

## REQUEST FOR EXPRESSION OF INTEREST (REOI)

### Financial market monitoring via social media and web extraction

**Project:** Development of a working prototype for market monitoring via social media (the "Project") for a financial authority (the "Agency").

**Description:** Advanced market monitoring tools that extend the existing capability of the authority, using machine learning models to go beyond basic sentiment analysis to surface trends, anomalies, and other patterns in the data that is relevant to monitoring of entities and activities flagged through other processes.

**Contracting Entity:** University of Cambridge, Judge Business School

**Countries and Agencies:** TBC before vendor selection and contracting

**Grant Value:** US\$100,000

**Publication Date:** 24 February 2023

**Expression of Interest Deadline:** 09 March 2023 23:59 GMT Time (UTC +0)

**Project Implementation Dates:** May 2023 – November 2023

**Procurement Process Managed by:** Cambridge SupTech Lab at the Centre for Alternative Finance, the University of Cambridge Judge Business School

**Submission:** Email all documents to the Cambridge SupTech Lab's Launchpad at [suptech-launchpad@jbs.cam.ac.uk](mailto:suptech-launchpad@jbs.cam.ac.uk) with subject line "REOI: Cambridge SupTech Lab – Web-Based Market Monitoring Project"

**Language:** All submissions must be written in English.

### The Cambridge SupTech Lab

The Cambridge SupTech Lab ("the Lab") at the Cambridge Centre for Alternative Finance, the University of Cambridge Judge Business School, accelerates the digital transformation of financial supervision.

While financial services are becoming increasingly global, digital and complex, analogue processing and antiquated technologies in data gathering, validation, storage and analysis erode the analytical capabilities of supervisory agencies, who are often too late in protecting consumers from fraud and seeing signs of stress in the financial system, or miss the underlying causes. This is all happening while financial crime remains a trillion-dollar issue, and public agencies face new challenges such as the regulation and supervision of crypto assets, and monitoring environmental, social and governance (ESG) aspects of the financial industry's business.

The Lab aims to meet financial sector supervisors' needs by developing with them new methodologies and processes that further market oversight and empower consumers, and to deploy suptech applications that generate relevant, reliable, timely insights to inform their decisions. From research to executive education, to technical assistance, to crafting production-grade suptech solutions, we are committed to supporting the emergence and acceleration of the suptech ecosystem and to empowering a new generation of innovation leaders seeking to digitally transform financial supervision.

For more information about the Lab, please visit <https://www.cambridgesuptechlab.org>

### The SupTech Launchpad

To accelerate the growth of the suptech marketplace, the Lab partners with financial authorities and technology vendors to co-create and deploy cutting-edge, scalable suptech applications. Our team helps detail the technical specifications, de-risk procurement for all parties, and provide project management support and hands technical assistance including security testing. Furthermore, we provide the vendors with coaching and opportunities to engage with investors and other stakeholders.

The most transformative Proofs of Concept (POCs) developed through the [Innovation Leadership Programme](#) are selected for agile prototyping and deployment. The process involves global competitions to crowdsource ideas from technologists and identify the best implementation partners among both other financial authorities and vendors. The Lab's Launchpad largely builds on the experience of the RegTech for Regulators Accelerator (R<sup>2</sup>A), which successfully developed [groundbreaking applications](#) by introducing an agile mode of collaboration between financial authorities and vendors.

For more information, please visit <https://lab.ccaf.io/launchpad/>

## I. Project Description

To comply with its mandate to preserve financial stability and identify weaknesses and risks in the entities it supervises in a timely manner, the Agency has developed as one of its fundamental pillars, a demanding and permanent financial supervision function aligned with values such as prevention and prudence. Through supervision, the Agency verifies adequate compliance with the regulatory framework and evaluates the risk profile and management of the supervised companies. The results of supervision can lead to cross-cutting and sectoral actions, so the development and implementation of its regulatory and supervisory activities are comprehensive, i.e., considering its supervisory approach, objectives and mandates, as well as the external environment and the risks existing in the supervised systems.

In that scenario, it is important to improve market conduct monitoring tools and process, with the use of a social network monitoring tool that uses artificial intelligence ("AI") and machine learning ("ML") to process data. This would automate the process of data collection, storage and processing of information, allowing for simplified reports through the use of language processing models, as well as real-time alerts regarding conduct contrary to the Agency's regulations and financial legal framework.

Within this context, the Agency is seeking advanced market monitoring tools that extend the existing capability of the authority, going beyond basic scraping and sentiment analysis to surface trends, anomalies, and other patterns in the data that is relevant to monitoring of entities and activities flagged through other processes. This web-based market monitoring application introduces advanced monitoring tools for analysing data gathered via social media and other channels, building on initial efforts already in place at the agency. When an issue relating to a supervised entity or activity is raised and resolved via a consumer complaint, inspection, or otherwise, that is not always the end of the story. Further monitoring ensures the issue remains resolved.

## II. Description of Required Solution

### Basic Requirements

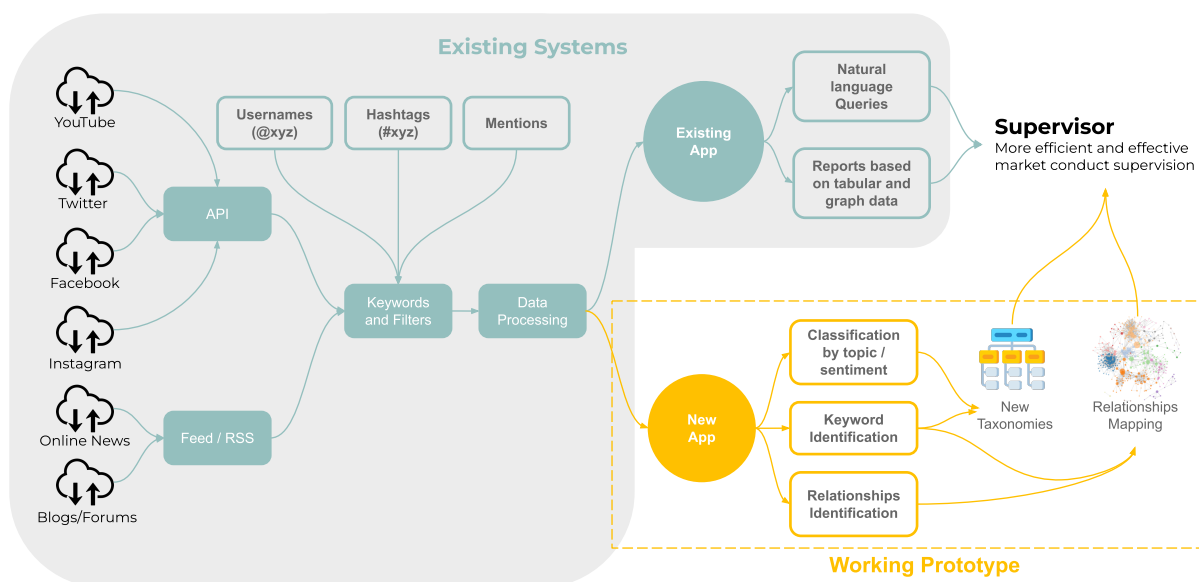
The Web-Based Market Monitoring Project will support the Agency in different aspects of their functions, enabling them to:

- Improve early warning systems, as social media monitoring provides early warning signals for potential financial misbehavior, allowing the Agency to act quickly.
- Improve the risk management as social media monitoring provides real-time information about market conditions and emerging risks, allowing the Agency to make informed decisions and manage risk effectively.
- Accurately identify and track consumer (negative) sentiment.
- Capture and analyzing vast amounts of data in real-time.

- Establish effective keyword and hashtag searches by AI.
- Automate data collection and reporting.
- Integrate analytics into the tool to provide insights and recommendations.
- Develop custom dashboards and reports to be used with the external provider's platform.
- Create automated alerts to notify supervisors of potential issues or opportunities happening in real time (potential misconduct).
- Identify and monitor emerging trends.
- Integrate data from multiple sources (specifically with traditional claims & complaints).

### High-level architecture

- The vendor gathers the data from social networks through their APIs.
- This information is filtered through previously defined keywords and other filters such as username and hashtags.
- This filtered information is processed using dictionary words and rules in order to get a sentiment score which can be positive negative or neutral.
- This sentiment score is assigned using an algorithm based on the previous rules however this sentiment analysis can be prone to errors and has limitations (it does not accurately identify the post's sentiment).
- Information would be processed using new technologies such as text analytics, graph databases and AI/ML, which would enable better keyword identification, which in turn would enable the creation of new taxonomies and word relationships, producing more relevant and accurate results, reports and alerts.



The existing app identifies information on mentions registered in social networks and other sources such as newspapers and blogs related with the main supervised entities of the financial system. The service provided allows capturing through algorithms the publications linked to market conduct issues, assigning them a sentiment (neutral, positive and negative).

Notably, however, the existing tool used is not ad hoc for supervisory work, the reports issued with the use of the information collected are static and limited, requiring manual validation not only by the provider, but also by the supervisory team to identify a possible misconduct by the supervised entities.

Likewise, the results obtained from the monitoring information cannot be automatically integrated with other sources of direct information, requiring more time and effort from supervisors to obtain integrated data.

### Key Technical Requirements

The solution must provide for the following requirements:

- Allow establishing parameters and identifying information in an easier and more dynamic way through natural language queries
- Provide for the application of advanced analytics techniques during initial processing of social media posts
- Automate and enrich the identification of keywords, classification of posts by specific subject or topics and discover new relationships between keywords
- Provide more precise information and improve the market conduct supervision model, especially the direct users monitoring mechanisms and dashboard
- Allow market conduct supervisors to:
  - Easily interact with and provide information with “natural language queries” for real-time monitoring and reports creation (i.e. a faster and more efficient tool to process data using AI Automatic reports)
  - Extract the real subtext of the data, filtering out noise, and providing alerts to identify possible misconduct.

### Scope limitations

In Scope:

- A market-level suptech application with the purpose of improving the Agency's market conduct supervision process through the monitoring of social networks and other indirect sources, in order to obtain and process financial consumer data

- Information to assist in decision making related to market conduct supervision actions, through agile, detailed and accurate reports on indications of misconduct identified by users' questioning on social networks, Internet and indirect sources such as complaints and denunciations
- Documentation including key lessons learned with respect to engagement with the Agency and Lab's staff, innovators, and other important stakeholders.

Out of Scope: No limitations explicitly identified.

### III. Key vendor requirements

The Project requires a vendor with the capacity, relevant experience, and resources to design, develop, test, and deploy a prototype with the purpose of web scraping of open sources, DSM, analysis, and integration with current chatbot solution.

#### General Launchpad qualifications

- Specificity: the competition is result oriented and the proposed solution needs to have a high level of detail and granularity with respect to the expected output.
- Precedent: the applicant needs to work on a novel solution.
- Geography: the vendor can be based in any jurisdiction. Data needs to be stored in an infrastructure compliant with the needs of the financial authority.
- Collaboration: the development of the solution should be conducted in collaboration with the team designated by the financial authority in each distinct phase.
- Sufficient experience to build an application that can serve the data needs of the financial authority, e.g.:
  - Integrate with existing application within a governmental entity
  - Allow the migration of existing data
  - Provide real time, on-demand support and the ability to generate reports or summary
  - Provide a high standard of application security.
- Demonstrated ability to:
  - Manage product life cycle
  - Develop, complete, implement, maintain, and deliver the appropriate technologies
  - Properly write documentation

- Maintain an enterprise ecosystem.
- Experience with:
  - Suptech solutions
  - Working with regulators and financial authorities
  - Frontend Development and UI/UX Design (as needed)
  - Data products and practical applications of analytics and data science (e.g., AI/ML)
  - Software engineering
  - Application architecture and devops
  - Program/project management and business analysis
  - Agile methodologies for application development
  - Application integration and performance tuning/optimization.
- Knowledge including:
  - Cybersecurity and secure application development coding standards
  - Best practices in relevant fields to the solution at hand.
- Resources:
  - Technical expertise on related knowledge and experience
  - Sufficient staffing and computing resources required by the identified feature and time requirements and constraints
  - Sufficient specification of online, on-premises, and/or hybrid computing resources
  - Adequate project management staffing based on requirements
  - Software, hardware, network, and cloud computing licenses and subscriptions to cover development, implementation, and warranty period.
- Project management: methods that leverage agile delivery methodologies for project planning, design, building and testing, stakeholder engagement, and effective risk management to ensure on-time completion of the project without budget overrun.

#### Engagement-specific requirements

- Strong Technical Skills: advanced technical skills in artificial intelligence, machine learning, and data science, as well as experience developing and deploying AI/ML applications.



- Understanding of Social Network Data: a deep understanding of social network data, including how it is generated, collected, and stored, and the types of insights that can be gleaned from it.
- Experience with Big Data Technologies for public data: proficiency in processing and analyzing large volumes of social network data, and have experience working with big data technologies.
- Knowledge of Privacy and Security Concerns: a strong understanding of privacy and security concerns and be able to develop AI/ML applications that respect users' rights and protect their data, given the sensitive nature of social network data.
- Creative Problem-Solving Skills: ability to approach social network supervision challenges with a creative and flexible mindset, and be able to develop innovative solutions that address the unique needs of different stakeholders.
- Strong Communication and Collaboration Skills: excellent communication and collaboration skills, and ability to work effectively with interdisciplinary teams to develop and implement AI/ML applications for social network supervision.
- Business Acumen: a good understanding of the business needs and goals of social network supervision, and be able to develop AI/ML applications that align with these goals and drive value for the organization.

#### Experience:

- Relevant Work Experience: several years of professional experience developing AI/ML applications, with a strong portfolio of successful projects.
- Strong Technical Skills: a deep understanding of AI and machine learning algorithms, as well as hands-on experience with programming languages such as Python, R, or Java.
- Data Science Experience: strong data science skills, including experience working with large data sets, knowledge of data visualization techniques, and the ability to design and implement complex data models.

#### Knowledge:

- AI and ML knowledge: a deep understanding of the fundamental concepts and algorithms of AI and ML, including supervised and unsupervised learning, deep learning, and reinforcement learning.
- Big Data Analytics Knowledge: experience working with large data sets, including knowledge of big data technologies such as Hadoop and Spark, and the ability to process and analyze big data to uncover insights and patterns.
- Data Science Techniques: experience applying data science techniques, such as data visualization, data cleaning, and feature engineering, to help develop and improve AI/ML models. Also, a good understanding of data privacy and security best practices.



#### IV. Project Award

The successful applicant will:

- Be awarded **US\$100,000** to develop and test the required solution. This is a fixed-sum contract, which is to cover all the applicant's expenses related to the development and testing work, including staff time, hardware, software, travel, and all other project-related expenses.
- Receive tailored coaching
- Be invited to the Lab's pitch day to connect with funders and to a demo day to present their products to potential clients
- Be listed in the Lab's online [Vendor Database](#)
- Be mentioned in a case study published by the University of Cambridge to share lessons from the project
- Engage with the supotech community through the Lab's hackathons and techsprints
- Be introduced to the global community of regulators and supervisors, investors, academics and development partners that are collaborating with the Cambridge SupTech Lab during other events hosted by the Cambridge Centre for Alternative Finance (CCAF).

#### V. Vendor Selection and Project Implementation

##### Timeline

Following the receipt of Expression of Interest submissions from qualified vendors, a Request for Proposals (RFP) will be formally issued to three shortlisted vendors on March 19<sup>th</sup>, 2023. The invited vendors will have ten days to submit their proposal. The winner of the RFP will be announced on May 2<sup>nd</sup>, 2023, with work commencing within three weeks. This Project will ultimately deliver a prototype that will be tested by the Agency no later than November 2023.

##### Key features of this initiative

- Blind review process: A panel of expert reviewers will score anonymised proposals without knowing the name of the vendor submitting them.
- Competitor scorecard: Applications will be assessed by the panel using a set of scoring metrics and weighing the relative importance of each attribute.
- Rapid turnaround time: We will select the winning vendor and award 50% of the **US\$100,000** within 48 days from submission of the final proposal. The last 50% of the award

will be granted in one installment upon completion of the deliverables according to projected timelines.

### Project Structure

The Project has four phases, elaborated below:

1. Kickoff and interface design, including technical integration specifications
2. Development of a working prototype
3. Integration and development of additional data analytics and/or visualization
4. Testing and signoff of the working prototype

Throughout all project stages, vendors are expected to meet weekly with key stakeholders of the Agency as well as the Lab's Launchpad team, to ensure close coordination and agility.

#### 1. KICKOFF & INTERFACE DESIGN, INCLUDING TECHNICAL INTEGRATION SPECIFICATIONS)

During the first phase, the selected vendor (in coordination with the Lab's Launchpad team) will gather requirements from the Agency and produce an initial Design Document that includes integration and user interface specifications. This living document should include specifications for the client-facing portion of the prototype, communicated in a manner such that clients of the prototype can understand how to integrate with systems and processes, submit data, and use the system without necessarily understanding the entire architecture behind the software. This includes specifications for the data integration (analytics and visualization) phase of the project as well. This Document is to be shared as needed with any other key stakeholders (e.g., vendors of relevant software used by the Agency, any financial institutions needing to integrate) to allow them to develop integrations and/or adapt existing software during the development phase. The design document should also include criteria for user acceptance testing for use during the testing phase.

#### 2. DEVELOPMENT OF A WORKING PROTOTYPE

The selected vendor will build a working prototype that delivers on the requirements laid out in the REOI, RFP, Project Agreement, and any agreed-upon modifications to scope agreed during the previous Phase 1.

The Web-Based Market Monitoring system will first receive a small subset (a representative sample) of all data required in a controlled environment. Initial data being submitted via the system will be sample data to start, with real data only being introduced to the system once proper security protocols and data sharing agreements are in place. Starting with a prototype and a small data set will allow the vendor to quickly identify and address any unforeseen issues early in the Project development cycle.

Once agreements are in place, the model can be trained, tested, and validated on real data extracted from social media and other channels.

The working prototype will also facilitate candid discussions among Project stakeholders regarding issues such as model interpretability, potential externalities, and the like.

### 3. INTEGRATION AND DEVELOPMENT OF ADDITIONAL DATA ANALYTICS AND/OR VISUALIZATION

Once the working prototype has been developed, tested, and accepted, the vendor will provide any analytics and visualization tools defined during the design stage.

Additionally, the selected vendor should (i) assess the needs of the Agency to understand which dashboards, reports, and statistics are most useful and/or difficult to produce under the current web-based market monitoring system; and then (ii) propose and develop a prototype mechanism for extracting and visualizing this information from data consumed, processed, and produced by the working prototype. This could involve creating custom queries, scheduling the generation of reports, and outputting in various formats.

The final UI of the prototype must be done before the final tests of Phase 4.

### 4. TESTING & SIGNOFF OF THE WORKING PROTOTYPE

Once the proof of concept has been completed to the satisfaction of the involved parties, integration and testing with real institutional data can begin. The working prototype will be tested with the Agency, based on any user acceptance criteria defined during the design stage, to allow the vendor to ensure that the prototype Detector can handle the acceptance, reporting, and detection of fraud within real data before it is scaled into full production. This approach also minimizes the risk of interruption due to unforeseen technology failure and serves to inform estimates of the cost to scale the prototype to a production-grade service.

A cyclical final test of the prototype and improvements by the developers must be done until the product is adherent to the functional specification document.

## VI. Rules and guidelines

### Submission Requirements

Interested vendors must demonstrate that they are qualified to perform the services required for the Project. In particular, interested vendors are asked to address the following requirements in their submission in a format capable of being read by Microsoft Word, which should together be **no more than 8 pages in length** (minimum 11-point font):

1. Company background (including technical and managerial capabilities of the executives).

2. A list of past projects representative of the experience of the firm and the executives.
3. Information on the qualifications of key staff to be involved in the Project, including whether they have experience working on suptech projects.
4. Summary of your working experience in the following geographies: (1) Brazil; (2) Philippines; (3) Indonesia; (4) Ghana; (5) Peru; (6) India.
5. Examples of prior experience related to the development of this solution prototype or similar technological solutions.
6. Indicative development/implementation schedule based on the timeline set out above.
7. Detailed description and examples of technical prowess that address both components of the key vendor requirements in section III.

All materials should highlight the relevance to the required solution described in this document.

Supporting documents may be submitted (e.g., company brochures, case studies, CVs, etc.) as attachments to the submission email. No page limits apply to these attachments, but evaluation will be based primarily on information included in the main body of the EOI.

Any questions prior to the submission deadline should be sent via email to Cambridge SupTech Lab's Launchpad at [suptech-launchpad@jbs.cam.ac.uk](mailto:suptech-launchpad@jbs.cam.ac.uk) with subject line "REOI: Cambridge SupTech Lab – Web-Based Market Monitoring Project".

### Disclaimer

This document is not a request for proposals. The Lab reserves the right to edit, invalidate, terminate, and/or reissue this REOI at any time and for any reason. The Lab also reserves the right to select a vendor through an alternate method and/or adopt an alternate timeline for vendor selection that differs from the method and/or timeline described in this document, the websites of the Lab and Launchpad, and any other communications related with the process. Furthermore, the Lab expressly disclaims responsibility for any costs incurred by any vendor in responding to this REOI, regardless of whether the REOI is edited, invalidated, terminated, and/or reissued at any time and for any reason.