DATA SCIENTIST AND DATA ENGINEER POSITIONS

at Alleanza Contro il Cancro (ACC)

ALLIANCE AGAINST CANCER

A network of ~50 Italian Research Hospitals (IRCCS; Istituti di Ricovero e Cura a Carattere Scientifico) coordinated by the Italian Ministry of Health is seeking motivated data engineer and data analyst for a coordination position in the context of a large collaborative project aimed at creating an IT-platform for the sharing, storage and analyses of health data (Health Big-data project). The successful candidates will work in tight collaboration with the IT and data scientists of the participating IRCCS and of Politecnico di Milano, which is partner of the project.

Scientists interested in this position will find detailed instructions on how to prepare and submit their application following the link **https://www.alleanzacontroilcancro.it/**. For further info please write to **careccia@alleanzacontroilcancro.it**. *Deadline is October 26, 2020.*

General goal of the project is the creation of a technology platform to collect, share and analyse clinical and scientific data from patients of each of the participating IRCCS.
The project is managed by the three largest Italian networks in oncology, neuroscience and cardiology **(Alleanza Contro il Cancro; Neuroscienze e Neuroriabilitazione; Cardiologia)** and supervised by the Ministry of Health.

Specific aims are:

1 Strengthening of the IT platforms of the participating IRCCS, guaranteeing mining, integration and interoperability of clinical and scientific data;
2 Creation of a centralized IT platform, guaranteeing connectivity among IRCCS and analysis of shared data. Different types of data will be collected, including omic and EMR data, clinical imaging and radiomics, patient-reported data. Main emphasis will be on mining of non-structured data and AI-based data analytics. We are now seeking motivated scientists to coordinate the IT-infrastructure and data-analysis activities of this project. The candidates will be based at one of the IRCSS of the three networks. This is an excellent opportunity to conduct cutting-edge IT- and Data-science research and to pioneer the emerging field of health big-data collection, storage and analysis.

REQUIREMENTS

Data Scientist: The following skill-set is highly desired:

1 Knowledge and experience in statistics, data mining and machine learning;
2 Proficiency in programming languages such as Python, R, C/C++;
3 Experience in the use of HPC platforms;
4 Experience of working with the main data analysis Python/R libraries: scikit-learn, numpy, pandas, CARET, mlr etc;
5 Knowledge of Deep Learning frameworks such as Tensorflow, Keras and Pytorch is considered a plus;
6 Experience in team working in a geographically distributed and multidisciplinary team;
7 Strong interest in the application of data science to the healthcare domain.

Data Engineer: The following skill-set is highly desired:

1 Modeling and development experience in data management and distributed architectures and use of frameworks such as HDFS, Kafka, Cassandra, Kubernetes;
2 Experience in the development of applications for the management and scalability of Big Data Projects;
3 Experience of software development in collaborative environments (Jira, Confluence, GIT);
4 Experience with continuous integration, unit-testing, integration-testing, infrastructure as code;
5 Experience in writing and maintaining documentation for data management and software development;
6 Experience in team working in a geographically distributed and multidisciplinary team;
7 Experience in team management and basic experience in project management;
8 Strong interest in the application of big data solutions to the healthcare domain.