

## About Natalino Busa



Chief Data Scientist, Enterprise Data Architect and Public Speaker on Data Science and Big Data Technologies. Natalino Busa provides consultancy services for the definition, design and implementation of big/fast data solutions for data-driven applications such as predictive analytics, personalized marketing, and security event monitoring.

O'Reilly author and advocate of Open Source, and Apache Software Foundation. Passionate about Spark, Flink, Cassandra, Kafka, Akka, Hadoop. Previously, served as Enterprise Data Architect at ING in the Netherlands. Blogs regularly about big data, analytics, data science and scala reactive programming at [natalinobusa.com](http://natalinobusa.com)

### On the web:

<http://www.linkedin.com/in/natbusa>  
<http://www.twitter.com/natbusa>  
<http://www.slideshare.net/natalinobusa>  
<http://natalinobusa.com>

### Contact Data

Name: Natalino Giorgio Busà  
Date of birth: 29-01-1973  
Address: Shenton One,  
1 Shenton Way, 068803  
Singapore, Singapore  
  
Mobile: +6590078702  
E-mail: [natalino.busa@gmail.com](mailto:natalino.busa@gmail.com)

## Summary

Chief Data Scientist, Data Technologist, Innovator, and Entrepreneur with 15+ years experience in development, management and research of distributed architectures and scalable services and applications.

Implements and executes on CEO strategies and visions delivering state-of-the-art architectures, and products. Proficient at translating marketing and product strategies into customer-centric services and experiences.

Seasoned technology lead, experienced in managing multi-functional local/distributed teams, for faster execution and time-to-delivery. People coach and motivator, stimulating work-and-learn environments, and technology growth in and around the company.

Passionate about startups, scale-ups and enterprises with a strong data-driven dna such as big data, advanced analytics, machine learning, reactive programming, and api services.

Promoted and managed innovation in several companies and domains such as, personalized marketing, fraud and cyber-security, cloud and virtualization, mobile and web applications.

## Highlights

Currently focusing on Machine data-driven and real-time services and applications:

- Data science, machine learning, AI
- Streaming computing, Big/Fast Data
- API and predictive services and applications
- Distributed, fault tolerant data architectures

Summary of past achievements:

- Set up, designed and architected the very first big data exploration environment at ING,
- Designed and managed 5 big data clusters at ING on a variety of technologies (cloudera, hortonworks hadoop, cassandra, spark, elasticsearch and akka clusters)
- Designed Predictive Analytical Models for Free-to-Spend Money, Total Lifetime Value
- Produced several propensity models on: churn, propensity, default, premium
- Defined a new data architecture architecture for the fraud and cyber security domain enabling faster response, advanced analytics and detection pipelines.
- Established Technology Strategist in the Data-Driven Domain,
- Invited speaker on 10s of conferences and events,
- Domain expert on Big Data, AI, and Data Science

## Experience

### **SVP, Enterprise Data Platform Architect, DBS** **September 2017 - Present**

Responsible for defining and driving the next generation platform for BI, Reporting, Analytics, Data Science and AI. Manage directly a group of architects, data modelers, and lead engineers executing the DBS group level data-first technology program. Define a unified strategy and operation model for data practices at DBS at group level across 18 different country domains. Evolve the current data architecture from on-premise to a multi-cloud infrastructure. Define metadata management, data governance and data access control methodologies. Bring together data science, ai, and etl processes in a single coherent framework by integrating vendor solutions, in-house engineering and open source tools.

### **Head of Applied Data Science at Teradata Benelux** **April 2016 – June 2017**

As Senior Solution Architect, I have provided consulting on Strategy, Business scans, Roadmap, Gap Analysis, and Technology Mapping for AI, ML, Fast and Big Data Systems. Always hands-on. Build demonstrators and proof of concept/value on data-driven applications. Engagements with most Belgian

and Dutch banks, and Telecom companies.

Promote Open Source Technologies to Teradata customers and within the Teradata organization. Mentor and grow a community of data scientists at Teradata. Speaker and Author on Data Science and Big Data.

**Author at O'Reilly Media**  
**Oct 2015 – present**

Authoring education material, such as articles, videos, tutorials on Big Data, Fast Data, Spark, Machine Learning, AI and Data-Driven Applications.

**Enterprise Architect at ING**  
**February 2012 – March 2016**

Tech lead and scrum master of a team of data and technology senior analysts, programmers, engineers, and statisticians. Exploring and Implementing first-of-a-kind big data solutions aiming to unravel predictive banking and personalized customer intimacy. Transforming services into experiences using data science, statistical languages, distributed computing, and dataflow architectures.

Explore and prototype innovative solutions for data analysis with a particular emphasis on data science, fault tolerance and distributed streaming processing. Provide advice and critical analysis for the selection of strategic and novel technologies and data processing platforms. Public speaker and tech community builder.

**Projects at ING:**

Streaming analytics: Coral is a real-time analytics and data science platform. It transforms streaming events and extract patterns from data via RESTful APIs. Built on Spray, Akka and Spark.

Real time Predictive Analytics: Define a new vertical architecture which enable the bank IT backend with real time predictive insight, using Spray, Akka, Scala, and R. Built on top of Big Data analytics as a real time extension to the already deployed distributed R framework.

Hadoop and distributed R: Rather than hyping big data, we rather put the framework on the workbench and measure how it performs. This project makes use of various big data technologies such as hadoop, hive, R on map-reduce, etc. This environment is meant as a workbench for testing and rating various applications and solutions ranging from predictive forecasting to product/customer classification, aiming for new and improved services and a better and more personalized user experience.

Redefine Fraud and Cyber Security with new components for streaming analytics for anomaly detection and pattern recognition of web traffic, internal threats and card/account usage.

**Lead Software Developer / Web Services Architect at Civolution**  
**October 2011 – January 2013**

Design, architect, configure, prototype and test new products and applications related to video/audio

fingerprinting, detection and monetization. Architectural lead for a scalable distributed hadoop like web service. Technology scouting for high-throughput, high-load, complex web services.

**Engineering Lead at Sagantec****July 2006 - September 2011**

Managing a group of seasoned/senior SW designers. Later managing the full operations of the european office, leading the full SW development cycle, from concept to release. Bridging the development team with Sagantec's sales, support and marketing teams.

Periodically interacting with key customers in Japan and US. Senior software engineer and product owner for Sagantec's analog migration and compaction tools. Leading an engineering team for the development of innovative EDA tools and features.

**Senior Researcher at Philips Research****March 1998 - January 2006**

Leading a 8 man project on advanced software techniques for video compression, aiming at low-power and predictive techniques for a better quality on band and power –limited hand-held devices.

Researching new architectural solutions and compilers for DSP systems. High Level synthesis of code to hardware specification and RTL / gate level design.

## Education

Università di Catania

Master Degree, Electronic Engineering, 1991 - 1997  
Catania, Italy

TiasNimbas Business School

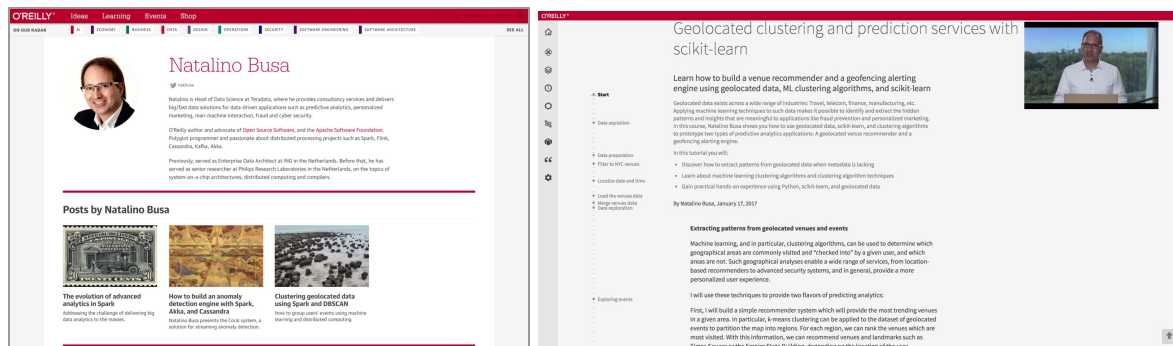
Attended: Master in Business Marketing, 2003 - 2004  
Eindhoven, The Netherlands

## On the Web

O'Reilly Author: <https://www.oreilly.com/people/b7f1d-natalino-busa>

Big Data and Data Science: <http://www.natalinobusa.com>

Databases and AI: <http://www.odbs.org/tag/natalino-busa/>



## Streaming Analytics

read my last article here:

<http://www.odbs.org/2017/01/streaming-analytics-for-chain-monitoring/>

## The “SMACK” stack

watch my last presentation here on the topic here:

<https://www.youtube.com/watch?v=Aeg5yEBuqGM>

## Geo-located Services and Data Science

Live tutorial here:

<https://www.safaribooksonline.com/oriole/>

## Image and Video AI:

To be seen at Strata London

<https://conferences.oreilly.com/strata/strata-eu/public/schedule/detail/57661>