INSTITUTE OF PUBLIC ADMINISTRATION

DEMYSTIFYING THE PROCESS OF

CREATING DIGITAL ASSISTANTS:

the experience of the
first state institution
with AI in Bulgaria

FEBRUARY 2025

Digital assistants are not something mystical or complicated to understand. They are tools with very specific capabilities and limitations. For their successful implementation, it is important to understand that they are assistants designed to solve specific tasks, not one-size-fits-all solutions. They require careful planning and ongoing maintenance, and their success depends on the quality of the information they have.

How does the digital assistant work?

Imagine an employee in an information centre - let it be Maria, who knows perfectly all the documentation of the institution, knows by heart all the most frequently asked questions and answers and when a citizen comes to her with a question, she can very quickly assist without time. The digital assistant works in a similar way: he has a "knowledge base" (this formal and informal knowledge of Maria) that is created by the experts in the institution. When someone asks a question, the assistant:

➢ First search this database for the relevant information.
➢ If sure of the answer, provide it immediately.
➢ If not completely sure, it directs to a person on the team.
➢ Can respond to multiple people at once, anytime - 24/7.

Unlike Maria, however, the digital assistant cannot make independent decisions or exceptions to the rules. Therefore, in more complex cases she always refers to an expert. Like Maria, the digital assistant learns from experience through the experts at the institution who help by monitoring his responses and constantly improving his knowledge base to become more and more useful. Just as professional athletes train almost every day, so do the experts train the assistant, it takes will and determination.

Experience with chIPA and technical aspects

The Institute of Public Administration's digital assistant, the **chIPA** chat-bot, in this version, has been developed with two main objectives:

1) Fast delivery of relevant information.
2) Assistance with questions about IPA training.

Our digital assistant was developed using GPT-4o-mini, and integrated into a system that combines a vector database and algorithms to extract relevant information. The IT experts we worked with, volunteered their time, created a vector database in which documents we had previously processed were stored.

An important point was the selection of information - we decided that the assistant should work with pre-selected and verified data instead of relying on information from the Internet and the Institute's website. We gathered relevant public information about the IPA, as well as frequently asked questions by phone and email, information about various initiatives and professional networks supported by the IPA.

In order to make the documentation comprehensible to the language model, it was transformed through vectorization, a process in which text is represented as mathematical vectors in a multidimensional space. This allows information retrieval not through exact word matches, but based on meaning proximity.

To steer the basic model in the direction we wanted, we used a specific formulation of the initial prompt/query that sets the entire tone of the assistant's communication. There we explicitly set the constraints that guide the correspondence and the information shared.

The digital assistant uses MMR (Maximal Marginal Relevance) algorithm to optimize information retrieval. The MMR algorithm helps to balance between:

- Relevance - how relevant a document is to the user question.
- Diversity - avoiding results that are too similar in order to present a broader context.

This balance is governed by a parameter that determines the relative weight of the two features. In our case, the focus is on relevance to ensure that the answers are as accurate and close to the user query as possible.

To ensure the safety and correctness of interactions, we have implemented the following measures:

- Protection against malicious attacks - filtering out queries that can manipulate the model or take it out of context.
- Tracking through logs - analyze user interactions to improve accuracy and detect potential problems.
- Session limit - each user session is limited to 50 requests, after which it restarts automatically. This prevents excessive context accumulation and potential exploitation attempts.

An important detail is that the IPA database and model arranged in this way can be "hooked up" to alternative Open AI AI systems such as BGgpt or others.

Challenges and solutions

When creating digital assistants, you will encounter various challenges - from technical issues to instances of assistant "hallucinations". This happened with chIPA as well. Here, it was particularly important that the implementation team followed a few important rules:

- Constant monitoring and adjustments, mainly outside standard working hours for administration;

- Initial involvement in the test of 5 people and the Executive Director personally, phased involvement of other IPA experts, recruitment of volunteers for testing from different institutions and communities in the public administration system;

- Extremely strict control over correct answers;

- Regular enrichment of the knowledge base;

- Adherence to the principles of ethics and integrity and to good cyber security practices,

as a result of which we have also drafted the terms of use for chIPA to reflect these principles https://www.ipa.government.bg/bg/usloviya-za-polzvane-na-chipa-chatbot-na-ipa.

The involvement of external experts and colleagues in the testing process proved to be extremely valuable not only for the validation of the model but also as a catalyst for its improvement. Their 'outside' view assisted in enriching the contextual knowledge base through different professional perspectives and identifying additional application scenarios.

Development

The model is designed as a "living" organism - it continues to learn and improve based on interactions with IPA users. A controlled approach to training is used to ensure the quality of responses, with any new information going through a validation process before being integrated into the model's knowledge.

The creation of a digital assistant is also an opportunity to develop the competences of the teams, because everyone involved in the process is learning how to overcome different challenges - lifelong learning in action☺ . Teamwork.

The IPA has bigger ambitions and will test the limits of AI as the main goal will be, along with the question answering activity, to personalize content and training, to pilot a digital assistant in the CAF implementation process, and all possible activities to improve the learning experience of civil servants.

And why should we, you may ask?

A digital assistant is one of your colleagues who:

• Responds on the spur of the moment to user questions.

• Answers repeated questions patiently.

• It helps many people at the same time.

• Provides equal access to information, all day and night.

• It helps you to work better, to concentrate fully on your actual work when you have deadlines and urgent tasks that require your full attention.

For the administration this means more efficient work and for citizens - more accessible services at all times.

The digital assistant does not replace employees, but helps them by taking over the bureaucratic work, the work on repetitive issues. This leaves us more time for the more important tasks that require human judgement and attention. Ultimately, the point is to make the administration's work more efficient through this new, digital assistant on the team, leaving time for the actual analytical work.

Authors. Sofia, February 2025

Supported by the European Digital Innovation Hub "Trakia".