



## WEB CONTENT MINING FOR INNOVATION INDICATORS AND METRICS

From early stage research to market introduction, innovation goes through many development phases. Today, methods for producing innovation indicators rely on the traditional outputs of these phases: scientific publications, patent applications, clinical trials, product launches, etc. Some indicators are also based on business and financial data (R&D expenditures, mergers and acquisitions, etc.) but in both cases the data sets are affected by time-lags and do not adequately cover emergent phenomena.

This PhD project aims to tackle the problem of automatically identifying and dynamically classifying corporate online data into innovation relevant categories, and to proactively mine information which is directly or indirectly linked to innovation activities. To achieve this, using Big Data tools and methods, new web mining algorithms to identify useful innovation-related information from different online channels have to be proposed. Metrics and indicators capturing the different innovation aspects and phases at the corporate level should therefore be designed. The outcome of this PhD project will empower the Expernova platform to provide unique innovating profiles of corporate companies in a « right-time » business intelligence context.

Candidates should hold a MSc (or equivalent) with a background in Machine learning, statistical modelling and natural language processing. Programming skills and software development experience would be beneficial.

Expernova SAS

503 315 137 RCS Montpellier | [contact@expernova.com](mailto:contact@expernova.com) | +33.(0)4.67.65.54.41

Siège social : 1 Place Francis Ponge, 34000 Montpellier