

Fabio BENEDETTI

PERSONAL DATA

PLACE AND DATE OF BIRTH: Pavullo nel Frignano (MO), Italy | 20 November 1985
ADDRESS: Via San Felice 22, 40122 Bologna, Italy
PHONE: +39 3391321589
EMAIL: fabio.benedetti@unimore.it
PERSONAL PAGES: [Linkedin](#) - [DBgroup](#) - [GitHub](#) - [Google Scholar](#)

WORK EXPERIENCE

- 15/11/2017 - Current* | Research a Development - Data Scientist at UNIPOL SAI
My job comprehends design and development of data centric prototypes for improving the business processes inside our group. These prototypes involve scalable infrastructures, given the **high volume of data**, **machine learning** techniques, and the use of innovative ways of **data visualisation**.
- 01/01/2013-14/11/2016* | Ph.d on ICT and Postdoctoral Fellow at UNIVERSITY OF MODENA AND REGGIO EMILIA
My research topic were **Linked Open Data** and **Big Data**. In these years I worked on LODeX, a tool able to automatically provide a high level summarization of a LOD dataset and a powerful visual query interface to support users in querying the dataset through a **Web Application**. During my Ph.d, I also cultivated my passion for **Data Science**, **Data Manipulation**, **Statistics** and **Data Analysis** following classes and working with real Datasets and realizing small projects.
As postdoctoral fellow I am developed novel techniques of inter-document similarity and **Information Retrieval** leveraging information contained in **Knowledge Graphs**.
- 08/08/2011 - 14/11/2016* | Computer systems designer and analyst at DATARIVER S.R.L, Modena
Internship 08/08/2011 to 27/11/2011 - Collaborator 30/11/2012 to 14/11/2016
www.datariver.it
During my internship I reengineered an existing application called RTM, used by the cancer registry of the province of Modena. I used **Talend**, an **ETL** software, to migrate the existing Database from **Microsoft SQL Server** to **Postgres**. After that, I developed the infrastructure of the new **Web Application** using the **J2EE** stack. In particular, I used these libraries and tools: **Maven**, **Hibernate**, **Spring** and **GWT**.
As collaborator I'm contributing in the development and customization of **OpenClinica**, an Open Source Java **Web Application** used for managing clinical trials.
- 15/05/2015-01/12/2015* | Research Visitor at KNO.E.SIS RESEARCH CENTER, DAYTON (OH), USA
knoesis.org
I spent a period abroad as a research visitor at the Ohio Center of Excellence in Knowledge-enabled Computing (Kno.e.sis) at Wright State University. During this period I developed a new technique called Context Semantic Analysis for estimating the similarity between document leveraging information contained in **Knowledge Bases** and I worked in a project where we used **Spark** for analyzing traffic sensor data.
- 09/08/2012 - 30/11/2012* | Business Intelligence at DOXEE S.P.A., MODENA
www.doxee.com
In Doxee I worked in the Quality Assurance department. My primary task was the development of reports, using **Pentaho** for the visualization and standard **SQL** for extracting data from **Oracle Databases**. During my collaboration I also solved some **Data Integration** problems using the platform **Pentaho Data Integration** as back-end of the reporting system.

EDUCATION

JULY 2014 **Web Science Summer School, Southampton University, Southampton, (UK)**
The School has been characterized by mixture of invited talks and tutorials on several aspects of data science, including data analytics, data publishing and interlinking, social media management, Open Data, and crowd-sourcing.

JULY 2012 **Master degree on Computer Engineering**
Department of Engineering “Enzo Ferrari”, Modena
Thesis: “Reingegnerizzazione di un framework per la gestione ed analisi di dati clinici nella forma di applicazione web database”

APRIL 2007 **Bachelor degree on Computer Engineering**
Department of Engineering “Enzo Ferrari”, Modena
Thesis: “Sviluppo di applicazioni Web attraverso il framework Ruby on Rails”

CERTIFICATES AND SPECIALIZATION COURSES

[M101J: MongoDB for Java Developers - M101P: MongoDB for Developers](#)
[Data Science - Coursera Computer Science & Engineering University Of Washington](#)
[Health Informatics in the Cloud - Coursera Georgia Institute of Technology](#)
[Machine Learning Stanford University on Coursera](#)
[Statistical Learning - Stanford Online Course](#)
[EdX Verified Certificate for Scalable Machine Learning with Spark](#)
[Learning Python for Data Analysis and Visualization](#)
[Deep Learning on Udacity](#)
[Functional Programming Principles in Scala](#)

LANGUAGES

ITALIAN: Mothertongue
ENGLISH: Fluent

JOB RELATED SKILLS

Languages: JAVA ^(high), PYTHON ^(high), JAVASCRIPT ^(high), SCALA ^(basic), R ^(basic)
Databases: SQL (MySQL, Postgres), NOSQL (MongoDB), GRAPH DATABASE (NEO4J, Virtuoso),
INDEXING (Elasticsearch, Lucene)
Web Application: J2EE, SPRING, HIBERNATE, GWT, TOMCAT, DJANGO, TORNADO
GUI: JQUERY, POLYMER & WEB COMPONENTS, D3.JS,HTML5 ^(basic), CSS ^(basic)
Data science ^{Python}: PANDAS, SCIKIT-LEARN, NUMPY, SEABORN, MATPLOTLIB, NETWORKX, SCIPY
Big Data: SPARK, PYSPARK, HADOOP ^(basic)

INTERESTS AND ACTIVITIES

Technology, Open-Source, Programming, Reading
Squash, Running, Traveling