

# From Scientist to Data Scientist

Joseph Pellegrino, PhD Quantum Optics  
Chief Data Scientist @ Kernix

From Science to Data Science, May 2017

# The Arrow of Time



2011

2012

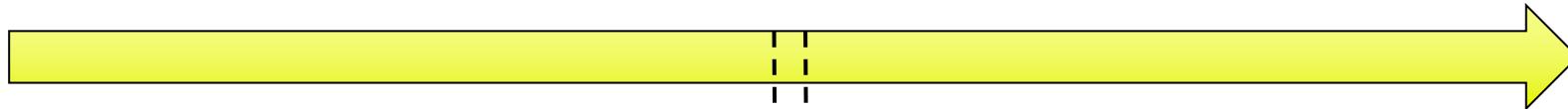
2013

2014

2015

2016

2017



Phase transition ?

PhD candidate in Quantum Physics  
@ LCF of IOGS



Data Scientist @ Kernix

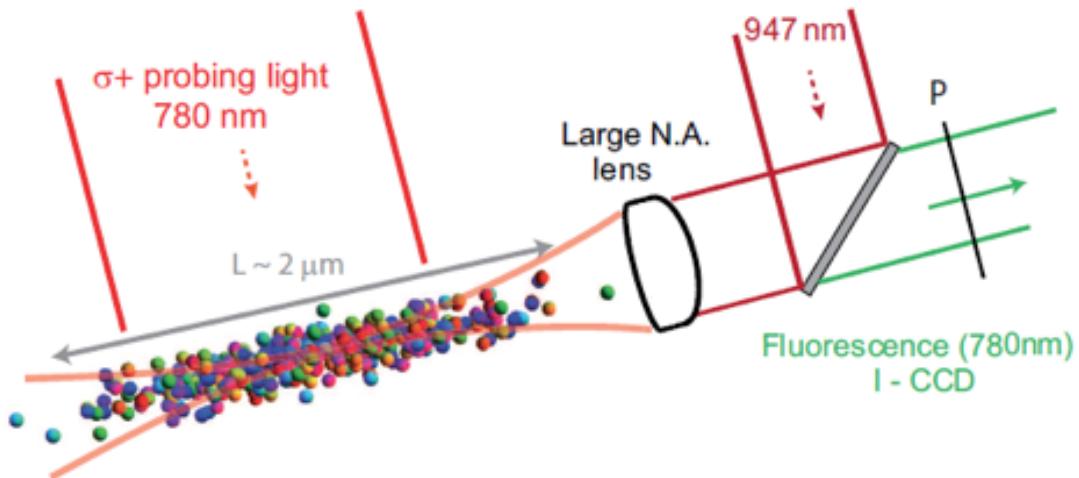


The background features a complex network graph composed of numerous nodes (small circles) and connecting lines. Some nodes are larger and filled with a medium blue color, while others are smaller and outlined in white. Dashed circles of the same color are overlaid on the network, creating a sense of depth and highlighting specific clusters of nodes.

**From Scientist ...**

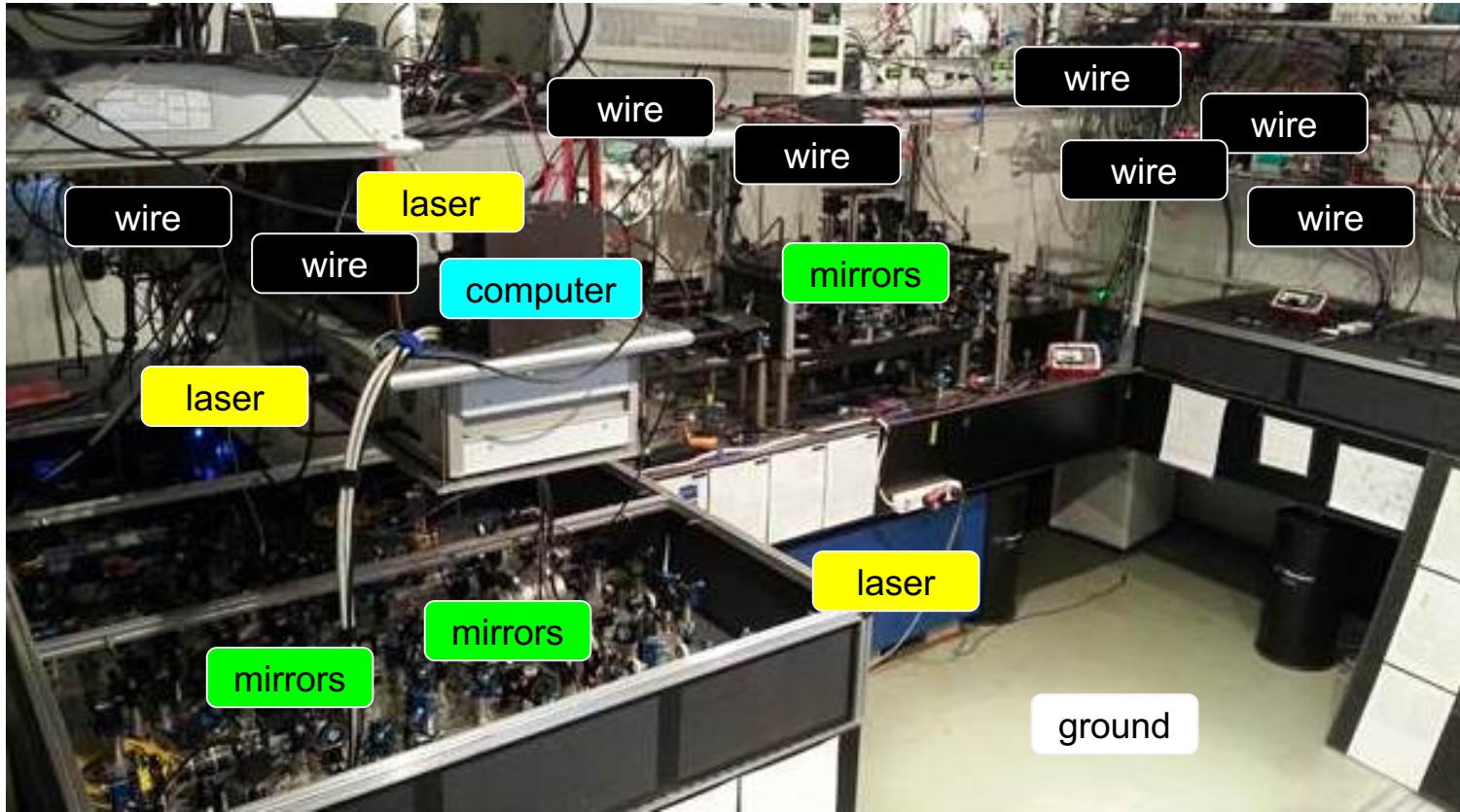


How does small and dense clouds of cold atoms scatter resonant light ?



I was looking for **collective** behavior induced by  
**interactions** between entities...

# Welcome in the cold atom lab !



# Some scientific skills



## *Experimental*

Optics (detectors, lasers, electro & acousto-optic modulator, ...)

Electronics (control systems, monitoring systems, ...)

Vacuum engineering (ultra-high vacuum chamber, ionic pumps,..)

Computer system (synchronisation of experiments, data acquisition & analysis, ...)

## *Theoretical*

Classical and quantum optics

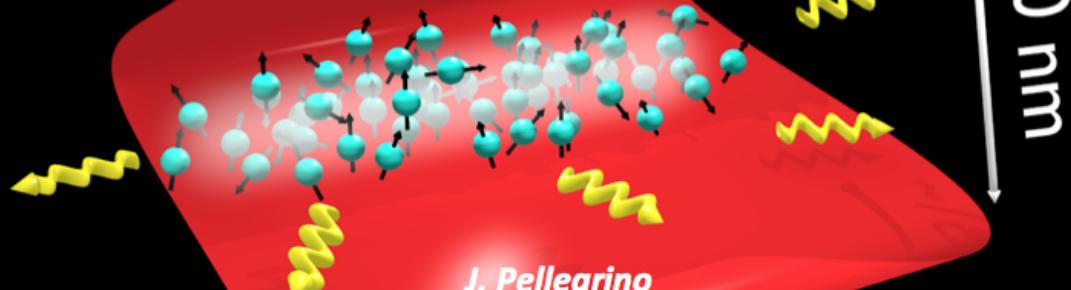
Atomic physics

Statistical physics

Signal processing



*Diffusion collective de lumière  
résonante par des nano-nuages  
denses d'atomes froids*



J. Pellegrino

*Soutenance de thèse de doctorat – 28/10/2014*

# Phase transition ?



What I love :

Understanding things

Physics & Mathematics

Solving hard problems

Building LEGO systems

Programming

Helping people to do great stuff

What I would like to do:

What I love

Something in phase with our everyday life

Something quickly useful for people

And one day, I heard about the “Big Data” trend ...

# Looking for data science positions ...





... to Data Scientist.

# Kernix: Digital Factory and Data Lab



I was looking for **collective** behavior induced by  
*interactions* between entities...





DATA LAB



AROUND THE WORLD IN A SOLAR AIRPLANE

**SOLARIMPULSE**

## Clients



**Cdiscount.com**  
**PRICEMINISTER**

**Performics**

## Collaborations



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE



**cap-digital**



## EXPERTISE

### Data Pipelines

Cop21

TerraRush

**Terra Rush**

### Predictive maintenance

ERDF



### Data Vizualisation

SolarImpulse

**SOLARIMPULSE**  
AROUND THE WORLD IN A SOLAR AIRPLANE

### Recommender systems

PriceMinister



WikiDistrict

Clickalto



HobbyStreet

### Marketing Automation

Performics

**RADIUM ONE**

RadiumOne

### Open Data

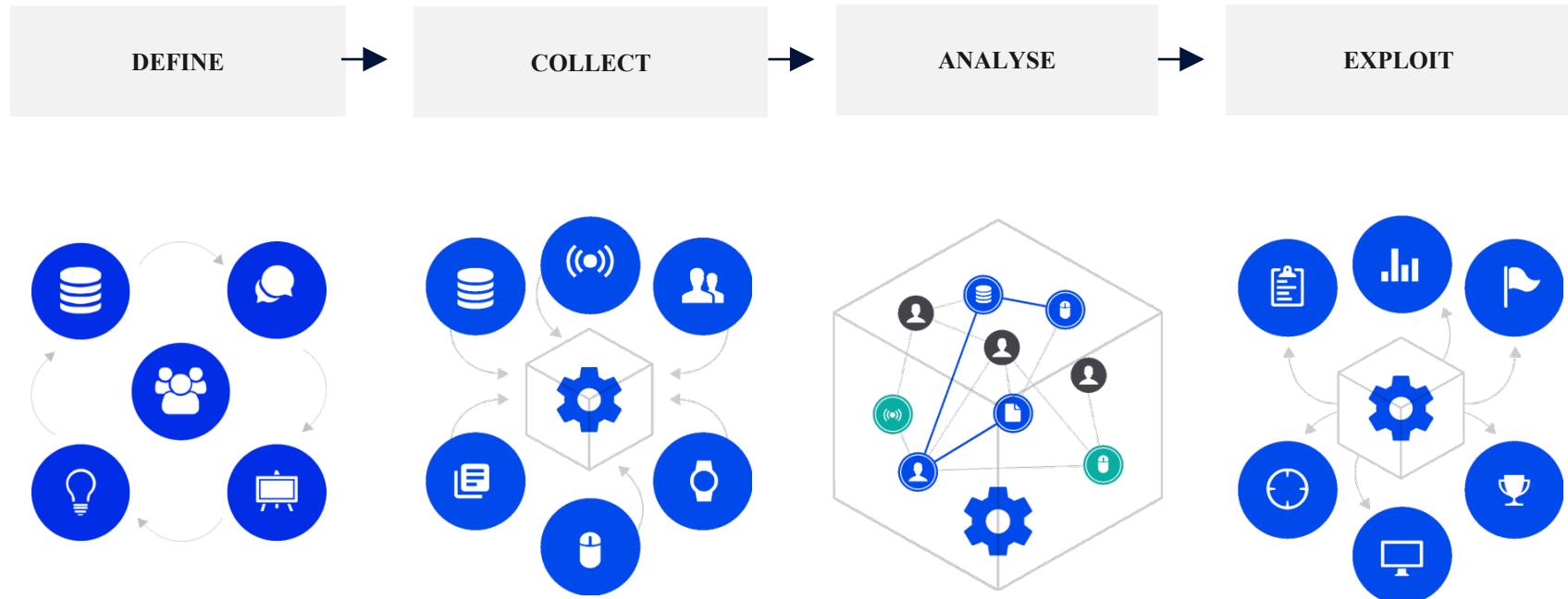
Accessible.net



**accessible.net**  
L'accessibilité pour tous, partout et en un clic

12

# Steps of a Data project



# New skills to develop



## Hard skills :

Programming languages  
(Python, Javascript)

Programming style (Object Oriented, Functional)

Databases (MongoDB, Neo4j, MySQL, Redis)

Machine learning

Image processing

Natural Language Processing

Graph mining

## Soft skills :

Project management

Business understanding

Team management

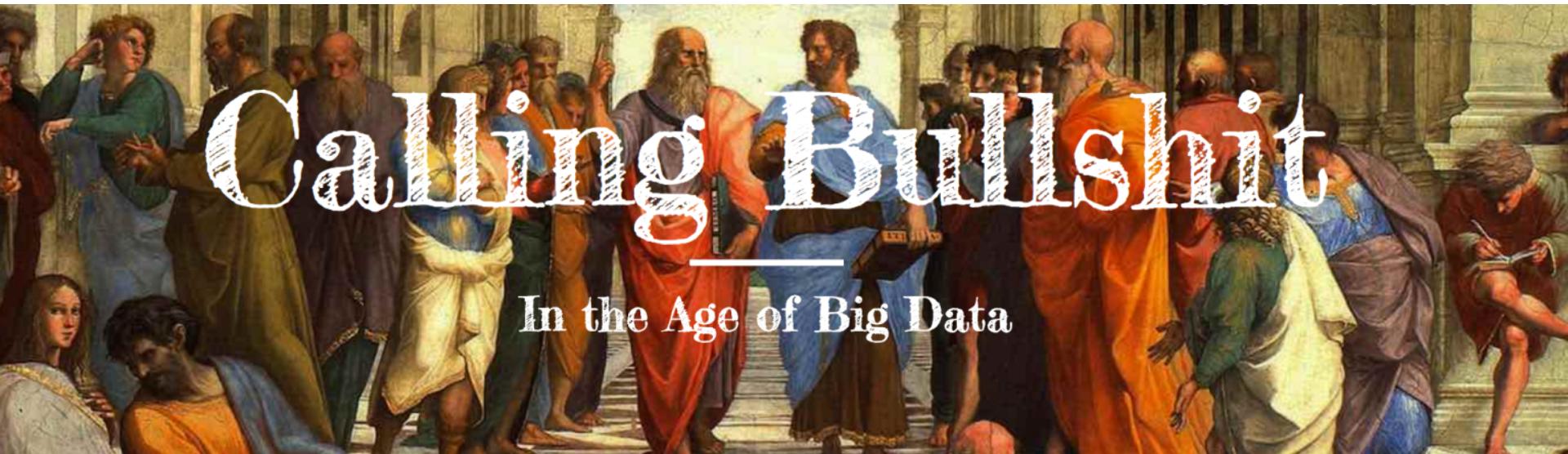


Science

Data Science

# Calling Bullshit

In the Age of Big Data



Critical and analytical thinking : in order to eradicate bullshit

# Are Data Sciences an ephemeral hype ?



Olympic Games, 1896





A network graph is displayed on a dark blue background. The graph consists of numerous circular nodes of varying sizes, some with solid blue fills and others with white centers. These nodes are interconnected by a web of thin, light-colored lines representing edges. Some nodes are surrounded by dashed circles, indicating a specific subset or focus area. The overall pattern suggests a complex system of connections and relationships.

Thank you for your  
attention