# FAO_black_20

# Food and Agriculture organization of the United Nations

### **Terms of Reference for Consultant Category B \***

|  |  |
| --- | --- |
| **Name:** |  |
| **Job Title\*\*:** | NLP and Text Mining expert |
| **Division/Department:** | Statistics Division, ESS/ES Department |
| **Programme/Project Number:** |  |
| **Duty Station:** | HQ (Rome, Italy) |
| **Expected Start Date of Assignment:** |  | **Duration:** |  |
| **Reports to:** | ***Name:*** |  | **Title:** | Research and Innovation Team |
| \* Please note: If this TOR is for Consultant / PSA.SBS contract, the minimum relevant experience required **for the assignment** is as follows:

|  |  |  |
| --- | --- | --- |
| 1 year for a category C | 5 years for a category B | 12 years for a category A |

\*\* Please enter a short title (max 25 chars) for this assignment. |
| General Description of task(s) and objectives to be achieved |
| Statistics is a core function of FAO and represents a highly visible area of the Organization’s work. The goal of FAO’s Statistics Division (ESS) is to provide timely and reliable data on hunger, food and agriculture to facilitate the design and monitoring of evidence based policy decisions by member countries, as well as serve as the foremost authoritative source of statistical standards and state-of-the-art methods in agricultural and food statistics. Within the Statistics Division, the Research and Innovation team is responsible for the development of methods, standards, tools and norms used mainly within the organization with the objective to improve the efficiency of FAO’s statistical system and the quality of its output. Rapid technological development requires ESS to innovate in a variety of areas to modernize the statistical business process and meet the increasingly demanding needs for fast, accurate, easy and cost-effective data and analysis. Therefore, it is part of FAO’s strategy to engage with non-official, Big Data sources and to rely on data science methods to solve the current information gaps problems. The final objective is to expand the quantity, quality and range of the statistical and analytical products of the division. Within the Research and Innovation team, a Data Lab for Statistical Innovation has recently been established to lead the Division’s work related to data science applications including Natural Language Processing (NLP) and text mining with the objective to solve research problems in agriculture statistics and policy (e.g. policy analysis, use of statistics in policy making) by leveraging existing information. The text mining expert will join this new Data Lab.Under the overall supervision of the Director, ESS, the direct supervision of the Senior Statistician and in collaboration with the concerned technical division, the primary responsibility of this position is to employ big data and data science techniques across a range of solutions. The specific roles and responsibilities include, but are not limited to the following:* Identify suitable text sources;
* Develop solutions using models for semantic understanding and text similarities applied to agriculture policy documents or other relevant topics for the organization;
* Designs, builds, and evaluates multilingual NLP algorithms that are accurate, scalable, and production-quality by applying statistical, rule-based, and knowledge base NLP techniques
* Designs and develops NLP applications that leverage interactive machine learning for AI-assisted chart review and monitoring services, machine active learning and self learning
* Be proficient in quickly developing models, as well as taking prototypes to full production;
* Perform operationally-relevant data analysis
* Publish well documented, reproducible work;
* Contribute with the ESS analytical and visualization products;
* Any other duty as required.

Education: * Master's degree in Computational Linguistics, Engineering, Bioinformatics, Physics, Statistics or related field with strong computational elements.

Core skills:* Experience in two or more of these or similar Natural Language Processing topics: word categorization and tagging, syntactic parsing, word sense disambiguation, topic modeling; contextual text mining, application of machine learning to NLP, semantic similarity, phrasal semantic analysis, text matching and similarity, word embedding, lexicon normalization, named entity recognition
* Experience with data mining, and analytics techniques
* Experience in programming in R, Python or similar languages,
* Experience with NLP libraries such as NLTK, openNLP, Stanford-NLP, WordNet or other NLP software.
* Experience with data visualization and presentation, turning complex analysis into insight
* Ability to make recommendations and advise decision-makers on highly technical issues;
* Focus on delivery and on meeting tight deadlines.
* Highly collaborative and able to work with individuals from many different teams
* Good writing and editing skills, with a strong command of English and an ability to convey complex ideas in a clear, direct, and lively format.
 |
| key performance indicators |
| Expected Outputs: | Required Completion Date: |
| * Full production solutions developed NLP and text mining techniques in areas of agricultural statistics and policies with its set of tested automatic processing routines
* Documents and manuals on the solutions properly published
 |  |