# **Proposal ARTH**

### **Automated Rail Traveller inquiries Handling**

April 4th 2023



▶RS:/01 ▶RS:/01

▶R5:/0211TR /ON ▶R5:/0211TR /ON



### **Project challenge**



ARTH outlines an automatic end-to-end product to automate and scale the processing of train passenger requests received by the ÖBB. The proposal is composed of successive **modules** that implement different processing stages. An implementation of each module is sketched in this proposal presenting a minimum viable prototype of the whole process. Once validated, each of them can be improved to refine its functionality or to expand its capabilities.

ARTH comprises the following modules:

- **1. Data gathering:** the currently existing form will be replaced by an online form that is filled via the web to digitalize the process and enforce data validation from the very beginning. Submissions in paper will be also accepted to avoid any kind of discrimination.
- 2. Data extraction and early request filtering: data gathered by the form is inserted into a database, structured using natural language processing (NLP) technologies and fed into a classification algorithm that filters the enquiries that do not correspond to passenger rights' requests.
- 3. Request validation: the details of the request are validated by comparing the provided data with the data existing in the ÖBB database.
- 4. Request categorization: based on the validated data of the enquiry, the system assigns a category to the request.
- 5. **Compensation proposal**: a different compensation procedure is applied depending on the category assigned to the enquiry. The degree of autonomy of each procedure depends on its complexity and also on the risk of the involved compensation. In all cases, the system will be overseen by human analysts that will be able to iterate with the proposed solution with a suitable tailored front-end.

It is important to remark that the handling and storing of personal data will be done in **strict accordance with the European directive 95/46/EC** which sets the regulatory framework for the processing and storage of personal data.

The training of the described algorithms needs a relevant and sizeable dataset of historical passenger requests together with their categorization, compensation, mitigations and other available data on their currently manual processing.



### **Backend (rest-API)**



- Modern and reliable rest-API backend fully **compatible with existing systems** (e.g. MS Dynamics) ⇒ easy to use for existing ÖBB employees.
  - Scalable solution with dynamic resources allocation: up-scaling during peak times and down-scaling during idle periods.

>

ARTH

### Frontend (UX)

Front-end definition and implementation will be defined and implemented following an agile methodology, and thus final functionalities will depend on invested effort and requirements.

ΰBB 144 - Commission 1000 040 ..... ŰВ Herosto VO It is important to create a 0.000 user-fiendly frontend tailored to the needs of the ÖBB personnel to provide a simple and intuitive user experience to the operators. ARTH

DRIJIA

### **ARTH Vision**

### Phase 1



### ARTH v0.1 Prototype

- Pipeline prototype
- Development of a first version of the different modules.
- Execution of the pipeline with previous inquiries already processed by humans and comparison of the output.



### ARTH v1.0 MVP

- Integrated solution
- Complete pipeline deployed in production.
- Intensive human supervision through the designed frontend.
- Human validation required at relevant steps (e.g. aproving a compensation).

### **Future phases**



#### **ARTH v2.0 Solution**

- Complete solution
- Semi-automatic operation.
- Human feedback gathered for future improvement of the solution.
- Human validation only required at critical steps.

### **Development methodology**

 $\bigcirc$ 

Agile methodology: development short streaks of 10 working days (sprint) with two meetings

+ Flexibility and alignment of expectations.

- Hinimum risk: the Client can stop the project at any time with 2 sprints in advance.
- (!) Risk of **losing overall focus**  $\rightarrow$  high-level control meetings every 4 sprints.





### Working team

**Dribia** deploys **4 Data Scientist (DS)** teams **devoting a total amount of 1.5 FTE.** 



#### Business supervision - 1 Manager DS - 0,1 FTE\*

Alignment with business and high-level control of compliance with objectives.



#### Project Leader - 1 Senior DS - 0.4 FTE

Plans and manages the project, designs the algorithm, monitors the deployment. Main person in charge and contact person on our pARTH.



#### Developers - 2 Data scientists 50% - 1 FTE

Analyze the data, implement the algorithm, and write the documentation. The team consists of 2 Data scientists at 50 %.

## DRIJIA

At Dribia we believe that creativity requires dialogue and perspective. That is why we hold weekly seminars, in which the project team discusses with the entire Dribia team, offering fresh ideas.

### From **ÖBB** we will need:



#### Product owner - client lead (0,1 FTE\*)

Validation of the results and prioritization of the sprint definition. Internal facilitator. Responsible for the project on the client side.



#### IT support (? FTE\*)

Supports both access to the necessary data, documentation of tables and supports the definition and deployment in the development phase

#### Key users (? FTE\*)

Provides feedback from the end user point of view.

### Planning and Deliverables (MVP)

