

Data Science NLP internship

Position Type: Intern

Team: Data Science Team

Internship Subject:

Data library indexation through datasets classification

Duration: 6 months

Description:

nam.R's Data Library contains an extensive collection of datasets of open data over a wide range of topics. Further exploitation of these resources require that the organization of this library be overhauled and re-engineered.

The intern's main task will consist in exploiting existing metadata information from these datasets to classify and sort the Data Library. We are specifically, though not exclusively, envisioning Machine Learning (ML) and Deep Learning (DL) approaches, with a narrower focus on Natural Language Processing (NLP) techniques.

The final output of this classification will be used to build a new indexation for the Elasticsearch engine that powers nam.R's Data Library. Depending on the intern's qualifications and experience, this second step may or may not be included in their workload.

Please send your applications to the following email address: recrutement@namr.com