EXPECTATION

PERSONALIZED EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR DECENTRALIZED AGENTS WITH HETEROGENEOUS KNOWLEDGE

D4.1

Technical report detailing the developed user model and agent-based profiling [M20]

Version: 1.0
Last Update: 28/04/2023
Distribution Level: CO

The EXPECTATION Project Consortium groups Organizations involved:

Partner Name	Short	Country
University of Applied Sciences and Arts Western Switzerland	HES-SO	Switzerland
Özyeğin University	OZU	Turkie
University of Luxembourg	UNILU	Luxembourg
Luxembourg Institute of Technology	LIST	Luxembourg

Document Identity

Creation Date:	1 April 2023
Last Update:	1 May 2023

Revision History

Version	Edition	Author(s)	Date

Introduction

A Virtual Coach (VC) is an intelligent computational system designed to offer guidance and supervision to human users, akin to the support provided by a human health professional/coach, to help them achieve their goals. VCs are designed to leverage personalized coaching to the broad public and as a complement for human coaches and domain experts while offering high availability and granular tracking capability. VCs use human users' data to model their preferences and behaviors, generating plans and strategies to help users achieve their goals. In recent years, VCs have become particularly relevant in health-related applications, assisting the users to change their behaviors and adopt healthier habits (i.e., quit additions, adopt healthy nutrition habits, exercise promotion).

One of the fundamental components of the VCs system is the recommendation module that can generate personalized suggestions according to the users' characteristics, the surrounding context, the plan in execution, and the user's goal. The recommendation module employs recommender systems (RS) use casesate suggestions and make proposals based on available data. RS can be non-personalized or personalized according to whether the recommendations are the same for everyone or if they differ depending on the user's characteristics.

Personalization and customization are essential requirements in critical RS use-case like nutrition advice, where recommendations should consider the user's health condition, goals, medical restrictions, cultural factors, and allergies. Personalization allows RS to leverage relevant suggestions that fit better users' preferences and needs. In addition, personalized RS can evolve and adapt to the users' needs and create a long-term trusted relationship with the user through several interactions. Thus, personalization is necessary for building trusting relationships between VCs and users.

Personalized RS are mainly based on user profiling. The profiling process consists of processing and analysis of personal and historical data to discover users' characteristics and behaviors' patterns and thus predict their future behaviors and interests. Profiling is principally built on users' personal data, which requires special treatment to ensure their security and privacy, and whose analysis and treatment entail ethical, legal, and social implications that must be considered in the design and development of the system.

The EXPECTATION project aims to create Nutrition Virtual Coaches (NVCs) with the ability to offer tailored nutritional guidance, explanations, arguments, and persuasive strategies to encourage the adoption of healthier eating habits. The proposed NVC structure utilizes user profiling to understand and continuously enhance their recommendations, explanations, arguments, and plans. This profiling involves two parallel processes: reliance on user-specific data and leveraging behavioral insights from past user interactions.

The user profile summarizes all the knowledge the system has about the user and is structured by an OWL-based ontology database. In the context of an ontology expressed using the OWL structure, the "concepts" refer to the abstract representations of classes or categories of entities in a particular domain.

This document describes the user profile and the profiling processes alongside the ontology and data structures required to store, process, and analyze user data.

User model

The user model comprises the most relevant features that characterize the user and constitute the user profile. The user model integrates the primary user profile (i.e., name, surname, age range, living country, clinical gender) and the user's dimensions that contextualize the user in the nutrition domain (i.e., health condition, metabolism, food preferences, and nutrition goals). In the ontology presented in Figure 1, the user model is defined and related to the rest of the system's entities (i.e., food recommendations and explanations) through semantic connections.

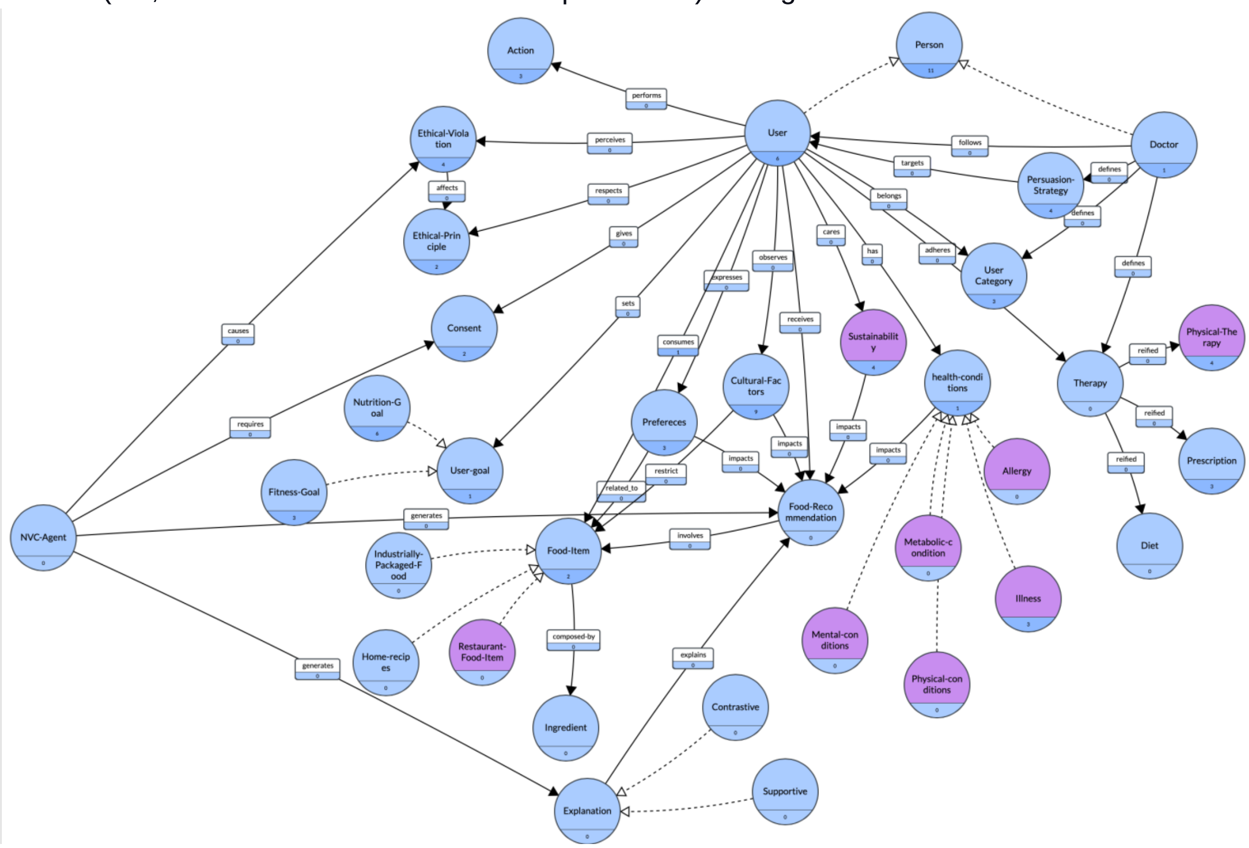

Figure 1: User entity representing user model in the system’s ontology.

Primary user profile

The primary user profile contains the user’s identification and punctual data, which includes the user’s identification application data (i.e., *email, username, user_id, password*), user's personal identity (i.e., user's *name*, *surname*), and punctual demographic data (i.e., *marital_status, age* *range, living_country, clinical_gender*). Figures such as *User* and *Doctor* are abstracted by the *Person* concept (having several traits in common within the ontological structure). Summarily, *Person* is a super-concept that could be any human actor of the system while a *Doctor* is a sub-concept of the *Person* that cares about healthcare aspects of the system while assigning user objective(s) to their given user. The *User* entity inherits several data attributes from the more general entity *Person*, used to characterize target users in the system. The person entity contains the following data attributes:

- *User_id:* Literal field that identifies the user according to the Pryv database (privacy-preserving and GDPR-compliant database [5].
- *Username*: Literal field that describes the user identity within the application.
- *Password:* Literal field that stores the password to allow access to the application and database account.
- *Email:* Literal field that stores the email address registered and employed to receive general application communication, recover passwords, and manage the user account.
- *Name:* Literal field that stores the user's first name, which identifies the user and can be used in the dialog by the coach to generate familiarity and personalized communication.
- *Surname:* Literal field that stores the user's surname, which can be used in the dialog by the coach to generate familiarity and personalized communication.
- *Clinical_gender:* Literal field that stores the user's gender, which is necessary to design the nutrition plan and characterize the user's metabolic response.
- *Current_location:* Literal field that describes the current user's location is required to provide relevant food suggestions according to the context and availability.
- *Age_range:* Literal field that stores the user's age range required to generate the nutrition plan and to calculate the metabolic response.
- *Living_country:* Literal field that stores the current user's country, necessary to provide culturally relevant food recommendations.
- *Country_of_origin:* Literal field that stores the user's origin country, and it is necessary to provide culturally relevant food recommendations and evaluate possible ethical or cultural violations (i.e., forbidden or disgusting food in certain cultures or countries).

Additionally, to the data attributes inherited from the *Person* entity, the *User* entity incorporate the following attributes:

- *Current_working_status:* Literal field that stores the current working status. It is necessary to calculate the daily calorie consumption and baseline physical activity level.
- *Marital_status:* Literal field that stores the marital status. It is necessary to calculate food portions in recipes.
- *Life_style:* Literal field status briefly describing lifestyle (i.e., exercise levels and frequency, physical activity, sleep patterns, etc.)
- *Weight:* Literal field to store the user's weight and its evolution. It is also necessary to calculate the Body Index Mass (BMI) and, based on it the user's goals, produce the nutrition plan.
- *Ethnicity:* Literal field to store ethical identity. It is necessary to identify the user's ethical and cultural context.
- *Height:* Literal field to store the user's height data. It is necessary to calculate the BMI principal indicator to determine the user's nutritional needs and current status (i.e., underweight, normal, overweight, obesity, extreme obesity).

User profile dimensions

The user model is completed with dimensions that characterize the user's current condition and state in the context of nutrition. Those dimensions are represented as entities in the ontology shown in Figure 1. The user's dimensions are described below:

Cultural Factors

The *Cultural Factors* entity describes the user's cultural context concerning food (i.e., strict dietary restrictions based on cultural or religious factors) and affects the food recommendation process. For instance, the nutrition virtual coach (NVC), also referred to as PERS (personalized explainable recommender system), needs to consider whether a person is a vegan prior to recommending a meaty meal. Cultural factors are composed of the following attributes:

- *Vegan_observant:* Literal field describes whether the user follows a vegan diet.
- *Vegetarian_observant:* Literal field describes whether the user follows a vegetarian diet.
- *Halal_observant:* Literal field describes whether the user follows a halal (Islamic law observant diet).
- *Koesher_observant:* Literal field describes whether the user follows a kosher (Jewish law observant diet).
- *Religion_observant:* Literal field describes whether the user follows a religious diet.
- *Drink_limitation:* Literal field describes whether the user has drink limitations (i.e., non-alcoholic, non-caffeine, non-lactose drinks).
- *Pescatarian_observant:* Literal field describes whether the user follows a fish-based diet.
- *Religion:* Literal field describes whether the user follows a particular religion and how this could affect their daily diet (i.e., fast, diet for religious days, among others).
- *Food_limitation:* Literal field describes whether the user has food limitations based on cultural or personal reasons.

Sustainability

The sustainability dimension describes the user's concerns about the environmental effects of the food recommendations. The sustainability entity comprises the following attributes:

- *Enviromental_harm:* Measure the environmental impact of the recommended food.
- *Eco_socre:* Score given to industrial packaged food based on production methods and package material.
- *Co2_food_print:* Measure the CO_2 amount generated to produce the food and the package.
- *Recyclable_packaging:* Literal that describes if the package is recyclable or not.

Ethical principles

Ethical principles characterize the user's ethical concerns and establish an ethical framework to evaluate the food recommendations, find ethical violations, and learn from the user's feedback to align recommendations with the user's values and ethical code.

- *Name:* String literal refers to an ethical principle followed by the user (i.e., respect, honesty, among others).
- *Description:* Literal describing the ethical principle and its effect on the dietary recommendations.

Actions

Those are the actions that users can execute on the system. Actions briefly describe the available interactions between users and the system. Actions are composed of the following attributes:

- *Action_type:* Actions that users can execute in the system (i.e., ask for food recommendations, ask for restaurant recommendations, track the consumed food, synchronize the Fitbit, provide feedback).
- *Location:* Geographical location where the action is performed to contextualize it.
- *Action_date:* The date and time when the action was carried out.

Preferences

This entity models the user's preferences related to meal time and is essential to synchronize recommendations with them. Preferences contain the following attributes:

- Breakfast_time: Preferred breakfast time employed to generate proactive breakfast recommendations.
- Lunch_time: Preferred lunchtime, employed to generate proactive lunch recommendations.
- Dinner_time: Preferred dinner time employed to generate proactive recommendations for dinner.

Additionally to the meal timing preferences of the user, the system also utilizes a preference modeling that utilizes Machine Learning models as explained in the Preference Modeling Section.

User goals

This entity describes the user's goals, which the NVC employs to produce plans and actions to achieve those goals. User goals are described by the goal category attribute and the *Nutrition_goal* entity (highlighted in bold) related to daily calorie ingestion and the *Fitness_goal* entity (highlighted in bold) related to physical activity and daily calorie consumption.

- *Nutrition_goal:* The nutritional goal is an entity that describes the targeted nutrients and calories ingested, and it is composed of the following attributes:
    - *Targeted_start_day:* Literal field to store the targeted day to start the suggested nutrition plan.
    - *Targeted_daily_nutrients:* Literal field to store the targeted nutrients per day according to the nutrition plan.

    - *Nutrition_goal_name:* Literal string field to store the nutrition goal name.
    - *Targeted_end_day:* Literal field to store the targeted end of the suggested nutrition plan.
    - *Targeted_daily_calories:* Literal field to store the targeted ingested calories per day.
- *Fitness_goal:* Entity that describes the targeted burnt calories through physical activity, and it is composed by the following attributes:
    - *Daily_walked_km:* Literal field to store the targeted daily km.
    - *Daily_burnt_calories:* Literal field to store targeted daily burnt calories.
    - *Daily_steps:* Literal field to store the targeted daily steps (i.e., 10000 steps per day to maintain a healthy habit).
- *Goal_category:* Literal describes the goal's type (i.e., lose weight, maintain the current weight, gain weight).

Health Conditions

Health conditions provide a holistic view of the user's physical and mental state. The *Health Conditions* entity is essential to plan treatments and give or avoid certain nutrition items. The *Health Conditions* entity is composed by the following entities and attributes (entities are remarked in bold).

- *Metabolic-condition:* Entity describing the user's metabolic conditions or disorders (i.e., metabolic syndrome, diabetes, hypoglycemia).
- *Mental-conditions:* Entity that describes user mental conditions or disorders that may affect diet and nutrition (i.e., Anorexia and bulimia).
- *Physical-conditions***:** Entity that describes the user's physical condition (i.e., active, sedentary, athletic, obese).
- *Illness:* Entity that stores the user's illness history that could affect the user's diet and nutritional needs (i.e., gastroenteritis). Illness entity comprises the following attributes:
    - Illness_name: Literal string field that describes the illness.
    - End_illness: Date field to store the end of the illness.
    - Start_illness: Date field to store the start of the illness.
- *Allergy:* Entity that stores the user's food allergies (i.e., allergy to peanuts).
- *BMI:* Literal float that stores the Body Mass Index (BMI), a single number that determines the user's weight condition (i.e., underweight, normal, overweight, obese, and extremely obese).

User Category

User category is defined by the doctor or domain expert (i.e., nutritionist), and it is employed to define the appropriate treatment and filter food recommendations according to the user's prototype.

- User_category_name: Literal string field to store the user category.
- User_goal: Literal field to store the user's goal per category.
- User_medical_conditions: Literal field to store user's medical conditions.

Agent-based profiling

Complementary to the user model, there is agent-based profiling which employs machine learning (ML) techniques to model the user's expertise, preferences, capacities, behaviors, and context to provide more relevant recommendations and plans to help the user to achieve their goals. Agent-based profiling constitutes the user's dynamic model and is based on data and subsymbolic (data-driven) algorithms. Below is presented a description of the methods and algorithms used to generate dynamic user profiles.

Preprocessing data

Before training ML models, data should be filtered, cleaned, and converted following the models' input specifications to extract useful information from them. Preprocessing is also an essential step for analyzing and extracting conclusions from data. The preprocessing step can vary depending on the target model and the data type. The following data types have been identified for this project:

- Tabular data: Structured data stored in matrix form, where columns represent features and rows samples (i.e., table with daily recommended food recipes with ingredients and calorie ingestion).
- Textual data: Semi-structured data represented as text in natural language (i.e., textual message in a chat between the user and chatbot).
- Image data: Semi-structured data represented as pixels' color intensities in an image (i.e., food photograph).
- Audio data: Semi-structured data representing a sound through digitalized waves (i.e., audio record sent by the user reporting daily calorie ingestion).

Machine learning models

In the Expectation project, different machine learning predictors are employed to improve the user experience and system's performance, providing relevant recommendations based on users' data collected through interactions, questionnaires, and public databases.

Rule-based personalized explainable recommender system

Rule-based personalized explainable recommender systems can produce explainable by-design recommendations and enforce strict rules like avoiding allergic food items or following dietary restrictions (i.e., vegan diet, Halal diet, kosher cuisine, among others). A PERS has been developed in the EXPECTATION project. PERS system employs the user model to generate queries to filter out those food items that can produce an allergy response or damage the user's health. Then, preferences (breakfast, lunch, dinner) and nutrition goals are employed to produce the initial recommendations, which are scored according to the user's preferences and presented to the user with a given explanation. Then, the user can provide feedback about the recommendation or explanation, which opens a negotiation process that, based on the given feedback, updates the system and the underlying rules, improves the ontology reasoning, and refine the queries and explanations. Figure 2 shows the FIPA interaction and negotiation protocol designed for the PERS subsystem.

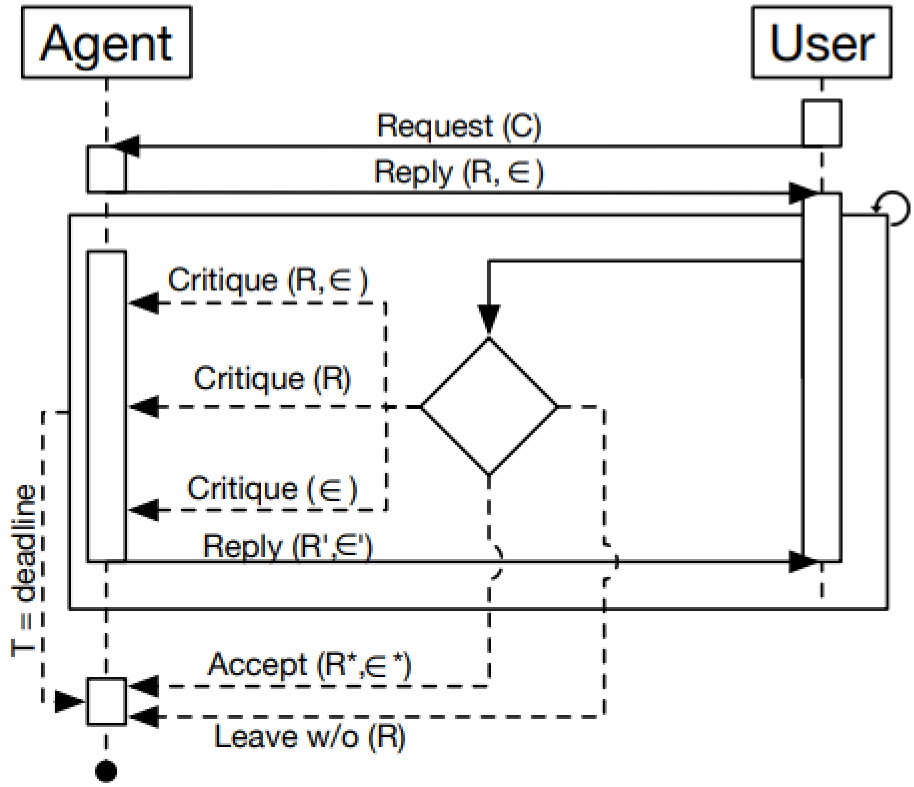

Figure 2: FIPA interaction and negotiation protocol designed for PERS subsystems [2].

Preference modeling

The "cold-start problem" is where the recommendation system lacks sufficient data about the user or has limited access to other users' data, making it challenging to create user profiles using existing information while modeling a user's preferences. To overcome this challenge, we employ a technique called Active Learning (AL). The goal of AL is to maximize the acquisition of useful information while minimizing the effort required from users. AL works by selecting informative data from a large, unlabeled dataset, asking an oracle (usually the user) to label these data points, and then updating a machine learning (ML) model based on these labeled examples.

In our approach, AL is integrated into a process that combines unsupervised, semi-supervised, and supervised ML techniques. This process involves gathering user feedback, particularly on the system's explanations for recommended items. This feedback is collected from informative samples and used to refine the underlying ML model, which estimates user preferences. The user interaction we design serves a dual purpose: it personalizes the system by incorporating user feedback into the ML model and enhances user trust by improving the system's explanations for its recommendations. We conducted human experiments to evaluate its effectiveness in the short term and experimented with various AL strategies using synthetic user profiles based on two food datasets to analyze its long-term performance. Our experimental results clearly demonstrate the efficiency of our preference elicitation approach, even with limited user-labeled data. Moreover, it enhances user trust by providing accurate explanations for the recommendations made by the system. For more information, you may refer to [4].

User preferences can be modeled as a continuous variable P_c or as a binary variable P_d (i.e., one if the user likes the recommendation, zero otherwise) that describes the user's preferences for specific ingredients, recipes, and cuisine types. To model the preference, it is necessary: i) a dataset that relates users' and meals' features and ii) a machine learning model able to calculate P_c or P_d according to the case. ML predictor models and learns the latent user's preferences by mining the user's previous interactions' database. This is a data-driven approach and differs from rule-based methods in that it can exploit information encoded in historical or aggregate data (i.e., previous user interactions and other users' selections). To generate data-driven personalized food recommendations, we can employ different methods and approaches, which we have selected the following:

- **Collaborative filtering methods:** These methods identify similarities between users and food items, assuming that users with similar characteristics will have similar preferences. Collaborative filtering has several drawbacks, like the cold start problem (i.e., the user is new in the systems, and there are not enough interactions to model their preferences) problem and privacy leaks. To overcome these limitations in the Expectation project, we have designed multi-agent collaborative filtering that preserves the user's privacy, sharing only aggregated knowledge (i.e., rules describing user preferences).

- **Deep learning methods:** Deep learning (DL) models have shown a remarkable ability to exploit latent knowledge in massive amounts of data, and it has been applied successfully to several tasks, including classification, regression, reinforcement, and ranking. In the Expectation project, we have employed deep learning models to learn the user's preference based on previous user-item interactions. Despite its generalization ability and its flexibility, deep learning models have some disadvantages: First, it is not possible to apply strict rules to filter out allergenic ingredients or apply other strong restrictions, but penalization can be applied to reduce the probability of occurrence, yet never wholly removed; The second, DL models usually are opaque and are hard to explain. To overcome these limitations, we have complemented DL models with PERS allowing strict filtering to ensure user's security and health, and as a second countermeasure against DL opacity, we have provided explanations through logic rule sets extracted from DL predictors employing the method Deep Explanations and Rule Extractor (DEXIRE) [3], which can explain the internal behavior of a DL predictor by looking in its neuron activation patterns.

Calories ingest forecast

To schedule the daily user calories ingestion, the system needs to take into account several variables, including the user's goals (i.e., nutritional and fitness), previous user behaviors (i.e., previous meals, previous calories ingested and burned), user's health conditions (i.e., BMI, allergies, illness, physical conditions, metabolic conditions, among others). Based on the collected data about the user and the food, the systems should forecast the next day's expected calorie ingestion C_i (target variable). Expected calorie ingestion is a continuous numerical variable that can be modeled through equations derived from a thorough knowledge of the nutrition domain or employing heuristic methods and regression models (i.e., Linear regression, moving average, seq2seq forecaster, among others).

References

1. Sabbatini, F., Ciatto, G., Calegari, R., & Omicini, A. (2021, September). On the Design of PSyKE: A Platform for Symbolic Knowledge Extraction. In *WOA* (pp. 29-48).
2. Berk Buzcu, Melissa Tessa, Igor Tchappi et al. Towards Interactive Explanation-based Nutrition Virtual Coaching Systems, 29 June 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-3110083/v1]
3. Contreras, V., Marini, N., Fanda, L., Manzo, G., Mualla, Y., Calbimonte, J. P., ... & Calvaresi, D. (2022). A DEXiRE for Extracting Propositional Rules from Neural Networks via Binarization. *Electronics*, *11*(24), 4171.
4. Canturk, Furkan & Aydogan, Reyhan. (2023). Explainable Active Learning for Preference Elicitation. 10.48550/arXiv.2309.00356.
5. https://www.pryv.com/