

# Ciência de Dados: Por que ainda precisamos falar sobre ela?

**Painel – Ciclo de Seminários PESCS  
22 / 9 / 2021**

Daniel Rattón Figueiredo  
PESCS/COPPE/UFRJ



# Ciência de Dados e Eu

- 2005, EPFL: Estudo da evolução da rede de emails
  - rede com todos os emails @epfl.ch por 1 ano
  - busca por padrões para classificar spam

- Problema de desembaranhar skis ( $k = 3$ )
  - sob nevasca e no plano inclinado



Michal Piórkowski  
Head of Data Science  
Philip Morris International

Eu!

- Primeiro contato com *Ciência de Redes*



**WARNING!**

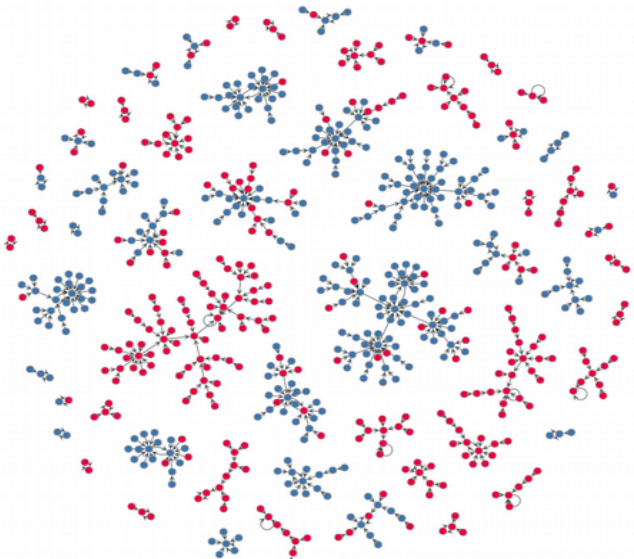
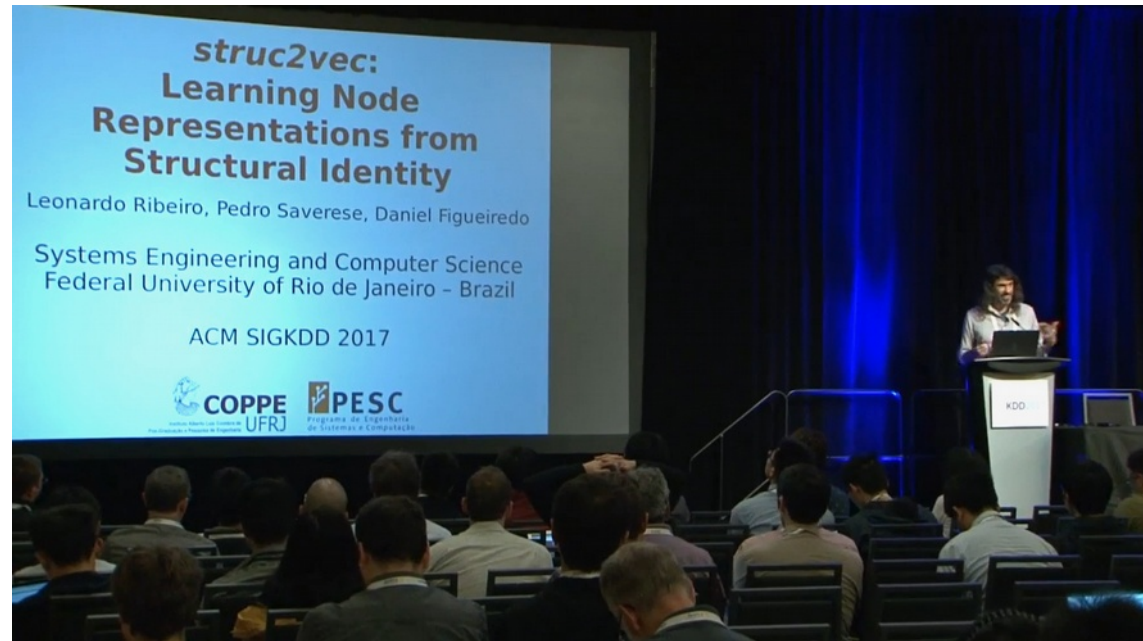
NETWORK SCIENCE IS EXTREMELY  
CONTAGIOUS. ONCE YOU LEARN IT, YOU  
START SEEING NETWORKS EVERYWHERE.

D. ZINOVIEV

# Ciência de Dados e Eu

- 2008, UFRJ: Disciplina de Redes Complexas (*Ciência de Redes*), oferecida anualmente
  - busca por leis e padrões em redes reais
  - explicação teórica para conectividade empírica

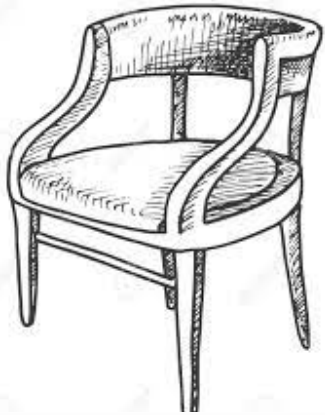
- 2017, KDD: *papel estrutural dos vértices*
  - artigo bem recebido, tese premiada CTDIAC



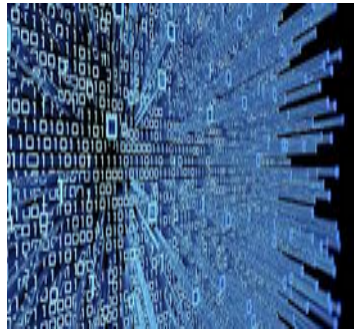
- 2021, Tese: identificação e classificação de grupos coordenados no Twitter (*bots*)

# Cientista de Dados em Ação

Produto



Informação



+

+

Cientista de Dados



=

Produto Melhor



pixtastock.com - 56717341

Statistics

Programming

Communi-  
cation

Business

Hot  
Air

The  
Data  
Nerd

The  
Hacker

The  
Stats  
Prof

R  
Core  
Team

The  
IT  
Guy

The  
Good  
Consultant

Drew  
Conway's  
Data  
Scientist

The  
Accountant

Comp  
Sci  
Prof

The  
perfect  
Data  
Scientist!

The  
Number  
Cruncher

Head  
of IT

Ana-  
lyst

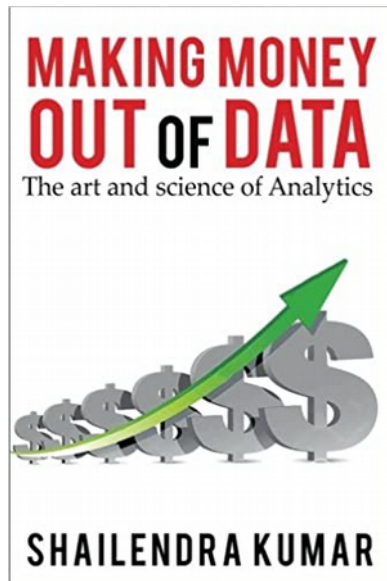
The  
Salesperson

- Gerar informação transformadora: *analytic insight* → *business insight*
- Mistura de *expertises multidisciplinares*

**Veio para ficar!**

# Ciência de Dados

## Conquistas



- Muitos casos de sucesso na indústria
- Computação assume protagonismo na sociedade da informação

## Desafios

### Big Data & Data Science Projects Failure Rate

GARTNER ESTIMATED

**85%**

of big data projects fail (2017). The initial estimation was 60% (GARTNER 2016)

THROUGH 2020

**80%**

of AI projects will remain alchemy, run by wizards whose talents will not scale in the organization. (GARTNER 2018)

THROUGH 2022

**20%**

of analytic insights will deliver business outcomes. (GARTNER 2018)

- Expectativas bem maiores que as entregas
- Formação de profissionais qualificados
- Padronização de metodologias tecnológicas