A corpus of Swedish conversations with a librarian

Magdalena Sandahl[1]
magdalena.sandahl@kultur.goteborg.se

Simon Dobnik[2]
simon.dobnik@gu.se

Mehdi Ghanimifard[2]
mehdi.ghanimifard@gu.se

[1]Stadsbiblioteket Göteborg
Gothenburg City Library, Sweden

[2]CLASP and FLOV
University of Gothenburg, Sweden

Abstract

We describe a novel Swedish corpus of social and task related dialogues that has been collected from the interactions of library patrons with the library staff. The corpus is intended to aid the development of language technology tools, including a dialogue system, that can be incorporated in the existing chat solution to facilitate and improve interaction between patrons and library staff. We describe the properties of the corpus collected so far, its annotation with topics and evaluation of annotation.

1 Introduction

During a reference conversation a librarian may be asked to place a reservation for a book. The patron may go on and ask which book the librarian is reading right now. In this way, the conversation evolves from a simple request to a chitchat about personal genre preferences. A dialogue of this kind may occur when a user is seeking help at a reference desk in the library, via a phone call or via chat software. The goals of this project are to develop dialogue technology tools, including a dialogue system, to reduce the workload and to increase efficiency and quality of the responses. The idea of chat-bots in libraries in not new (Mohammed Ali, 2019), there are few projects in the field: for example, 'Bizzy' from University Library of Oklahoma (Young, 2019) or ANTswers from UC Irvine Libraries (Kane, 2016). An interesting study of the likely impact on human employment by AI/robotics, with a specific focus on library workers and users can be found in (Phillips, 2017), where the feasibility and acceptability are discussed. Interactions between patrons and library staff are often task-oriented but also social.

There are limited dialogue corpora for Swedish, especially in the area of task-oriented dialogue, and to the best of our knowledge there have been no previous corpora of conversations in this domain. However, several other language technology resources exist for Swedish. For example, in February of 2020, the National Library of Sweden released three pre-trained language models based on BERT and ALBERT. These models are trained on approximately 20GB of text, i.e. 200 million sentences from various sources aiming to provide a representative BERT model for Swedish text.[1] One of the aims of this project is also to try to utilise such external resources for development of new dialogue systems for Swedish.

Stadsbiblioteket Göteborg (SB, Gothenburg City Library, Sweden) uses LibraryH3lp as a virtual service software platform.[2] The software is designed as a chat system, allowing reference service and interaction between librarians and library users. One of the platform's features allows storage of chat transcripts, annotating the data, i.e. tagging the chats with descriptive categories as well as downloading metadata in CVS format for further analysis. The platform also allows integration of external tools in conversations with patrons through APIs. The goal of the ongoing corpus collection and annotation is to aid the development of language technology resources such as topic modeller and chatbot that will be incorporated in LibraryH3lp, save some resources and predominantly heighten the service quality and quantity. By developing this corpus, however, we are also proving a novel and interesting conversational resource for Swedish.

2 Data collection and annotation

Corpus collection and annotation started in January 2020. The data is still being collected and incrementally annotated. As of the middle of June 2020, the dataset comprises over 1600 conversations with

[1]https://github.com/Kungbib/swedish-bert-models

[2]https://libraryh3lp.com/

Proceedings of the 24th Workshop on the Semantics and Pragmatics of Dialogue, July 18-19, 2020, Online, hosted from Massachusetts, USA.

ca 15 chats added daily. The dataset is incrementally annotated for topics relevant in the domain of library conversations, for example: *book and subject search*, *app Libby* and *purchase requests* and up to now the annotation has converged on 30 topics which are incrementally revised. All annotation was carried out by the same librarian and was later validated by other librarians (Section 3). As our goal is development of tools that would assist communication between patrons and librarians, we opted for annotating entire conversations for topics. Depending on the detected topic, the tools would either answer the question or chitchat with the patron or rout the query to a librarian (Smith et al., 2011; Pulman et al., 2010).

Here is an example of a summarised conversation. A patron gets a written overdue library book notice. She contacts the library through chat to ask about the notice, claiming that she had already returned the book. After a short conversation, checking the patron's library account and explaining the rules for overdue fines, the chat evolves from a task-oriented to social-oriented when the patron comments that they enjoyed interacting with a librarian through chat: "Ok! Tack för hjälpen! Så bra att kunna chatta!" This chat is annotated with the following topics: (i) check out, extend loan and check in, (ii) overdue fines, and (iii) social chat. The tags refer to simple actions (*naxos*, *app PressReader*) or complex ones (check out, extend loan and check in) and related to and motivated by how the work of librarian is organised. In addition to the topics tags, a dialogue may also be flagged with the label *admin* (when the conversation is carried out between two librarians) or *split chat* (when either the librarian or the user does not respond or when the interaction gets split into two and continues as a new one due to technical problems).

3 Corpus validation

As mentioned above, the tag set is not predefined and can be expanded. Hence, after additional three weeks from the beginning of data collection, a poll was sent out to 21 librarians working with the LibraryH3lp chat service, in order to validate the corpus and verify the topics. The poll was anonymous and all librarians responded to it. They received 18 authentic anonymised chat interactions that were randomly selected from the saved data.

The chats were presented as two tasks. In the first task which consisted of 15 dialogues, each respondent had the option to choose one of the four topics that in their opinion best describes the conversation. They also had a possibility to create their own tag instead in case they would disagree with all four pre-given choices. Three of four topics were relevant to the conversation but to a different degree (the goal was to measure annotator preference, e.g. *app Libby* or *suggestions, opinions and questions*), the fourth one was chosen at random to check if the participants are providing useful answers. The results in Table 1 indicate that dialogues are not restricted to a single topic and must be assigned multiple classes which has implications for computational modelling. In the second task of the study, the librarians were presented with three chats which they were asked to tag freely. The results were compared with the original annotations and ten new additional topics were identified.

Dialogue	Majority topic	%
D1	låneregler	0.571
D2	app uTalk	0.952
D3	fysisk utrustning & verktyg	0.905
D4	bok- & ämnessökning	0.810
D5	bok- & ämnessökning	0.952
D6	reservationer	0.667
D7	TV-spel	0.667
D8	förseningsavgifter, krav & inkasso	0.857
D9	lån, omlån & återlämning	0.476
D10	programpunkter & kalendarium	0.952
D11	app PressReader	0.524
D12	fjärrlån	0.571
D13	app Legimus	0.952
D14	bibliotekskort & PIN-kod	1.000
D15	grupprum	1.000
D16	Utskrifter	0.095
D17	cineasterna	0.238
D18	bokcirkelkassar	0.143

Table 1: The agreement (in %) of domain experts on the most likely topic of each dialogue. The results indicate that several dialogues should be assigned more than one label as several topics are discussed.

4 Conclusions

Our future work will focus on (i) further extension of the corpus and refinement of the tags; (ii) application of the corpus in computational modelling; (iii) examining the relation between domain specific dialogue (specific tools and procedures at SB), task specific dialogue (conversations between patrons and librarians) and general social dialogue; (iv) releasing the corpus to the public domain so that it will be useful both for libraries and researchers in dialogue.

Acknowledgements

The research of Dobnik and Ghanimifard reported in this paper was supported by a grant from the Swedish Research Council (VR project 2014-39) for the establishment of the Centre for Linguistic Theory and Studies in Probability (CLASP) at the University of Gothenburg.

References

Danielle A Kane. 2016. The role of chatbots in teaching and learning. *E-learning and the academic library : essays on innovative initiatives*, 131:1–26.

Salihin Mohammed Ali. 2019. Bots in libraries: They're coming for your jobs (or is it?). In *ALIA Information Online 2019*, Sydney, Australia. Research Collection Library, University of California, Berkeley School of Law.

David Phillips. 2017. Robots in the library: gauging attitudes towards robotics and AI, and the potential implications for library services. Master's thesis, City University of London, January.

Stephen G Pulman, Nigel Crook, Ramon Granell, Simon Dobnik, and Manjari Chandran-Ramesh. 2010. Deliverable D4.2.3: Report on topic detection with Dirichlet Process Priors. Technical report, Department of Computer Science, Oxford University, Oxford, United Kingdom.

Cameron Smith, Nigel Crook, Simon Dobnik, Daniel Charlton, Johan Boye, Stephen Pulman, Raul Santos de la Camara, Markku Turunen, David Benyon, Jay Bradley, Björn Gambäck, Preben Hansen, Oli Mival, Nick Webb, and Marc Cavazza. 2011. Interaction strategies for an affective conversational agent. *Presence: Teleoperators and Virtual Environments*, 20(5):395–411.

Jeffrey R Young. 2019. Bots in the library? Colleges try AI to help researchers (but with caution). Technical report, EdSurge.