

# Industry careers for physics PhDs – a data scientist's view

Kalyan Allada

# Unscientific Survey

Do you hope to pursue a career in academia ?

# Uncertain prospects for postdoctoral researchers (Nature, Dec 2020)

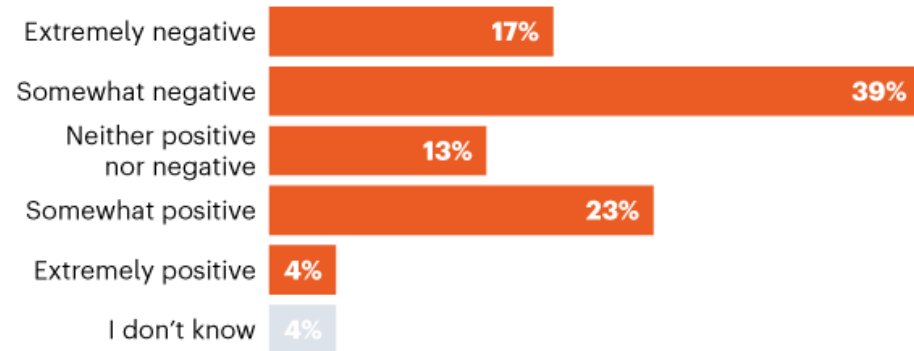
## ACADEMIC AMBITIONS

Despite hiring freezes and layoffs at universities around the world, most respondents aspire to a career in academia. And although they turn to peers and principal investigators (PIs) for career advice, they still have a negative view of their prospects. A large majority feel that previous generations of postdocs had it better.

### Do you hope to pursue a career in academia?



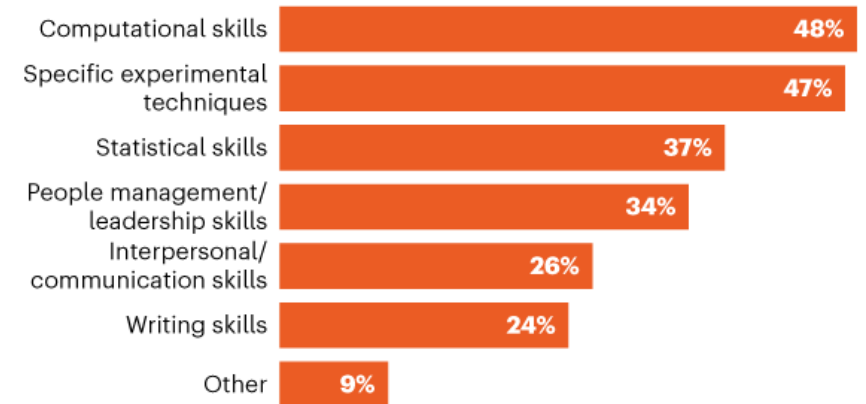
### How do you feel about your job prospects?



### What do you think is the biggest challenge for your personal career progression?



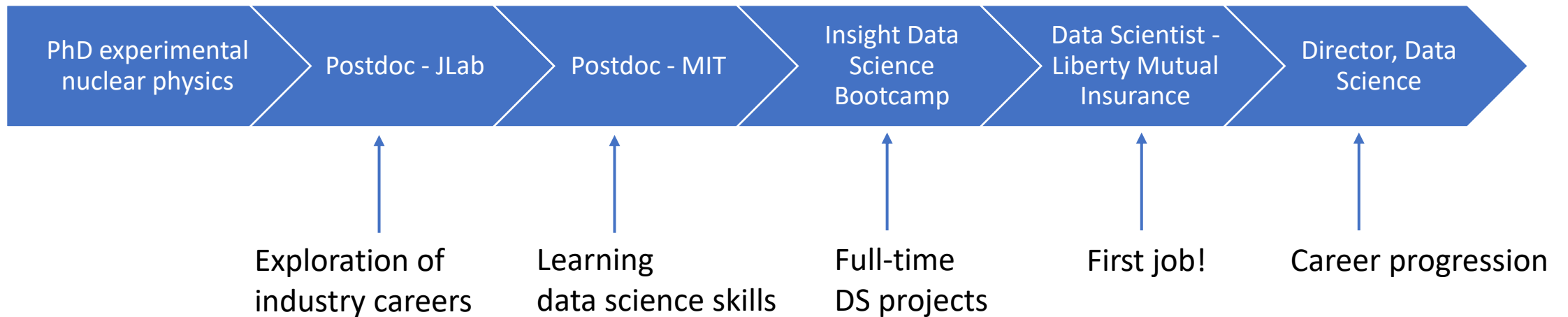
### Which skills do you feel you're lacking in?



Uncertain prospects for postdoctoral researchers, Nature, Dec 2020

<https://www.nature.com/articles/d41586-020-03381-3>

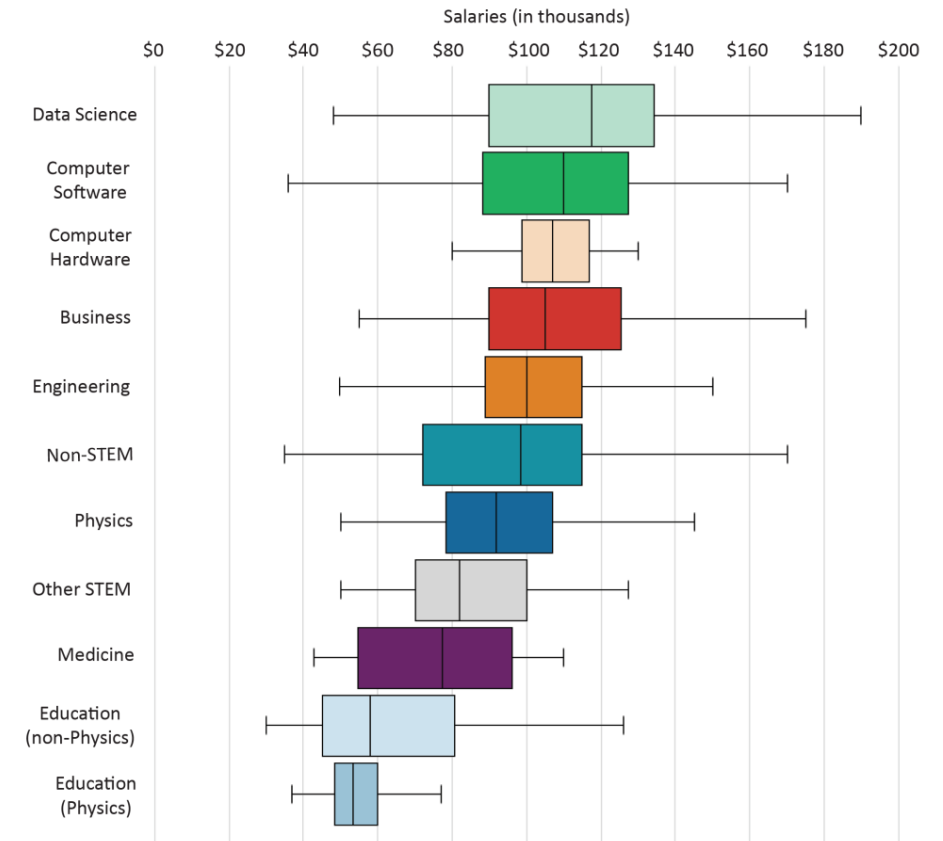
# My journey – Experimental physicist to data scientist



# Industry career options for PhD physics grads/postdocs

- Data Science
  - ML/Data engineer
  - Data scientist
  - Applied scientist
  - Research data scientist
- Software/hardware engineer
- Medical physics
- Quantitative finance
- Staff scientist at FFRDCs
- Physicist at a defense contractor

Starting Salary Ranges for New Physics PhDs in Potentially Permanent Positions, Classes of 2014 through 2018



The full starting salary range is represented by the lines extending to each side. The box represents the middle 50% (25th to 75th percentile) of salaries. The vertical line within the box represents the median starting salary in the field

Identify a career field that excites you

# I am excited about building and working with AI language models

## The New York Times

### *Meet GPT-3. It Has Learned to Code (and Blog and Argue).*

The latest natural-language system generates tweets, pens poetry, summarizes emails, answers trivia questions, translates languages and even writes its own computer programs.

## The New York Times Magazine

### **A.I. Is Mastering Language. Should We Trust What It Says?**

OpenAI's GPT-3 and other neural nets can now write original prose with mind-boggling fluency — a development that could have profound implications for the future.

## The Washington Post

### **The Google engineer who thinks the company's AI has come to life**

AI ethicists warned Google not to impersonate humans. Now one of Google's own thinks there's a ghost in the machine.

## THE NEW YORKER

MASTERPIECES JULY 11 & 18, 2022 ISSUE

### **DALL-E, MAKE ME ANOTHER PICASSO, PLEASE**

*The creators of an artificial intelligence that can produce almost any art work imaginable—from "cheeseburger lamp" to "the rest of mona lisa"—sift through their latest requests for original images.*

# GPT3 models – some of the most powerful AI systems

AI generated original image



**Prompt to AI:** *An astronaut playing basketball with cats in space in a watercolor style*

SCIENTIFIC  
AMERICAN®

ARTIFICIAL INTELLIGENCE | OPINION

## We Asked GPT-3 to Write an Academic Paper about Itself— Then We Tried to Get It Published

An artificially intelligent first author presents many ethical questions—and could upend the publishing process

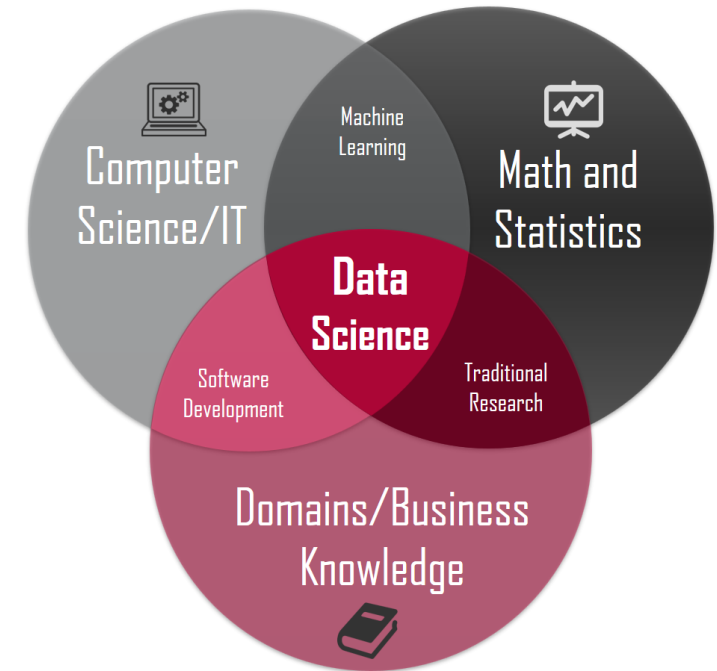
By Almira Osmanovic Thunström on June 30, 2022

<https://hal.archives-ouvertes.fr/hal-03701250v1>



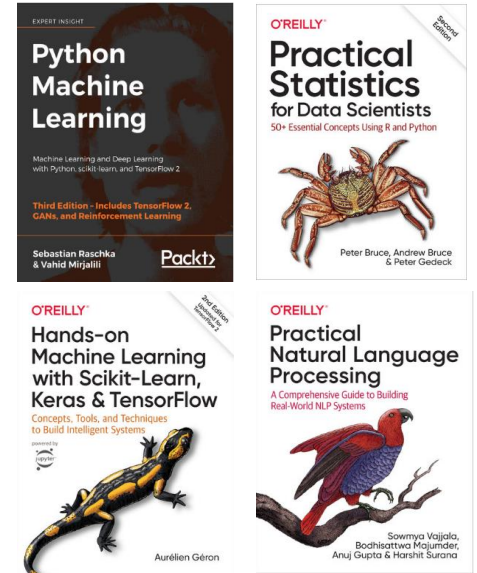
# Core qualities of a data scientist

- Good understanding of ML/DL algorithms
- Programming skills (mostly python or R)
- Business acumen to understand what problem could be solved
- Critical thinking and problem solving
- Attention to details in order to discover nuances in data
- Agile mindset – thinking iteratively to maximize impact of your work
- Excellent communication skills to explain findings to non-technical business partners
- Collaborative



# How to prepare for a data science career?

- Transition to python programming
  - Learn software engineering skills - OOPS, unit testing etc.
- Learn about ML algorithms and when to use one
- Learn about deep learning architectures and use them in your projects
  - Computer vision/NLP: Transformers, RNN, CNN etc.
- Brush-up on fundamental statistics
  - Sampling distributions, probability, hypothesis testing, linear regression, classification etc.
- Learn the DS lingo!
  - Better to have wide, but shallow understanding of many concepts, rather than deep understanding of a few concepts
    - Allows for building quick solution
  - Almost opposite of academic research!



kaggle coursera

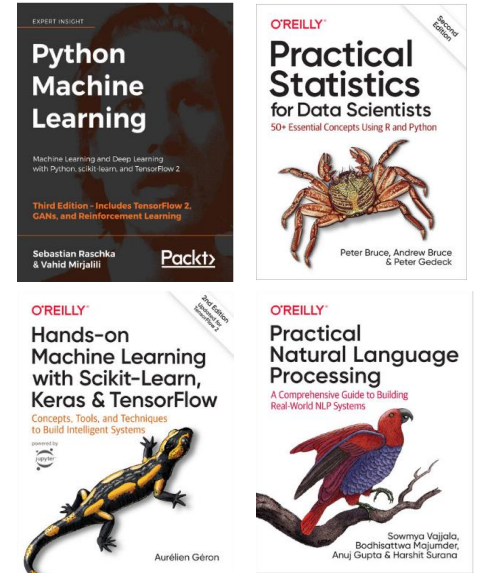
UDACITY edX

datacamp fast.ai

towards  
data science

# How to prepare for a data science career?

- Build end-to-end DS projects using publicly available data
  - Self motivated by deep interest and passion
  - Original research ideas/effective story-telling with data
  - Work with ML/DL libraries such as XGBoost/PyTorch/Tensorflow
- Create a GitHub profile and showcase coding skills and insights from your project
- Write about your project learning experience (blogs etc.)
- Find summer internship or volunteer for pro bono work
  - <https://connect.informs.org/probonoanalytics/volunteers/volunteer-process>
- Create a 1-page resume highlighting personal DS projects/ML skills



kaggle coursera

UDACITY edX

datacamp fast.ai

towards  
data science

Questions?