if a driver is going to sleep](#) in the next hour, using an ensemble of models with tidymodels.
- I created a webserver in Julia using [Oxygen.jl](#) to read data from several APIs and write them in our databases. A previous version of this code was written in R and each iteration took almost 4 minutes; it was reduced to 5 seconds in Julia using some parallelism magic.
- I participated in the creation of many computer vision algorithms to detect cellphones, drowsiness, potholes and gestures using keras and YOLO.
- I interviewed and hired people, managed the team and developed the abilities we needed to finish our projects.

## TOOLS AND SKILLS

### R

- [tidyverse](#): I use several of these packages everyday; I am “fluent” in dplyr, purrr, tidyr, stringr, tibble and many others. I prefer to use echarts4r instead of ggplot2. With [dbplyr](#), I can translate many complicated transformations to SQL in a painless way.
- [shiny](#): I explored a lot of the shiny ecosystem since 2018: from shinydashboard to the new [bslib](#), using [shinyWidgets](#) and [toastui](#). I can create dashboards, CRUDS, user interfaces, display photos/videos and maps. Async and parallelism are a bit painful but can be done with future and promises.
- I am also very familiar with igraph to manipulate graphs, leaflet to plot maps, http2 to make API calls, plumber to create APIs, reticulate to interface Python.
- I can create packages, documentation and tests. Unfortunately, my best packages are private but a public example can be found [here](#).

### JULIA

Coding in Julia is a joy and make R/Python look like unpolished old languages used by barbarians.

- I started learning Julia to write some Topological Data Analysis algorithms and the result was impressive, even with my initial poor skills.
- I am the owner of the [JuliaTDA organization](#) which aims to bring the Topological Data Analysis ecosystem to Julia. I wrote the [Mapper and Ball Mapper](#) algorithms in Julia and the documentation in Quarto, and later also wrote the [ToMATo](#) algorithm.
- Building JuliaTDA, I was able to get familiar with parallelism, code piracy, structs, multiple dispatch, profiling (memory allocations and algorithm efficiency) and plotting tools like Plots and Makie.

- The [book](#) “Hands-On Design Patterns and Best Practices with Julia” is my bible, and I am rewriting all my previous code using what I learned there.
- Since I like the tidyverse so much, I am now part of the [TidierOrg](#): a 100% Julia organization inspired by R data tools. I am the main contributor to [TidierIteration.jl](#), a Julia version of R’s purrr.
- Since I like Quarto so much, I created a package that helps writing Julia documentation with it: [QuartoDocBuilder](#). Its [own documentation](#) is written in Quarto!

## AWS

- I can create and manage EC2 instances, use S3 to storage and query data, create databases in RDS and create dashboards with metrics.
- I created a facial recognition system of unlabeled drivers using S3 and Rekognition.
- Recently I moved several R/Julia/Python scripts from EC2 to ECS services, providing a reproducible and isolated environment for each code. While doing this, I learned how to create images with Docker, building them with EC2 Image Builder and storing it all in ECR.

## PYTHON

- I can read code and port it to another language when necessary; I never had the need to write pure-Python code, but can learn it easily if needed. Using R’s `reticulate` and Julia’s `PythonCall.jl`, I was able to use many Python packages elsewhere.

## PUBLISHING

- I used a lot of [RMarkdown](#) to generate a wide range of dashboards, reports and analysis. Now I prefer to use [Quarto](#) because of its many cool features (sites, blogs and dashboards) and easiness to use with Julia.
- I am writing a book called [Topological Data Analysis with Julia](#) while I implement several TDA algorithms in JuliaTDA.

## TALKS, WORKSHOPS AND PAPERS

- Paper: [Remote sensing to quantify potential aquifer recharge as a complementary tool for groundwater monitoring](#) (2024, published in [Hydrological Sciences Journal](#)), co-author.
- Workshop: [Topological Data Analysis workshop](#) (2024, at the [XXIII Brazilian Topology Meeting](#)).
- Talk: [Topology meets the real world: how geometry can help us analyse finite metric spaces](#) (2023, at the [Workshop of Algebraic Topology and Applications](#)).
- Paper: [Motivic clustering schemes for directed graphs](#) (2020, published in ArXiv, with Facundo Mémoli).