

# Industry-Driven Master Certificate in Data Science

Gianluca Reali



UNIVERSITÀ DEGLI STUDI  
DI PERUGIA



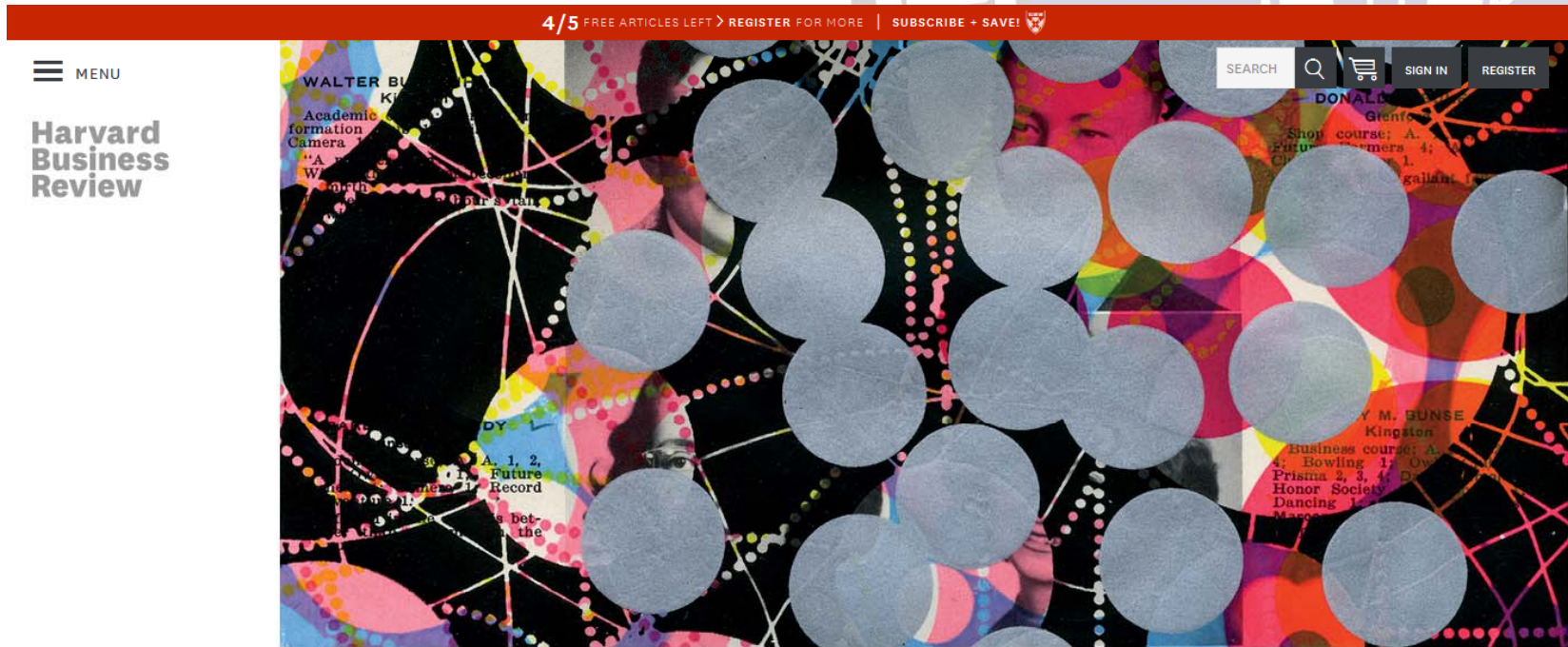
# Outline

- External Contexts
- Academic Program Design Approach
- Analysis of the LinkedIn Data
- Mix of Competences
- Academic programs at the University of Perugia
- Conclusion and Further Activities



# External Contexts

SCS



ARTWORK: TAMAR COHEN, ANDREW J. BUBOLTZ, 2011, SILK SCREEN  
ON A PAGE FROM A HIGH SCHOOL YEARBOOK, 8.5" X 12"

DATA

## Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

### WHAT TO READ NEXT

[Big Data: The Management Revolution](#)

[Where the Digital Economy Is Moving the Fastest](#)

[A Formula to Stop You from Overcommitting Your Time](#)

Bari

12-11-2015

# External Contexts

## Gartner Hype Cycle for emerging technologies



2014



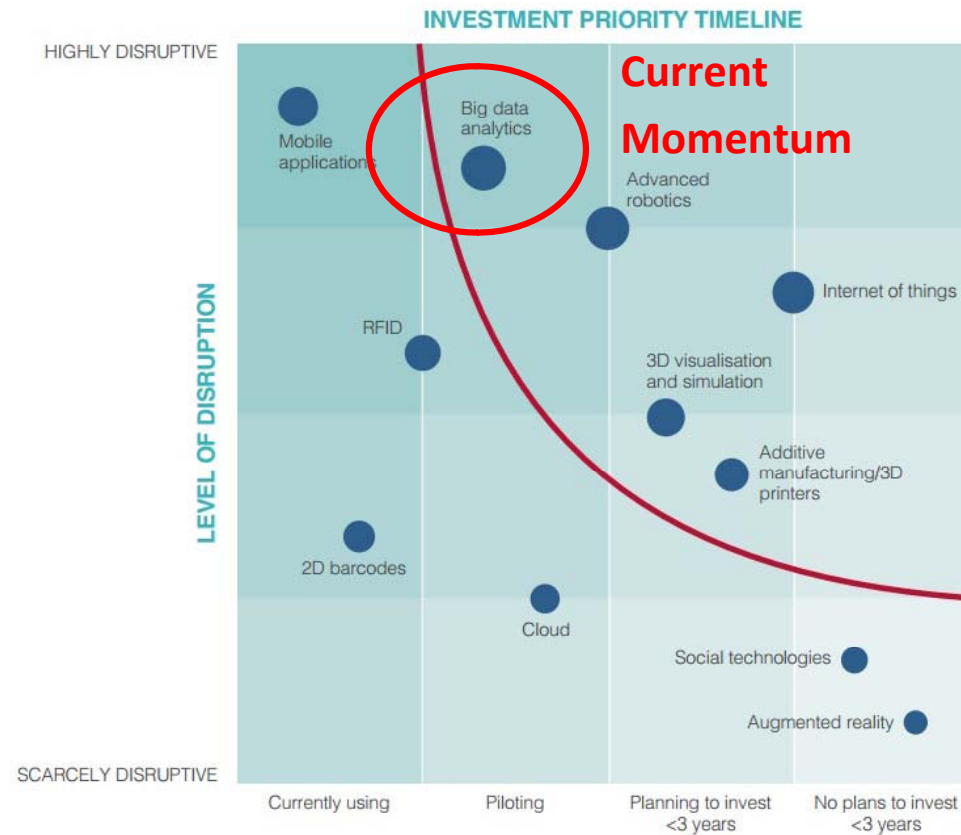
2015

# External Contexts

## Focused Market Analyses

### 8 | Technology investment priorities and their future impact on manufacturing

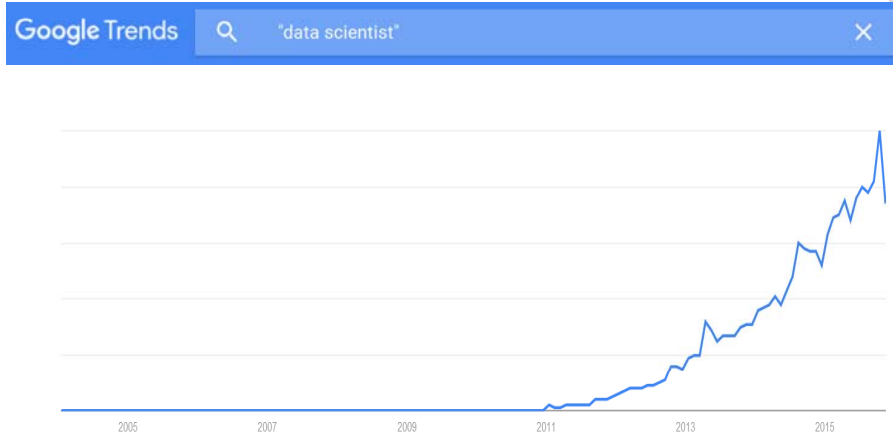
Bubble size reflects level of impact on top KPIs



**Source:** SCM World + Big Data Analytics, Mobile Technologies And Robotics Defining The Future Of Digital Factories, Louis Columbus, Forbes, **February 2015**



# External Contexts



- Retail
- Financial Services
- Healthcare
- Energy
- Telcos
- Web/Social/Media
- Government
- Manufacturing

LinkedIn search results for "Data Scientist" jobs. The search bar shows "Data Scientist" and the results section indicates "39,499 results for Data Scientist". The results list several job openings, including positions at YOOX GROUP, ClearPeaks, Tuft & Needle, and CV-Library Ltd.



# External Contexts

## Related academic initiatives worldwide

School	Degrees		
<b>Undergraduate</b>			
George Mason University	B.S.		
Illinois Institute of Technology	B.S.	<b>Certificate</b>	
Oxford University	Adv. Diploma	iSchool @ Syracuse	Grad Cert. 5 courses
<b>Masters</b>			
Bentley University	M.S.	Rice University	Cert.
Carnegie Mellon		Stanford University	Grad Cert.
DePaul University	MS.	University of California San Diego	Grad Cert. 6 courses
Georgia Southern University	M.S. 30 cr	University of Washington	Cert.
<b>Ph.D</b>			
<b>Masters</b>		George Mason University	Ph.D.
Louisiana State University <a href="http://businessanalytics.lsu.edu/">businessanalytics.lsu.edu/</a>		IU SoIC	Ph.D.
Illinois Institute of Technology			
Michigan State University	M.S.		
North Carolina State University: Institute for Advanced Analytics	M.S.: 30 cr.		
Northwestern University	M.S.		
New York University	M.S. 1 yr		
Stevens Institute of Technology	M.S.: 36 cr.		
University of Cincinnati	M.S.		
University of San Francisco	M.S.		
UC Berkeley	M.S.		



# Academic Program Design Approach

- Analysis of data
- Interaction and coordination with interested companies.
- Analysis of the existing academic programs and similar projects
- Reference profile from IWA Italy Chapter, “DATA SCIENTIST” - June 30, 2014.

<http://www.skillprofiles.eu/stable/g3/profiles/WSP-G3-024.pdf>



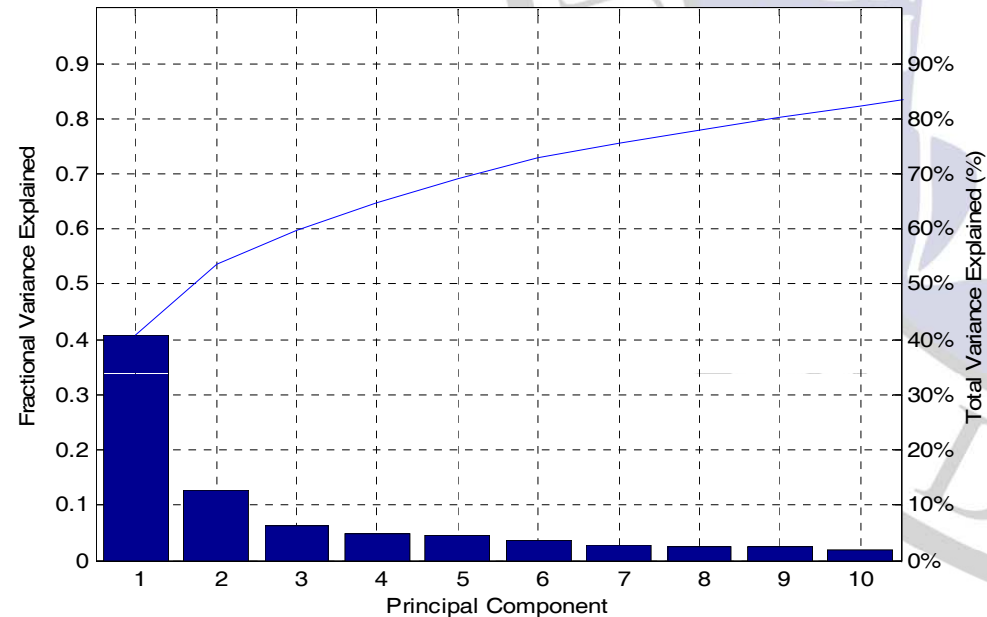


# Analysis of the LinkedIn data

- Thousands of job offers
- Data Collection and keyword identification: php, JAVA, Ruby, Machine Learning, Statistics, Information Retrieval, Graph Analysis, Big Data, Hadoop, hBase, Hive Pig, Scala, HTML, SQL, Python, MATLAB, Data Mining, Bachelors degree, BS degree, MS degree, M.S, B.S., PhD, Applied Mathematics, Computer Science, database, text mining, SAS, neural network, NLP, Computer Science, Electrical Engineering, Operations Research, Linux, Unix, Teradata, C/C++, Spark.
- Dimensionality reduction through PCA
- Extraction of the most significant competences

# Analysis of the LinkedIn data

## PCA results

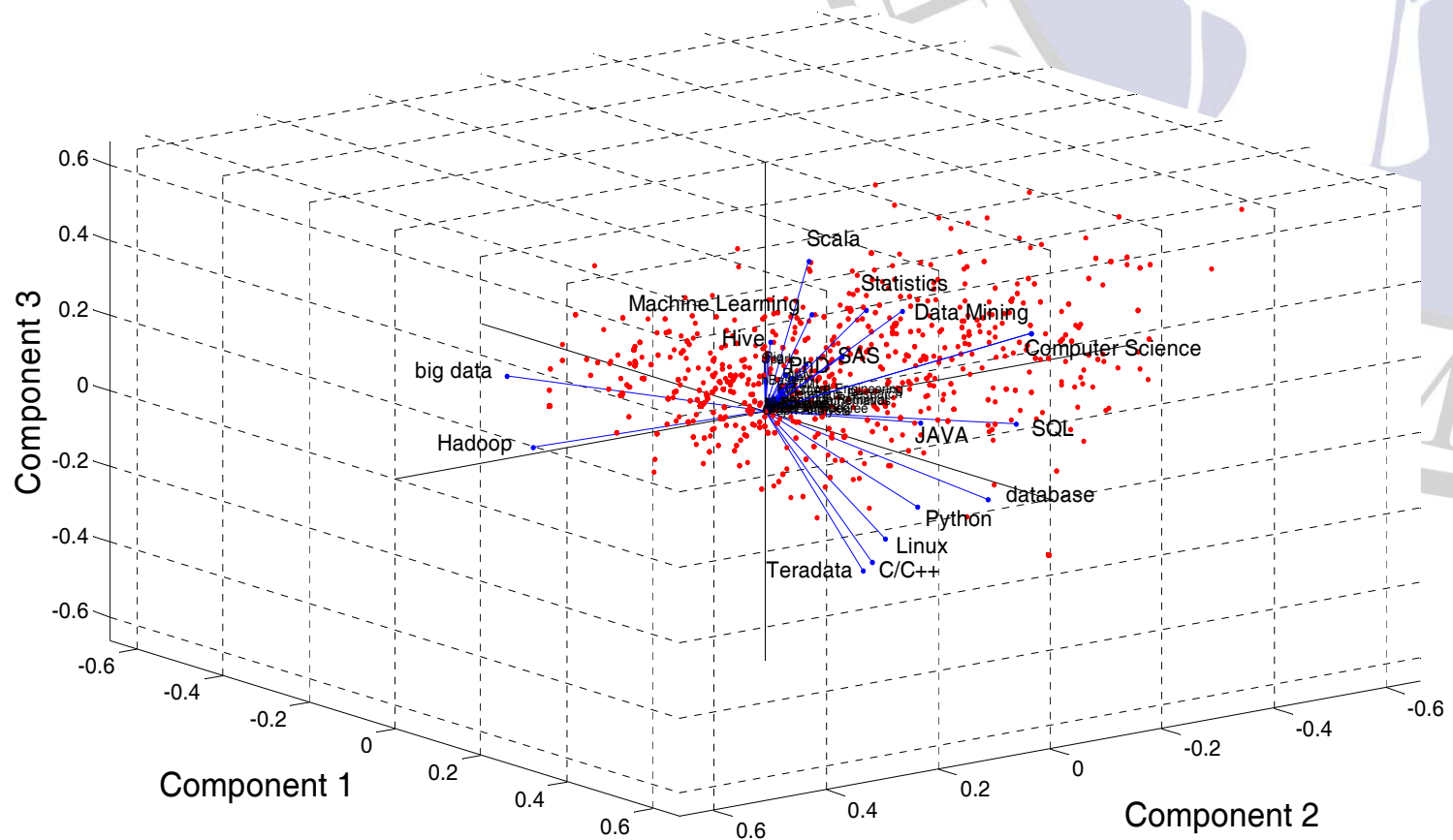


- The two most significant components cannot capture most of the process energy.
- The third component is necessary for collecting a significant percentage of energy.
- The remaining energy is still significant. It shows the heterogeneous nature of Data Science.



# Analysis of the LinkedIn data

## PCA results





## Programming & Digital Technologies

- Computational System (concurrency and distributed systems)
- Programming Languages and Paradigm (C, R, OpenMPI, Python)
- Tools (Big Data, Cloud Platforms, Databases, Sensors,...)



## Domain-specific & Analysis skills

- Data sources (Open and Linked Data) curation
- Standards and Certification for the domain
- Interpretation skills (Knowledge extraction)



## Business-orientation & Communication

- Marketing and Market Analysis (Innovation leadership)
- Legal and Ethical elements
- Data Visualisation and Communication Skills



## Maths and Statistic competences

- Statistics and Probability; Algebra and Calculus
- Machine Learning
- Data Mining And Business Intelligence



# Data Science: Academic programs at the University of Perugia

## Master of Science degree in Computer Engineering and Robotics

- **Data Science Program**
  - Focus on data engineering and analytics with applications to business intelligence
- **Robotics Program**
  - Focus on cloud robotics driven by machine learning

## Post-graduate International, Industry Driven Master Certificate in Data Science

- Focus on Finance, Energy & Utilities, Telcos, Social & Media, Industry/Manufacturing, Services
- Industry driven.



# Detailed program of the Master Certificate

Hours ECTS

	Hours	ECTS
Fundamentals of Computer Science	14	2
Fundamentals of Statistics	14	2
Big data Processing and Tools	28	4
Business Intelligence	21	3
Cloud Computing	28	4
Data Science Tools	71	8
Networking and Data Security	28	4
Data Mining and Machine Learning	52	6
Visual Analytics	40	5
Communication and Presentation Skills	21	3
Data-driven Marketing	21	3
Seminars focussed on Finance, Energy & Utilities, Telcos, Social & Media Industry/Manufacturing Services	42	6
Stage in Industries	225	9
Final project preparation	150	6

➤ Designed to be completed in 12 months,

➤ over 1500 hours of lectures, lab experiments, business focused seminars, individual and group study and project design.

➤ Lectures in English

➤ International guest lecturers, from academy, industry, service providers.



# Conclusion and Further Activities

- **Design of the program in Data Science at the University of Perugia, tailored to the Industry Needs.**
- **Corresponding PhD research activities**
- **Continuous monitoring of results and expectations**
- **Candidate to become champion in the Edison Project**  
**<http://edison-project.eu/>**
- **Open to further collaborations**