D5.2 – Advanced Coach Interaction Design

2019.10.10

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

www.nestore-project.eu

Deliverable ID:	WP5/D5.2/T5.2
Deliverable Title:	D5.2 - Advanced Coach Interaction Design
Responsible partner:	HES-SO
Contributors:	LU-CIM
Nature:	Demonstrator
Dissemination Level:	PU
File:	D5.2 Demonstrator_Template V1.0.docx
Revision:	V3.0
Due date of deliverable:	2019.08.31
Actual Submission date:	2019.10.10
call	European Union's Horizon 2020 Grant agreement: No 769643
topic	SC1-PM-15-2017 Personalised coaching for well-being and care of people as they age

Document History

Revision	Date	modification	Author
0.1	2019.07.30	Document Creation, initial TOC	Leonardo Angelini (HES-SO)
1.0	2019.07.31	Adding Emotional Engine content	Martin Sykora (LU-CIM)
2.0	2019.09.28	Adding conversational agent content	Leonardo Angelini (HES-SO)
3.0	2019.10.10	Integrating reviews and images	Leonardo Angelini (HES-SO)

Approvals

Date	Name	Organization	Role
2019.10.08	Cinzia Mambretti	PoliMI- FPM	Reviewer
2019.10.08	Ilenia Ghenio	AGE	Reviewer
2019.10.10	Giuseppe Andreoni	PoliMI	Scientific Coordinator

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

Short Abstract

This document presents the results of Task 5.2, describing the demonstrator of the conversational agent and emotion engine of the virtual coach, which allow the user to have, respectively, written and spoken conversations with the virtual coach and to understand users' affect states in spoken or written sentences.

Introduction to the demonstrator, scope of the document

The conversational agent and the emotion engine are used through different interfaces of the NESTORE Virtual Coach, in particular, the chatbot interface integrated in the NESTORE main app and the NESTORE tangible coach. The conversational agent allows to manage simple conversations with the NESTORE Virtual Coach, in written form, through the chatbot interface, and in oral form, through the tangible coach. It allows to recognize users' intents, to manage context information and it handles four different languages (English, Italian, Dutch and Spanish). Conversations are used both for coaching (suggesting activities and encouraging the user based on feedback about coaching activities) and for collecting data from the user (questionnaires, food intake self-reporting, questions about user preferences). The emotion engine allows to recognize user's affect states in specific phrases written or pronounced by the user. It is also able to handle the same languages of the conversational agent. This document presents an overview of the architecture and of the main functions of each engine. The theoretical background, the design process and the implementation details of Task T5.2 has been partially presented in D5.2.1 and will be detailed in an extended version of this document (not publicly available).

Positioning of the deliverable in the NESTORE development plan

T5.2 has been carried out as part of Work Package 5, from M4 until M24. The preliminary report D5.2.1 has been provided at M12. An extended version of D5.2.1 report will be provided at M26.

Key Words

Conversational agent; chatbot; emotions; affect detection; multi-language management.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

Table of Contents

Table of Figures

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

1 Introduction

The aim of task T5.2 is to give to the NESTORE Virtual Coach conversational and emotion recognition capabilities, in order to make it more user-friendly and to build empathy with the user. Indeed, as highlighted also in D5.1, this is particularly important for e-coaches. Co-design activities carried out in T7.2 and T7.3 also highlighted the willingness of many users to interact through voice with the virtual coach. Building a conversational agent able to understand the user's voice and text messages was therefore a key task of the NESTORE project.

1.1 Interlinks with other Work Packages

WP5's role is to provide the user with the coaching recommendations and plans elaborated in the Decision Support System (DSS, developed in WP4). D5.2 is the demonstrator of T5.2, task of WP5 devoted to the design and development of the conversational and emotional interaction with the NESTORE Virtual Coach. Recommendations and plans provided by the conversational agent are provided by the DSS, based on data collected in WP3. Coaching assumes different forms, which are based on the behaviour change techniques identified in T5.1 and orchestrated according to the HAPA model (Schwarzer et al., 2011). In T5.2, coaching is presented to the user in form of conversations. Conversations can be started both by the user, asking a question to the coach (or saying hello), and by the coach, through a notification in the app or via the tangible coach. The app, developed in T5.6, contains a custom interface for the chatbot and allows to have conversations with the coach in written form. D5.6, contains additional details about the app. The tangible coach constitutes a vocal interface for the conversational agent. Details about the tangible coach are presented in D5.3. Recommendations and feedback provided by the conversational agent are based on experts' recommendations, developed throughout WP2 and WP4 and reported in T5.6 as part of the pathway and coaching activity design. Conversations started by the coach are always prompted by the Decision Support System (DSS), developed in T4.4. The proposed conversations, in particular those implemented in the tangible coach, have been partially co-designed with users in WP7 (T7.2 and T7.3). Conversations in Italian, Dutch and Spanish have also been translated and adapted with the help of pilot sites' staff. Finally, it is worth noting that the conversational agent is used not only for coaching but also for monitoring different aspects of the user's wellbeing, complementing the work of WP3 sensors. Data collected though the chatbot (questionnaires, preferences, and food-intake) and the tangible coach (emotionally-rich sentences) are sent to the cloud and analysed by WP4 modules.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

2 Demonstrator

2.1 Overview of the conversational agent and Emotion Engine in the NESTORE system

The following diagram shows the Conversational Agent engine and the emotion engine in the NESTORE system architecture.

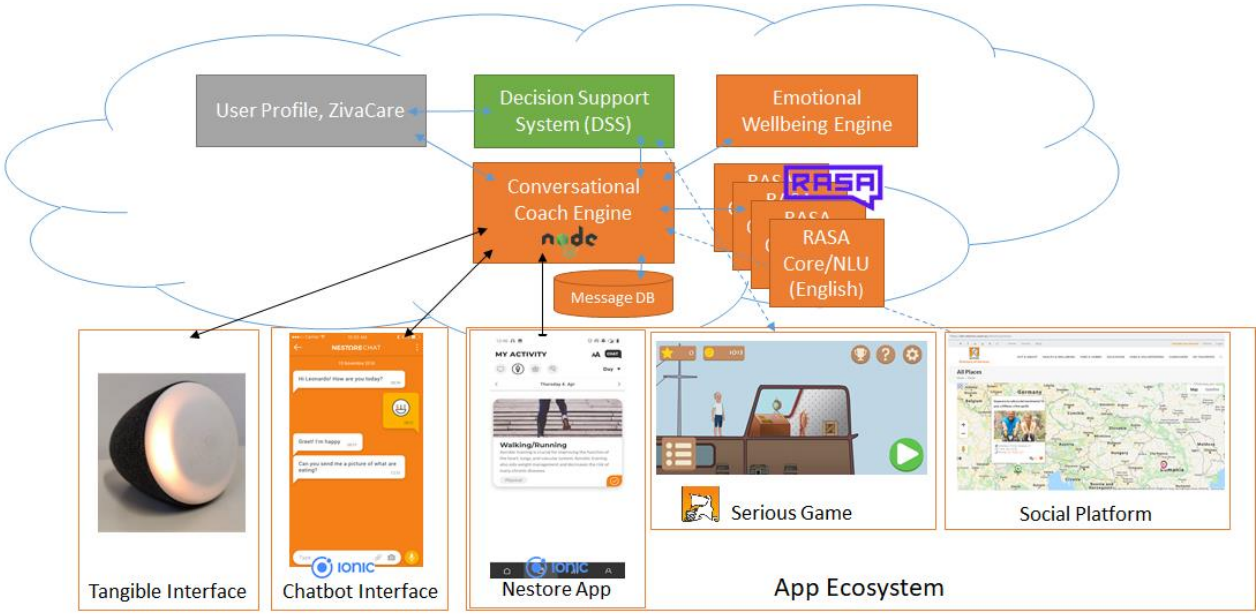


Figure 1. Architecture of WP5 interfaces with particular focus on the conversational agent and emotion engine.

The conversational agent (described in Section 2.2) animates conversations in both the tangible and the chatbot interface. Conversations can be prompted by the DSS and data are collected through the conversational agent and stored in different NESTORE APIs and in the DSS. Affect-related state analysis is prompted by the conversational agent through the tangible coach and relies on the Emotion Engine, described in Section 2.4.

2.2 Conversational Agent, Architecture and Editing interface

The engine of the conversational agent is built in Node.js[1] and is accessible both from the chatbot and the tangible coach, giving to the user the possibility to interact both through text (via the chatbot) and voice (tangible coach).

The conversational agent supports both scripted conversation flows, typically used when the coach starts a conversation (after being prompted by the DSS), and intent recognition, typically used when the user starts a conversation. Intent recognition are managed by 4 different instances of RASA[2] Natural Language Understanding (NLU) units, one per supported language (English, Dutch, Spanish and Italian). The Spacy library[3] inside RASA allows to overcome typing errors of the user or alternative word spellings, using existing language models.

[1] https://nodejs.org/

[2] https://rasa.com/

[3] https://spacy.io/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

In particular the conversational agent is implemented as shown in Figure 2:

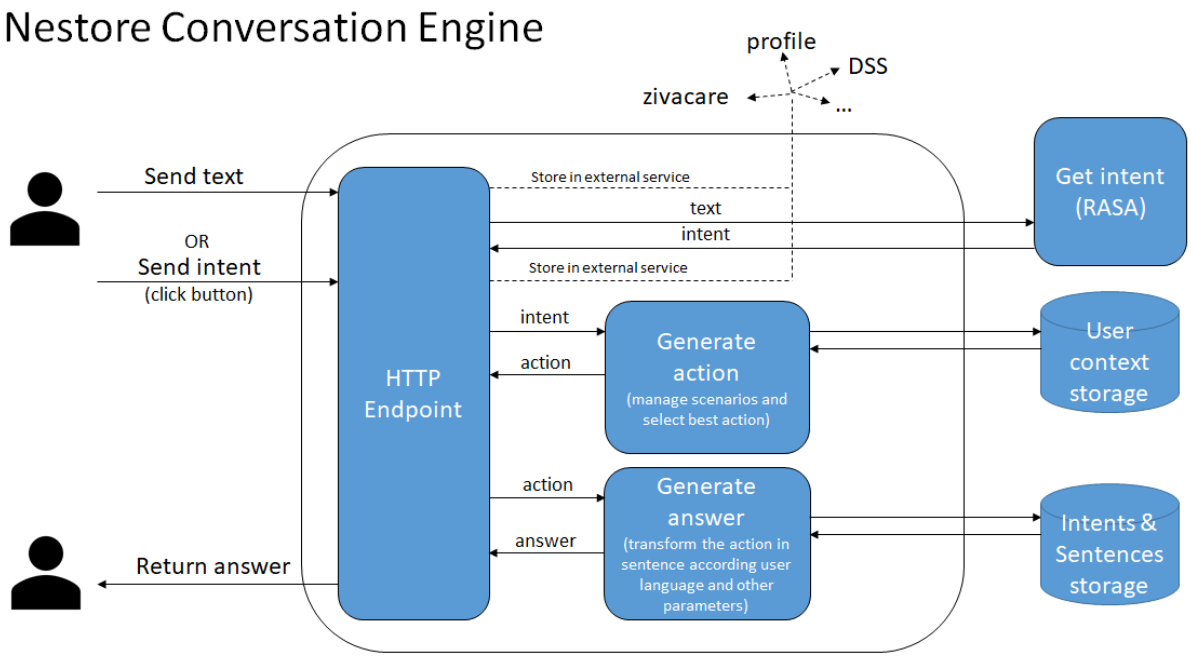
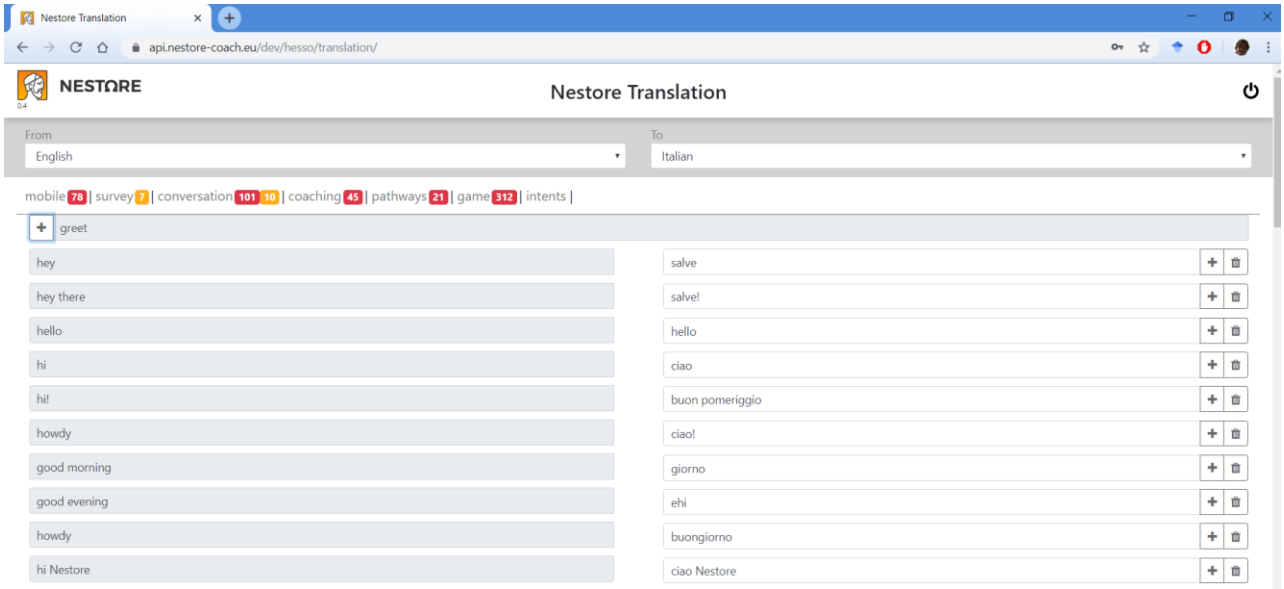


Figure 2. Internal architecture of the Conversational Engine

The HTTP endpoint allows to receive requests from both chatbot and tangible coach and allows to return an answer that, respectively, can be displayed in the chatbot interface or spoken aloud by the tangible coach. Once an intent is recognized by RASA, a first module inside the conversational agent defines the best action in respect to the detected intent, according also to the context of the conversation (current conversation scenario, previous answers, etc.). Based on the selected action, the second module generates an answer retrieving from the message database the corresponding translated sentence. Any variable related to the user profile or to the user's score is retrieved from the different NESTORE APIs and filled in the generated sentence. (Figure 2). All sentences and intents of conversations are stored in a database and are easily accessible for translation through a custom developed web interface, shown in Figure 3. This interface allows the pilot staff to carefully translate experts' recommendations and standardized questionnaires and, at the same time, to reach a high level of customization for the recognition of intents, providing example of phrases typical of their language.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

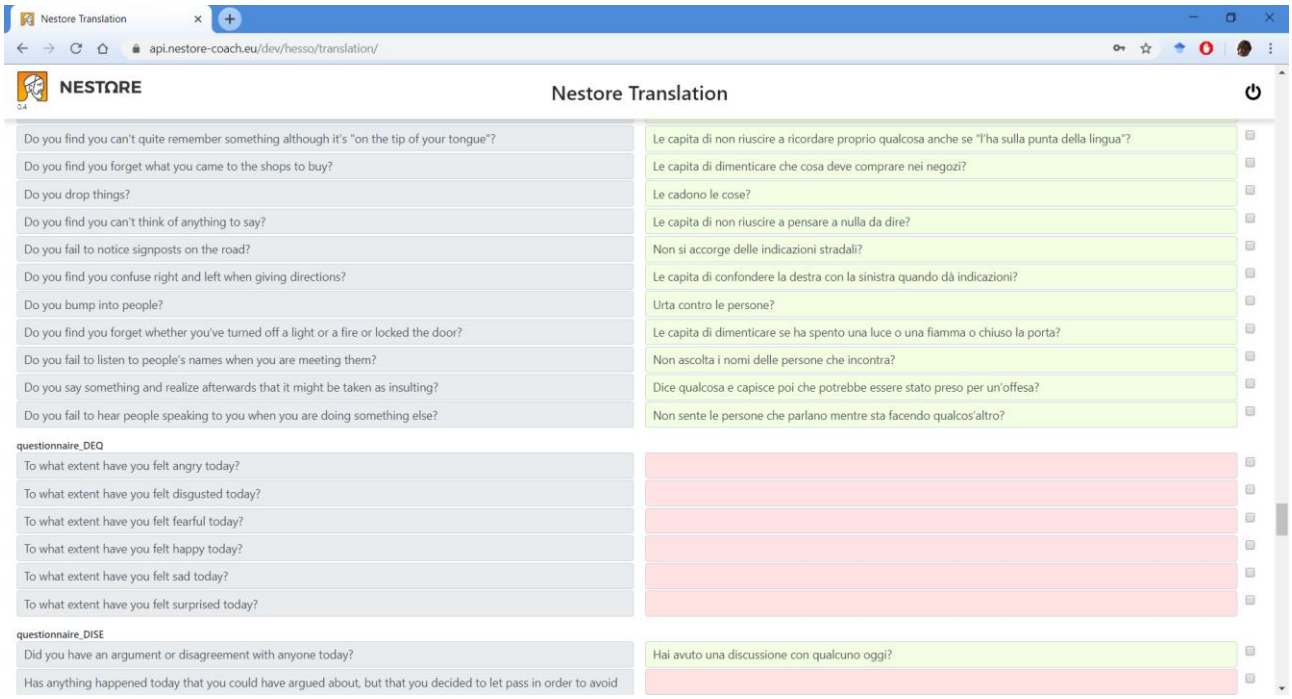


Figure 3. Web Interface for translation of intents (top) and standardized questionnaires (bottom). In green, completed translations, in red, missing translations, in orange, translations to be updated or reviewed.

Any modification or translation added to the interface is directly applied to the running system, with the only exception of intents. Indeed, intents in the translation interface are used to train a model that has to be manually deployed in the RASA server.

The aforementioned interface allows also to easily add new questionnaires, specifying among different types of answers. (Figure 4)

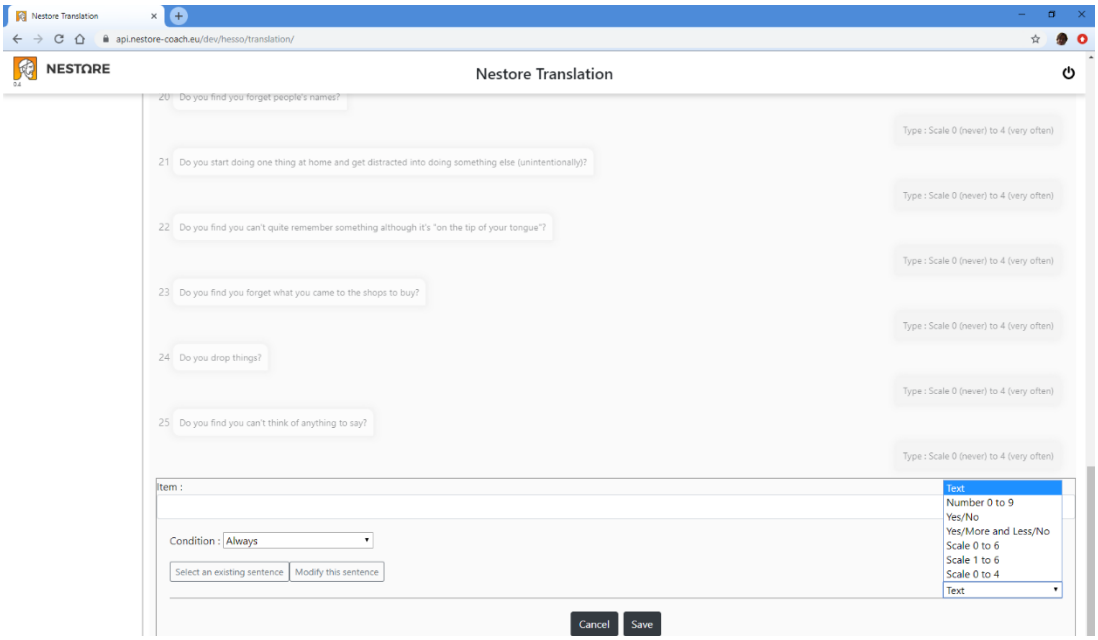


Figure 4. Web Interface for adding and editing questionnaires.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

2.3 Designed Conversations

Different conversation flows have been designed for the chatbot and the tangible coach interface, in order to allow the user to engage with both interfaces, yet providing a coherent intervention among all the users.

In the chatbot the user can:

- Get introduced to the NESTORE system and learn about how to interact with the chatbot
- Answer questions about preferences and habits (allowing the virtual coach to fill the user profile)
- Receive reminders of activities that should be scheduled or done
- Being prompted for filling in standardized questionnaires proposed by expert to monitor the user's status. These questionnaires are filled in a different view, to avoid polluting the conversations with the virtual coach, which should remain as friendly as possible
- Review activities done during the day and receive motivational messages based on the answers (Did you do the activity? Did you enjoy the activity? If not for which reason?)
- Enter food by sending a photo to the chatbot

In the tangible coach the user can:

- Ask general questions about NESTORE (Are you real? How can I talk to you? When should I talk to you? Etc.)
- Ask about the activities scheduled for today
- Ask how s/he's doing in a particular domain, receiving feedback about his/her score in the pathway chosen for that domain. The user can also ask for feedback in the emotional domain
- Talk about today's feeling, in order to monitor affect-related states

2.4 Emotion Engine

In this task LU-CIM has designed, implemented and fully populated an emotion/affective semantic model (i.e., encoded as an RDF/RDFs/OWL linguistic ontology, which we refer to as; EMOTIVE Wellbeing), which is at the core of a newly developed emotion engine, bringing together emotion models, for the sensing of raw affect variables from communicated, text based, content.

This emotion engine is exposed through an authenticated and secure REST API service (over HTTPS/TLS encryption), running on Loughborough University IT Services Infrastructure at https://nestore.lboro.ac.uk/ (swagger based documentation is available here[4]). The REST API provides automated sensing and recognition of up to 24 emotion related affective states. Several, longer-term, emotion analytics are also provided via the API, such as inter/intra-person trends and distribution of affect. Multilanguage support is now available, due to neural machine translation (NMT) and the following languages (in brackets, ISO 639-1 language code) are currently supported besides English, and these are Italian (it), Dutch (nl), Spanish (es) and Catalan (ca).

The particular affect-related states detected are listed below:

- Basic Emotion (6) → Anger, Disgust, Fear, Happiness, Sadness, Surprise.
- Emotion Related State (6) → Awe, Confusion, Frustration, Inspiration, Motivation, Shame,
- Emotional Attitude (8) → Empathy, Faith, Grief, Hate, Hope, Jealousy, Love, Trust.
- Associated Stress (4) → Acute Stress, Chronic Stress, Eustress, No Stress. *(these are represented as properties within the above states)*

The design, and population of the semantic model with a rich set of terms, phrases and phrase patterns of affect bearing expressions was a substantial undertaking. The model significantly expands and builds on the prior work in EMOTIVE (Sykora et al. 2013). Various important aspects were considered, such as the different conceptualisations and prominent models of affect, empirical results in emotions research, as well as work in the field of sentiment analysis, keeping in mind the feasibility and relevance to NESTORE virtual coach and

[4] https://app.swaggerhub.com/apis/martinsykora/wellbeing/1.0.0

characteristics of its user population. Basic emotions represent the more immediate and momentary emotional states, while emotion related states are affective states that are not considered basic nor do not have an autonomous physiological reaction, yet they do have a meaningful valence and arousal associated with them. Emotional attitudes on the other hand represent more longer lasting attitudes that tend to persist over time and have an affective dimension. Besides these 20 affective variables, each emotion, state and attitude is also mapped to a likely type of stress experience (e.g., eustress, acute stress, etc.), which is derived from work on dimensional emotion model of core-affect. The semantic EMOTIVE Wellbeing model represents these affective types and links, as well as intensifiers, conjunctions, negators, perceived emotional magnitudes and valid parts-of-speech tags which are helpful in disambiguating non-affect bearing expressions.

The semantic model currently contains over 7'000 statements, with over 1200 classes. We have leveraged a number of lexical resources and databases in its construction, including; ANEW (Bradley and Lang, 1999), the Circumplex vocabulary by Choudhury and Counts (2012), EmoLex (Mohammad and Turney, 2013), or the SentiStrength-2 Vocabulary (Thelwall et al., 2012), to name a few. We have also incorporated recently reported work, such as by de-Silva et al. (2018) from care online support groups, including various emotion expressions found through a process that was based on seed-words' similarities in a 200-dimensional word-embeddings space from a large corpus of online support groups posts text (4'795'428 posts). Being able to leverage such previous work meant the coverage/recall of the model would be improved.

The final technology stack of the solution, including the REST API service includes the Flask Framework, Apache, CentOS. As well as the individual elements, described below and highlighted in figure 5, which are all integrated within the API processing pipeline:

a. Neural Machine Translation – Besides English, also Italian, Dutch, Spanish and Catalan are currently supported via Microsoft Azure NMT (Microsoft.CognitiveServicesTextTranslation - selected region Europe/United Kingdom/South UK).
b. Emotion Recognition Model – deployed on this server, running a Python NLP and linguistic ontology model based processing engine for the efficient real-time affective language sensing from text based data.
c. Database (temporal storage) – deployed on Loughborough University IT infrastructure. Technology: MariaDB. Only pseudo-anonymised personal data for basic processing.

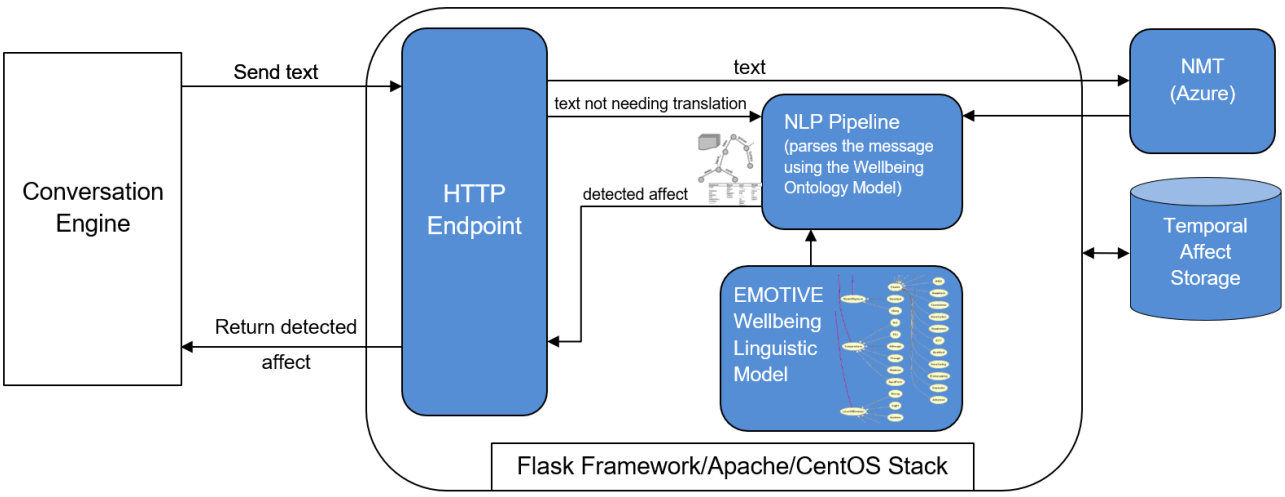


Figure 5. Internal architecture of the Emotion Engine

2.5 Link(s)

The NESTORE chatbot is accessible through the NESTORE app.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant

The NESTORE app is available for Android on the Google Play Store through a closed beta program. In order to test the app, it is necessary to join he beta program at this link:

https://play.google.com/store/apps/details?id=institute.humantech

The app is compatible with every Android device equipped with Android Version 7 or later.

A tangible coach device is needed in order to speak (in oral form) with the NESTORE conversational agent. Further details about this device are available in D5.3. A video of a conversation with the tangible coach is available at this link: https://drive.switch.ch/index.php/s/LGbB0GgzhMir13z

2.6 References

De Silva, D., Ranasinghe, W., Bandaragoda, T., Adikari, A., Mills, N., Iddamalgoda, L., ... & Gray, R. (2018) Machine learning to support social media empowered patients in cancer care and cancer treatment decisions. PloS one, 13(10), e0205855.

Mohammad S. and Turney P. (2013) Crowdsourcing a Word-Emotion Association Lexicon, Computational Intelligence, Vol. 29, No. 3, pp. 436-465

Choudhury M. and Counts S. (2012) The Nature of Emotional Expression in Social Media: Measurement, Inference and Utility, Technical Report: Microsoft

Bradley M.M. and Lang P.J. (1999) Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings, Technical Report - Center for the Study of Psychophysiology - Emotion and Attention, University of Florida

Schwarzer, R., Lippke, S., & Luszczynska, A. (2011). Mechanisms of health behavior change in persons with chronic illness or disability: The Health Action Process Approach (HAPA). Rehabilitation Psychology, 56 (3), 161-170.

Sykora M., Jackson T. W., O'Brien A. and Elayan S., 2013. Emotive Ontology: Extracting Fine-Grained Emotions from Terse, Informal Messages, International Journal on Computer Science and Information Systems, Vol. 8, No. 2, pp. 106-118

Thelwall M., Buckley K. and Paltoglou G. (2012) Sentiment Strength Detection for the Social Web, Journal of the American Society for Information Science and Technology, Vol. 63, No. 1, pp. 163-173