MEMORANDUM OF UNDERSTANDING (MoU)

FOR
DATA ANALYTICS – PROTOTYPE ABOUT
Internal virtual assistant for social media communicators
and mythbusters – Phase II

between
THE DIRECTORATE-GENERAL FOR COMMUNICATION (DG COMM)
of the European Commission

and
THE DIRECTORATE-GENERAL FOR INFORMATICS (DG DIGIT)
of the European Commission

Reference: DIGIT-00928-00

The Directorate-General for Communication, hereinafter “DG COMM”, represented by [Name], Director-General DG COMM

and

the Directorate-General for Informatics, hereinafter “DG DIGIT”, represented by Mr [Name], Acting Director-General DG DIGIT

HAVE AGREED as follows:

1. PREAMBLE

As stated by Commission President Ursula von der Leyen in her Political Guidelines, “Digital technologies, especially Artificial Intelligence (AI), are transforming the world at an unprecedented speed. They have changed how we communicate, live and work. They have changed our societies and our economies”. One of the wider applications of AI in communication is in the form of virtual assistants, or chatbots. Gartner predicts that by 2022, 70% of white-collar workers will interact with conversational platforms on a daily basis. This expected growth is on par with the increase of millennials in the workplace. Business Insider experts predict that by 2020, 80% of enterprises will use virtual assistants.

DG COMM has been working towards incorporating artificial intelligence and machine learning into the Commission’s social media communication and mythbusting work in the form of an “internal virtual assistant” in support of the work of social media communicators and mythbusters. The aim is not to replace human capacity but to
develop AI solutions that support frontline staff by providing an extensive pool of knowledge on which they can draw instantly, thus improving the quality of their interaction with the public.

DG COMM and DG DIGIT have been working together to develop the prototype of an in-house virtual assistant that could be further deployed across the Commission. The work on the prototype so far (Phase I) has demonstrated that it is possible to create an internal virtual assistant that is highly useful for Social Media communicators and mythbusters in their daily work. Working closely with DG COMM, DG DIGIT set up the necessary infrastructure, developed a virtual assistant prototype based on the question-answer format in English and implemented the e-Translation tool from DG DGT, so that the custom chat interface for experimentation and demonstration can today operate in all 24 EU official languages.

DG COMM now envisages a second phase of the prototype in which users from the Social Media Network would be brought in – in order to test and feed the knowledge base of the virtual assistant. The underlying knowledge base of the tool would be strengthened in its machine-learning components, as well as further expanded and tailored to the needs of the Social Media community. The second phase would include the following:

- Internal testing of the virtual assistant and feeding it with more questions and answers (DG COMM);
- Collecting feedback from users (DG COMM) and incorporating it in an improved version of the virtual assistant (DG DIGIT);
- Developing the algorithms, active-learning and text generative capabilities of the virtual assistant (DG DIGIT);
- Exploiting machine learning on the knowledge graph to experiment with content validation, fact-checking, and indicating of trending/hot topics (DG DIGIT).

Looking much further into the future, this internal virtual assistant could be the launchpad for an external, EU-wide version capable of interacting directly with citizens on all our platforms, in all 24 languages. An EU-wide version has the advantage of ensuring interoperability, being responsive to specific needs, and represents a common approach which will mitigate risks inherent to AI solutions within the Commission. Deploying any such external citizen-facing function would be subject to political approval, inter-institutional negotiation and extensive trialling.

2. SUBJECT

The subject matter of this MoU is to define the framework of collaboration between DG COMM, as business owner, and DG DIGIT, as project manager and system supplier, on the development of a prototype for an ‘Internal virtual assistant for social media communicators and mythbusters – Phase II’, as described below.

3. SERVICE

3.1. Solution definition

3.1.1. Improvements

3.1.1.1. Performance: Make the needed architectural and infrastructural changes to support up to 200 users. Improve on accuracy and response time.
- Identify Bottlenecks
- Improve Neo4j graph algorithms
- Improve Deep Learning NLP algorithms
- Evaluate change of Graph DB
- Evaluate work impact on accuracy only with keyword search
- Review architecture to assure scalability

3.1.1.2. Collect Feedback: Integrate a tool to collect and manage feedback from multiple users.
- Evaluate and implement way to collect feedback
- Integrate JIRA API
- Form in application
- Store feedback reply matching accuracy indicated by user to be used by active learning.

3.1.1.3. Additional data sources: Provide the ability to inject data into the knowledge graph from internal EC data sources.
- Introduction of new sources with minimal impact on performance
- Adjust architecture to new data structures
- Investigate overlap with other projects (document ingestion, text extraction) and reuse components
- Implement spellchecking and data cleaning on random quality data (tweets, social media posts, ...)

3.1.2. Active Learning: Provide the ability to continuously evolve the bot by considering the interaction with the users.
- Graph Neural Networks (predict/classify nodes instead of search) [https://docs.dgl.ai/index.html](https://docs.dgl.ai/index.html)
- Generate parts of query using ML
- Integrate user satisfaction feedback

3.1.3. Generative Chatbot: Provide the ability to generate responses and summaries based on information in the knowledge graph.
- Generate Reply function
- Generate Summary function

3.1.4. Additional knowledge graph learning (time allowing): Experimentation of the ability to validate content against the knowledge graph and to provide a determination of trending and “hot” topics.
- To investigate the inclusion of multimodal learning from media
- Targeted content validation based on entities (figures, dates, names and roles, ...), using entity extraction and query generation
- Experiments on the determination of trending topics and the lifespan of “hot” topics (discovering expired topics)
- Consideration of requisite architecture for the additional features

3.2. Work packages

WP1 - Improvements:
- Integrate user satisfaction feedback on question reply:
  - change schema DB and injection scripts;
  - adapt API;
  - adapt Assistant;
  - adapt Chatbot;
• Build metric for measurement of accuracy and performance and register base line:
  • scripting for extraction from DB;
  • scripting for measurement and statistics;
• Identify bottlenecks:
  • use metric for performance to identify bottlenecks;
• Assure Scalability:
  • address identified bottlenecks;
  • redesign Architecture to assure scalability;
• Evaluate and Integrate feedback form:
  • define approach;
  • integrate back office front and back end if needed;
• Chatbot Performance Scaling.

WP2 - Sources extension:
• Injection of data:
  • Adding new data sources / formats:
    • EC internal:
      o evaluate new sources, especially DG COMM;
      o define and implement approach for data extraction;
      o edit graph schema if needed;
      o adapt scripts for injection;
    • Tweets:
      o Adapt scripts to inject tweets;
      o Adapt Schema Knowledge graph;
• Extending UI for Injection and Graph Management

WP3 – Text-generative bot:
• Evaluate and implement generative approaches:
  • define approach;
  • define metric for validation;
  • change back and front end;
• Evaluate and implement Graph DL for active learning.

WP4 – Additional knowledge graph learning (time allowing):
• Facts knowledge graph (semantic db):
  • define and implement schema;
• Query layer for content validation:
  • experimentatal approach for fact checking and content validation;
  • expose as API to other layers;
• Experiments on trending and “hot” topics:
  • machine learning model for the determination of trending topics and lifespan of “hot” topics.

3.3. Deliverables

D1: Implement performance and accuracy improvements of front- and back-end
D2: Implement feedback monitoring tool
D3: Inclusion of additional data sources
D4: Text generation bot component
D5: Active Learning component
D6: Additional knowledge graph algorithms
4. Management

4.1. Duration

This MoU enters into force on the date of the last signature and remains valid for one year. Its duration may be extended by mutual agreement via amendment.

Any amendment to the present MoU shall be made in writing, and duly signed by both parties.

4.2. Addenda, evaluations and termination

Where a party to this MoU wishes to end its co-operation, it shall provide the other party a written notice at least one month before the intended date of cancellation.

In the event of termination, both parties shall take all appropriate measures to prevent business discontinuity, minimise costs and cancel or reduce its commitments.

4.3. Governance, roles and responsibilities, monitoring and reporting

Meetings at regular intervals will be held to take the necessary decisions on business, IT and budget matters at two levels:

- At the start of the project, the Heads of Unit of DG COMM.A1 and DIGIT.D1 will meet to agree on a detailed work plan and final budgetary sheet drawn up by DIGIT. At least once every three months, they will meet to take stock of implementation and budget consumption. They will approve the quarterly progress and budget report.

- At least once a month the expert teams in DG COMM.A1 and DIGIT.D1 will meet under the chair of Head of Unit of DG COMM.A1 to take stock of progress made towards monthly milestones, and to review budget allocation and consumption. They will report to the respective Heads of Unit in DG COMM.A1 and DIGIT.D1, who will have to agree on any substantial changes to the initial work plan and budget allocation.

DG DIGIT will be responsible for project management and implementation, including the necessary procurement procedures. DG DIGIT shall report, at regular intervals, on the use of the co-delegated appropriations, in particular:

- Once a month, DG DIGIT will report on budget consumption against the final budgetary sheet agreed at the start of the project, detailing time & means, quoted time & means, license and hardware purchase and hosting costs.

- Every three months, DG DIGIT will draw up a progress report accompanied with a quarterly budget report.

Both parties commit to implement this MoU in a spirit of cooperation, transparency and efficiency.
5. **Budgetary resources**

DG COMM will provide the total budgetary resources required for the development of the prototype as described under section 3.

The total amount is €157,384,90. This amount will cover all direct and indirect costs incurred in the development of the prototype. DG COMM will provide new budgetary resources of €131,384,90, and unspent appropriations of €26,000 were provided through MoU DIGIT-00894-00.

Any unspent appropriations at the end of the project will be used by DG DIGIT in the next phase of implementation, subject to the signature of a new Memorandum of Understanding.

6. **Funding modalities**

DG COMM will put the new funds referred to under Article 5 at the disposal of DG DIGIT via co-delegation on the appropriate budget line(s).

DG DIGIT will authorise the commitments and payments as well as the other underlying transactions incurred to deliver the services under this MoU.

These transactions will be managed under the internal control framework of DG DIGIT, which complies with the Commission's rules and guidelines.

7. **Data protection**

Personal data included in or relating to the present MoU and its implementation, including data related to the provision of the services, shall be processed in conformity with Regulation (EU) 2018/1725 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data.

Such data shall be processed solely for the purpose of the implementation, management and monitoring of the present MoU by authorised personnel, without prejudice to any possible transmission to the EU bodies charged with a control or inspection mission in applying Union law.

8. **Contact persons**

For DG COMM:
- Technical part: COMM.A1
- Administrative Contact: COMM A1
- Financial Contact: COMM A1

For DG DIGIT:
- Technical part: DG DIGIT D.1)
- Administrative part: (DG DIGIT D.1)
- Financial part: DIGIT-MOU@ec.europa.eu (DIGIT A.2)

Any changes in points of contact do not require an amendment to this MoU. They will be notified by e-mail to the other signing party.
9. **ANNEXES**

N/A

Signed electronically in Ares

**For DG COMM**

[e-signed]

Director-General DG COMM

Done in Brussels on (see Ares Visa)

**For DG DIGIT**

Digitally signed by:

(EUROPEAN COMMISSION)
Date: 2020-11-24 13:38:01 +01:00

Acting Director-General DG DIGIT

Done in Brussels on (see Ares Visa)