Chatbots4Mobile: Feature-oriented Knowledge Base Generation Using Natural Language

Quim Motger[a], Xavier Franch[a] and Jordi Marco[b]

[a]*Department of Service and Information System Engineering, Universitat Politècnica de Catalunya*
[b]*Department of Computer Science, Universitat Politècnica de Catalunya*

Abstract
Chatbots4Mobile is a research project from the GESSI research group (UPC-BarcelonaTech) which aims at designing and developing a task oriented, knowledge based conversational agent to support mobile users in the process of managing and integrating the functionalities exposed by their own application portfolio. To support the design of the required knowledge base, the project focuses on the application of Natural Language Processing (NLP) techniques to infer extended knowledge about the features exposed by a subset of mobile applications, including feature extraction from app-related documents, syntactic and semantic similarity analysis between features, and intent/entity classification focused on functionality identification from user requests. As next steps, we are focusing on the evaluation of embedded linguistic knowledge in large language models, as well as the application of granular sentiment analysis techniques to discern biased documents and process sentiment-based user feedback.

Keywords
feature extraction, large language models, conversational agents, chatbots, knowledge bases

1. Introduction

Dialogue systems (i.e., chatbots) are becoming ubiquitous tools designed to support users in a wide variety of domains, including e-commerce [1], business support [2], healthcare [3], education [4] and daily-life support [2]. A particular type of these systems are knowledge based chatbots, which provide dialogue-based access to a domain-specific, centralized information system specialized on the resolution of user inquiries and requests for a particular topic [2]. Designing effective and scalable knowledge bases becomes a challenging task from multiple perspectives, starting with the availability, collection, processing and structuring of domain-specific large amounts of data [5]. In this sense, mobile app repositories (i.e., app stores) offer a great research opportunity, providing centralized access to large metadata and natural language documents, including technical, proprietary documents from mobile app developers (e.g, descriptions, changelogs) and user-generated documents (e.g., reviews).

But beyond data availability, the state of the art in the Natural Language Understanding (NLU) field is still focused on multiple challenges related to the user experience (e.g., user adherence,

In: A. Ferrari, B. Penzenstadler, I. Hadar, S. Oyedeji, S. Abualhaija, A. Vogelsang, G. Deshpande, A. Rachmann, J. Gulden, A. Wohlgemuth, A. Hess, S. Fricker, R. Guizzardi, J. Horkoff, A. Perini, A. Susi, O. Karras, A. Moreira, F. Dalpiaz, P. Spoletini, D. Amyot. Joint Proceedings of REFSQ-2023 Workshops, Doctoral Symposium, Posters & Tools Track, and Journal Early Feedback Track. Co-located with REFSQ 2023. Barcelona, Catalunya, Spain, April 17, 2023.
jmotger@essi.upc.edu (Q. Motger)

satisfaction, learnability) with chatbots [6]. Focusing on knowledge based chatbots, potential mechanisms to overcome these challenges include designing adaptive knowledge bases [2], embedding personalization mechanisms into the model design to customize the knowledge base to the user [1] and design specific patterns for composite intent and entity recognition to resolve complex goals rather than single, atomic tasks [7]. To this end, large language models open a research window to effectively address some of these challenges by leveraging the knowledge embedded in these models towards knowledge base consumption and adaptive tasks [8, 9].

In this context, we present the project Chatbots4Mobile[1]. The main goal of this research is to study the transformation of mobile software ecosystems (e.g., smartphone devices, mobile apps, app stores) with semi-automatic cross-app feature integration capabilities using conversational agents (i.e., chatbots) to support the collection of explicit feedback towards a personalized experience for the user. In the following sections, we provide a brief background of the research group and the research project (Section 2), a summary of the system design and the work-in-progress lines of research (Section 3), as well as the research plan to continue contributing to each of these research lines (Section 4).

2. Research background

The Software and Service Engineering Group[2] (GESSI) is a research group at the Universitat Politècnica de Catalunya (UPC), located in Barcelona. Our team conducts research in multiple software engineering related areas, including requirements engineering (RE), software architecture, empirical research and open source software, among others. With respect to RE related research, the group recently participated in the Horizon 2020 OpenReq project (contract No. 644018), where the team contributed to develop, evaluate, and transfer highly innovative methods, algorithms, and tools for community-driven RE in large and distributed software-intensive projects[3]. From a scientific perspective, the GESSI research team focused especially on the design, development and evaluation of tools, services and processes based on the use of Natural Language Processing (NLP) techniques to support RE-related tasks, including requirements classification and dependency extraction from large domain-specific documents [10], improved management and automated detection of issue dependencies in large collaborative projects [11] and stakeholders recommender systems based on topic modelling and keyword extraction techniques [12], among others.

In alignment with our previous research in the field of NLP for RE, one of the scientific goals addressed by the Chatbots4Mobile project is to integrate metadata and natural language data collection, data modelling and knowledge extension techniques in order to build an NLU knowledge base for a domain-specific subset of mobile applications. The project aims at exploring the potential of state-of-the-art, pre-trained large language models to support multiple data-driven processes, mainly: (1) feature extraction from mobile app related natural language documents; (2) knowledge base generation to support mobile app feature oriented conversational processes; and (3) user intent and entity classification to enact specific cross-app feature integrations. The

[1]https://gessi.upc.edu/en/projects/chatbots4mobile
[2]https://gessi.upc.edu/en
[3]https://openreq.eu/

project started on January 2021, and its estimated end date is on April 2024.

3. System design

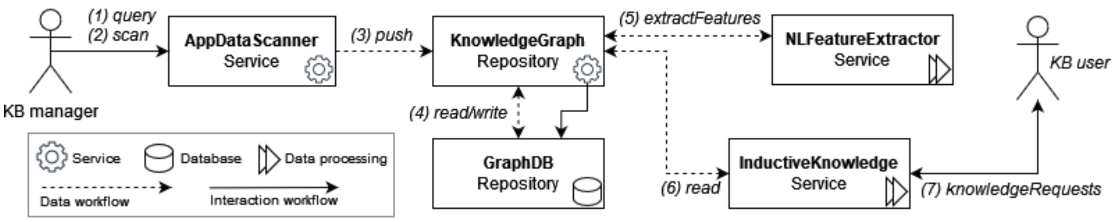

Figure 1: High-level overview of the Chatbots4Mobile system proposal limited to the NLP for RE research scoped in this report. All software components are available at https://github.com/gessi-chatbots

.

3.1. NLP-based feature extraction

Focusing on mobile software ecosystems, one of the main lines of research of the Chatbots4Mobile project is the automatic, data-driven elicitation of the set of features (i.e., functionalities) exposed by a given catalogue of mobile apps. This reverse-engineering process is fed with a large corpus of app-related natural language documents, including developer's documentation (e.g., summaries, changelogs, descriptions) and user-generated documents (e.g., user reviews). The feature extraction process is designed to serve as input for the generation of a feature-oriented knowledge base of mobile applications (see Section 3.2), as well as to fine-tune intent and entity classification models (see Section 3.3).

To this end, the group has developed the **NLFeatureExtractor Service** [13], a NLP-based feature extraction pipeline (available in the project repository). The tool leverages the embedded syntactic knowledge of a pre-trained, large language model (i.e., a RoBERTa-based model [14]) to apply linguistic annotations to the documents received as input. Based on these annotations, we use consolidated syntactic and semantic techniques (e.g., POS pattern recognition [15], dependency parsing [16]) to identify and extract the set of features covered by the given texts. Alternatively, we have explored the potential of fine-tuning large language models (e.g., BERT [17], T5 [18]) for this task. While there is an evaluation plan for both processes in progress (see Section 4.1), we already conducted some technical tests and verification tasks using a data-set of mobile apps and related natural language documents in the field of trail tracking and sports activity apps.

3.2. NLU-based knowledge base generation

The feature extraction process covered in Section 3.1 is used to conduct deductive knowledge strategies for a given catalogue of mobile apps based on the set of features extracted from app related documents. To support the design and population of this knowledge base, we designed and developed the following software components:

- **AppDataScanner Service**. A data collection service supporting the automatic extraction of mobile app natural language documents. The tool is designed to integrate multiple data sources through the combination of API consumption and web scraping techniques.
- **KnowledgeGraph Repository**. A centralized storage service based on a semantic graph database which serves as the data layer of the generated feature-oriented knowledge base.

In addition to the application of deductive extended knowledge strategies (i.e., **NLFeatureExtractor Service**), we extended the App Repository Service with inductive knowledge strategies (i.e., **InductiveKnowledge Service**). These inductive knowledge strategies are mainly based on analysing the syntactic and semantic similarities among the features exposed by the mobile apps covered in the knowledge base. Consequently, the knowledge base generation allows not only the automatic extraction of structured knowledge from natural language documents, but also explicit knowledge inference which is not explicit on a textual level [19]. The resulted, extended knowledge generation serves as a knowledge base for the design and development of a task-oriented conversational agent (for simplicity, the **KB user** in Figure 1), designed to assist users in the process of handling cross-app integrations of features from two different applications (see Section 2). To this end, we conducted a systematic literature review in the field of conversational agents, where we covered multiple scientific perspectives, including the technical infrastructure for task-oriented, knowledge-based conversational agents [5]. The aforementioned software components are already developed and available in the project repository.

3.3. NLU-based intent/entity classification

Beyond the extraction of static knowledge from a corpus of mobile app related documents, the NLP-based feature extraction process is also intended to support the recognition of features users are referring to during the conversational process with the mobile-based chatbot embedded in our system. Specifically, we are focusing on fine-tuning the intent and entity classification processes to match user requests and input messages with the set of mobile apps and features modelled in our knowledge base. Consequently, we aim at designing a customized, adaptive knowledge base specialized on a specific application catalogue to support users on specific requests about their mobile apps and the integration of features exposed by those apps. These tasks are currently in a verification and validation stage, and the associated software component (i.e., the back-end component for a task-oriented, knowledge-based chatbot in this domain) is still under development and available in the project repository.

4. Future research plan

4.1. NLP-based feature extraction

Concerning the feature extraction approach, there are three main activities planned in the short term. The first one is to conduct a quantitative evaluation analysis of the feature extraction pipeline using a data set of annotated natural language documents. We plan to combine publicly available user annotations (i.e., crowdsourced feature annotations made by real users available in sideloading repositories) with extended requested annotations of a sub set of documents made by experts and domain-related agents. The second one is based on exploring the internal knowledge embedded in large language models with respect to linguistic (either syntactic and semantic) knowledge, which can serve to improve the fine-tuning process of a deep learning based feature extraction process. This approach has been addressed by recent studies in the field for multiple linguistic tasks and different domains [20]. We also plan to extend the NLFeatureExtractor Service with a sentence-level sentiment analysis layer to filter out biased documents (i.e., subjective reviews) which might introduce noise data to the knowledge base.

4.2. NLU-based knowledge base generation

The resulted knowledge base and the underlying infrastructure to support its population and knowledge extension is already available to be consumed by a third-party software system. Hence, the next natural step is to integrate the knowledge base in order to be consumed by the conversational agent, which will be responsible for accessing and querying the knowledge base to resolve on demand user requests. Hence, after the component testing stage has concluded, we will be focusing on integration and system testing to verify the validity of the knowledge based predictions based on real user requests (e.g., queries about the level of similarity between two mobile apps, queries about the set of features exposed by a specific app).

4.3. NLU-based intent/entity classification

Motivated by the advances and next steps depicted in the previous sections, the next step on the fine-tuning process of feature-based intent/entity classification relies on the evaluation of the models and its integration with the rest of the software components and processes included in our system. Specifically, we plan to design a set of stories (i.e., dialogue patterns matching a specific interaction between the user and the conversational agent) which require knowledge base consumption of a domain-specific instance of the knowledge base. These stories will allow us not only to evaluate the validity of each of the sub-processes and components depicted below, but also to prepare an evaluation plan for the complete system in action with real users. Additionally, we plan to integrate sentiment-based feedback analysis to user requests (i.e., user intents) to extend the extracted knowledge with sentiment characteristics such as polarization, subjectivity and mood of the user. This information will be embedded in the NLU classification tasks to improve and personalize the cross-app feature integrations triggered by the system.

Acknowledgments

With the support from the Secretariat for Universities and Research of the Ministry of Business and Knowledge of the Government of Catalonia and the European Social Fund. This paper has been funded by the Spanish Ministerio de Ciencia e Innovación under project / funding scheme PID2020-117191RB-I00 / AEI/10.13039/501100011033.

References

[1] R. Bavaresco, et al., Conversational agents in business: A systematic literature review and future research directions, Computer Science Review 36 (2020) 100239.

[2] A. Janssen, et al., Virtual Assistance in Any Context: A Taxonomy of Design Elements for Domain-Specific Chatbots, Business and Information Systems Engineering 62 (2020).

[3] A. de Barcelos Silva, et al., Intelligent personal assistants: A systematic literature review, Expert Systems with Applications 147 (2020) 113193.

[4] R. Knote, et al., The what and how of smart personal assistants: Principles and application domains for IS research, Multikonferenz Wirtschaftsinformatik (2018) 1083–1094.

[5] Q. Motger, et al., Software-based dialogue systems: Survey, taxonomy, and challenges, ACM Comput. Surv. (2022).

[6] R. Ren, et al., Evaluation Techniques for Chatbot Usability: A Systematic Mapping Study, Proceedings of the International Conference on Software Engineering and Knowledge Engineering, SEKE 2019-July (2019) 479–484.

[7] S. Bouguelia, et al., Context Knowledge-Aware Recognition of Composite Intents in Task-Oriented Human-Bot Conversations, in: Advanced Information Systems Engineering, 2022, pp. 237–252.

[8] P. Xu, et al., MEGATRON-CNTRL: Controllable story generation with external knowledge using large-scale language models, in: EMNLP, 2020, pp. 2831 – 2845.

[9] D. Alivanistos, et al., Prompting as Probing: Using Language Models for Knowledge Base Construction, in: CEUR Workshop Proceedings, volume 3274, 2022, pp. 11 – 34.

[10] G. Deshpande, et al., Requirements Dependency Extraction by Integrating Active Learning with Ontology-Based Retrieval, in: Proceedings of the 28th International Requirements Engineering Conference, 2020, pp. 78–89.

[11] M. Raatikainen, et al., Improved Management of Issue Dependencies in Issue Trackers of Large Collaborative Projects, IEEE Transactions on Software Engineering (2022).

[12] C. Palomares, et al., Personal Recommendations in Requirements Engineering: The OpenReq Approach, in: Proceedings of the 29th International Working Conference on Requirement Engineering: Foundation for Software Quality, 2018, pp. 297–304.

[13] A. Gallego, et al., TransFeatEx: a NLP pipeline for feature extraction, in: REFSQ 2023, CEUR Workshop Proceedings, 2023.

[14] Y. Liu, et al., RoBERTa: A Robustly Optimized BERT Pretraining Approach, CoRR abs/1907.11692 (2019).

[15] T. Johann, et al., SAFE: A Simple Approach for Feature Extraction from App Descriptions and App Reviews, in: Proceedings of the 25th International Requirements Engineering Conference, 2017, pp. 21–30.

[16] C. Ma, et al., Content Feature Extraction-based Hybrid Recommendation for Mobile Application Services, Computers, Materials and Continua (2022).

[17] J. Devlin, et al., BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, in: Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019, pp. 4171–4186.

[18] C. Raffel, et al., Exploring the limits of transfer learning with a unified text-to-text transformer, J. Mach. Learn. Res. (2022).

[19] M. Vierlboeck, et al., Natural Language Processing to Extract Contextual Structure from Requirements, in: Proceedings of the International Systems Conference, 2022, pp. 1–8.
[20] A. Miaschi, et al., Linguistic profiling of a neural language model, in: Proceedings of the 28th International Conference on Computational Linguistics, 2020, p. 745–756.