



FREIE UNIVERSITÄT BOZEN  
LIBERA UNIVERSITÀ DI BOLZANO  
FREE UNIVERSITY OF BOZEN - BOLZANO



Faculty of Computer Science, Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100  
Bolzano, Italy

Tel: +39 04710 16000, fax: +39 04710 16009, <http://www.inf.unibz.it/krdb/>

*KRDB Research Centre Technical Report:*

## The BoB Dialogue Corpus

Manuel Kirschner

<b>Affiliation</b>	KRDB Research Centre for Knowledge and Data, Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100 Bolzano, Italy
<b>Corresponding author</b>	Manuel Kirschner <a href="mailto:kirschner@inf.unibz.it">kirschner@inf.unibz.it</a>
<b>Keywords</b>	Natural language processing, dialogue systems, interactive question answering, linguistic corpora
<b>Number</b>	KRDB10-4
<b>Date</b>	07-09-10
<b>URL</b>	<a href="http://www.inf.unibz.it/krdb/">http://www.inf.unibz.it/krdb/</a>

©**KRDB Research Centre for Knowledge and Data.** This work may not be copied or reproduced in whole or part for any commercial purpose. Permission to copy in whole or part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of the KRDB Research Centre, Free University of Bozen-Bolzano, Italy; an acknowledgement of the authors and individual contributors to the work; all applicable portions of this copyright notice. Copying, reproducing, or republishing for any other purpose shall require a licence with payment of fee to the KRDB Research Centre.

# 1 Introduction

This document aims to describe the BoB Dialogue Data corpus, a constantly growing collection of human-machine interaction logs collected by BoB, the Bolzano Library Bot on the web-site of the Library of the Free University of Bozen-Bolzano (<http://www.unibz.it/library>).

The main distinguishing features of our corpus are its **realistic domain**, i.e., the dialogues reflect actual interactions of genuinely motivated users of our University Library, and its **multilinguality**, since the system interacts in the three official languages of the Free University of Bozen-Bolzano, namely English, German and Italian. We would like to encourage the use of the BoB Dialogue Data by the research community. The corpus might be of particular interest to researchers in the fields of Linguistics, (Interactive or contextual) Question Answering (QA) or Dialogue Systems.

Please refer to <http://bob.iqa-dialogues.net> for the latest version of this document, which incorporates the most recently collected dialogues.

## 2 The BoB project

### 2.1 Overview

BoB, the Bolzano Bot, is a multilingual chatbot we developed and deployed on the web-site of the University Library<sup>1</sup> of the Free University of Bozen-Bolzano. It has been serving library users since September 2008. As most other chatbots, it is based on simple yet robust pattern matching techniques for analyzing user input. For each question pattern that matches a user's question (or, more generally, an utterance), BoB returns a canned-text answer associated with this particular question pattern. Our main goal when designing BoB was to create a tool that would allow us to collect Interactive Question Answering (IQA) dialogues from real, genuinely motivated users. The BoB web application (although without the question/answer repository, which was developed by the domain experts of our library) is available under an open source software license at <http://code.google.com/p/chatbot-bob>.

The BoB project is a joint venture between the KRDB Research Centre at the Faculty of Computer Science, and the University Library of the Free University of Bozen-Bolzano. Besides providing an additional service to library users, the project has led to different research activities, and so far was the basis for one PhD thesis [4] and two Masters theses [3, 7].

### 2.2 BoB's question/answer repository

BoB's question patterns and its system answers were developed by a team of domain experts at our University Library. They cover a wide range of library information topics, e.g., opening time, lending procedures and different library services. While being rather simplistic regarding the underlying theories of language and dialogue, it is an adequate baseline system that we developed in around one year's time, that is easy to tailor and maintain by library domain experts, and that allows us to collect realistic IQA dialogues in three languages. Since the development of BoB's question/answer repository involved considerable manual work by the library domain experts, it is not released to the research community as part of the open source BoB system. Please get in touch with the University Library ([library@unibz.it](mailto:library@unibz.it)) to inquire about licensing possibilities.

The mapping from user input to system response is done on the basis of regular expression patterns; in BoB, each regular expression pattern for matching certain user input is associated with its corresponding "canned-text" system response<sup>2</sup>.

---

<sup>1</sup><http://www.unibz.it/library>

<sup>2</sup>Note that system responses frequently "over-answer" the user question, i.e., they might contain the information originally requested by the user, but possibly within a longer sentence, or even a segment of a few sentences.

The regular expression used in the question patterns are based on Perl regular expression (perlre) syntax, extended with the boolean operators AND, OR and NOT. With the help of the latter, the question patterns can capture a wide range alternative formulations of a given question class, in that they can specify the presence or absence of certain keywords in the user question independently of their location in the question. Table 1 provides a typical question pattern for English, along with the corresponding system answer that BoB returns when that pattern matched the user question. Note that BoB’s question pattern syntax foresees the use of “macro references”, marked within hash signs (`#`) and highlighted here in blue, which were used by the domain experts to encode either syntactic or semantic word classes. The boolean operators that extend standard perlre syntax are highlighted in red.

Question pattern	((( <code>#WHY#</code> (am <code>#I#</code> not able can’t <code>#I#</code>  am <code>#I#</code> not allowed) <code>.*#ST_BORROW#</code> )) <code>&amp;&amp;</code> <code>(#ST_REFERENCE_LIBRARY#</code> ))    <code>(#ST_LITERATURE#</code> ( that)? (is are) <code>.*#ST_REFERENCE_LIBRARY#</code> )    <code>(^#IS# #ST_REFERENCE_LIBRARY# reference only</code> ))    <code>#ST_LITERATURE# #IS# #NOT# to borrow</code>    <code>#ST_LITERATURE#</code> <code>&amp;&amp;</code> <code>“can’t be borrowed”</code>    <code>#ST_LITERATURE#</code> <code>&amp;&amp;</code> <code>#ST_REFERENCE_LIBRARY#</code>
Answer	Well, some books like encyclopedias or compulsory readings for lectures can’t be borrowed or reserved, because they have to be available to students all the time.

Table 1: Example English question pattern and corresponding answer

BoB’s question/answer repository comprises (at the time of writing) around 750 Q/A pairs for “normal”, context-independent questions as the one shown in Table 1. An additional 500 Q/A pairs capture context-dependent user questions, or user answers to system-initiated questions — these special types of questions are described in the following section.

### 2.3 BoB’s dialogue-enabling features

There are two features of BoB’s answer search algorithm that extend BoB’s otherwise stateless question-answer mapping mechanism. These dialogue-enabling features were manually assigned to around 500 of the around 1,250 question-pattern and answer pairs in the question/answer repository.

**Question patterns defined specifically for context-dependent Follow-Up Questions (FU Qs)** One of the distinguishing features of the dialogues between a user and BoB that the latter allows the user to pose questions that are related in certain ways to the previous dialogue. We call every question in an IQA dialogue a FU Q if there exists at least some previous user question or system response, since all follow-ups are potentially related to the dialogue context.

At this point, we are interested only in the subset of FU Qs that are *context-dependent*, i.e., that *require* some additional information from the dialogue context in order to be fully specified and unambiguous. BoB’s repository of question patterns includes dedicated question patterns that specifically cover context-dependent FU Qs after a specific previous user question and system answer. These question patterns carry a specific “local” attribute, which is interpreted by the question pattern search algorithm in that it searches these patterns first, and only in the specific dialogue context for which the domain experts had encoded the pattern.

On the one hand, this obviously makes the QA task easier, but also gives the domain experts a chance to present the user different related options that she could pursue next.

What follows is an example taken from the BoB question pattern repository. After the user asked about the availability of guided tours to the library, she asks an elliptical – and thus context-dependent – FU Q<sup>3</sup>: “Where is the meeting point?”. This FU Q is matched by a “local”-marked regular expression pattern, which can be rather indiscriminately defined as accepting any string that begins with the word “where”.

**Question patterns defined for answers to system-initiated sub-dialogues** In an interactive QA or dialogue setting, both the system and the user can pose FU Qs. In the case of a system-initiated question, it most frequently serves as a request for more information from the user, when the system failed to determine a (unique) answer to her original question. This is known as a *clarification question* by the system. The goal behind the system-initiated clarification dialogue can be either constraint relaxation (when no answer could be found initially), or constraint refinement, i.e.: a further specification of the parameters, so as to yield a complete query to the question/answer repository.

In BoB, sub-dialogues are used to encode relatively short, predefined sequences of system questions, to which the user’s answer must come from a rather small, specifically predefined set of possible answers (e.g., “yes/no” for the simplest kind of question). From the domain expert’s point of view, they are used either to implement system-initiated clarification requests, or to guide the user through the library domain in a conversation style where the system takes the initiative. In the first case, an ambiguous user question would trigger a clarification sub-dialogue, while in the second case, the system response provides the user with a choice of related topics about which she could then ask in more detail. This kind of guidance should be especially useful for inexperienced users who do not know how to formulate their problem or question explicitly.

Sub-dialogues must be handled explicitly by the search algorithm, since their regular expression patterns must only be searched when the corresponding sub-dialogue is active. This prevents question patterns with very unspecific patterns (e.g., “/yes — ok/”) to be selected in the course of the global question pattern search.

**BoB’s answer search algorithm** We now describe BoB’s search algorithm, and how it works in conjunction with the two dialogue-enabling features introduced above. Every time the user enters a new question, a suitable question pattern has to be identified, so that the system response stored in that specific topic node can be returned as an answer. The search for a question/answer pair depends both on the user input and on the previously selected question/answer pair, along with the information whether BoB is currently in a sub-dialogue.

Conceptually, the search algorithm can be divided into two parts. In the first part, sub-dialogues and context-dependent FU Qs are dealt with on a local level. In the second part, the search for a system response is extended to the entire repository of non-local and non-subdialogue question patterns.

The search begins with step 1 of the algorithm, and continues to apply the following steps in order (unless explicitly stated otherwise in the algorithm), until a system response has been determined, the dialogue state has moved to a (new) question pattern, and possibly the sub-dialogue flag has changed. The first two steps of the algorithm cover the two special types of question patterns introduced above – sub-dialogues and context-dependent FU Qs – since they need to be treated specially and override the global tree-searching step 3.

1. [**Continue sub-dialogues**] Applies only when in sub-dialogue mode. Search all question patterns marked as “SD” (Sub-Dialogue), and connected to the previous user question/system answer, for a question pattern matching current user input.

---

<sup>3</sup>In the surface form, the FU Q is lacking the attribute “... for guided tours?”.

- (a) [**Match**] From a matching question pattern, return its associated system response. Exit.
  - (b) [**No match**] Leave sub-dialogue mode (sub-dialogue flag is disabled).
2. [**Context-dependent follow-ups**] search all question patterns connected to the previous user question/system answer that are marked as “context-dependent” (= “local”). If successful, return associated system response. Exit.
  3. [**Global question/answer repository search**] search all non-sub-dialogue and non-local question patterns in BoB’s question/answer repository for a matching pattern. If more than one question pattern matches, apply BoB’s machine learning-based answer re-ranker to the set of possible answer candidates, and determine the best answer.
    - (a) [**Match**] From the (highest-ranking) matching question pattern, return its associated system response. Exit.
    - (b) [**No match**] Return an apology message stating that BoB did not understand the question. Exit.

If step 3 identifies more than one matching question pattern for a given user input, BoB’s algorithm uses a supervised machine learning component to determine the most promising of all matching answer candidates, based in part on lexical similarity features between the user’s question and each answer candidate.

### 3 Description of BoB dialogues

#### 3.1 An example dialogue

We start off with an overview of some important aspects and properties that characterize the IQA dialogues collected through BoB. This is best achieved by providing a dialogue between a user (U) and BoB (B) as an example. We chose this particular dialogue because it exhibits many interesting phenomena due to its above-average length. For presentational reasons, we abridged the dialogue at three positions, namely at lines 10, 19 and 38. The dialogue is taken verbatim from the BoB log files, thus including all the original typos. We only made the following two formatting changes to improve the readability here: in the actual BoB web interface, words that in the example below are contained in angle (< >) or square ([ ]) brackets are in reality displayed as user-clickable links. When the user clicks on an “angle bracket” link, words contained therein are immediately submitted to BoB as a new question, appearing also in the chat history window as if the words were entered regularly by the user. On the other hand, clicking on a “square bracket” link opens a pop-up browser window displaying the URL that the library expert writing the BoB system answers had encoded as meta-data for the particular link at hand. Note that the typos contained in the example represent the actual typos made by the user.

```

1 B: Hello, I am BoB. I can tell you how to use the Library’s services e.g., how
    to search the OPAC, borrow a book or where to find things in the Library.
    Please ask me questions in English - you will help me to learn!
2 U: where can I find books about morphology?
3 B: The OPAC offers two possibilities to search by topics. Both can be found in
    the advanced search menu. Searching for subject is the first possibility.
    And the second one ...
4 U: how do I use opac
5 B: First of all: You can access the OPAC wherever you have internet access.
    More details depend on what you are looking for ...
6 U: i like to find and borrow a specific book
7 B: Could you rephrase your question, please.
```

8 U: how do I find and borrow a specific book

9 B: If you want to borrow books, DVDs or journals you need to take them from the shelf yourself and check them out at the information desk. If you want them to be sent to Brixen-Bressanone, Bozen-Bolzano, Bruneck-Brunico or Schlanders-Silandro you can [order] them via the OPAC.

10 ...

11 U: how will I know when the books I ordered are ready to be borrowed

12 B: If you stated while ordering, that the book should be reserved for you, we will contact you as soon as it arrives.

13 U: How will you contact me?

14 B: Gosh, I didn't quite get that. Strange, since I am actually pretty intelligent. Well, I will ponder upon that further and do some research, so the issue will make more sense to me on your next visit.

15 U: when are the libraries open

16 B: You can find our current opening hours on [this] website.

17 U: are there special holidays when the libraries are closed

18 B: On holidays my colleagues are allowed to stay at home. I'm of course always here. As well as our online offer. You can find our current opening hours on [this] website.

19 ...

20 U: for how long can I borrow library material

21 B: Loan periods depend on the type of medium. Do you have a <normal book>? Or is it something from the <textbook collection>? Or maybe a <journal>, a <dvd> or something ordered from another library via <Interlibrary Loan>?

22 U: normal book

23 B: Books can be kept for four weeks. And you can extend the loan period twice - again for four weeks. But this is only possible if the book is not already reserved by someone else.

24 U: what happens if i missed to return a book

25 B: First of all you have to tell us - that's important to avoid further overdue fines. If it doesn't reappear you have to buy a new copy. If that's not possible you have to pay the price of the book.

26 U: do i have to pay if I bring the book later than the due date

27 B: Gosh, I didn't quite get that. Strange, since I am actually pretty intelligent. Well, I will ponder upon that further and do some research, so the issue will make more sense to me on your next visit.

28 U: what are the prices if i give back a book too late

29 B: Overdue fines are 2,50 EUR per item per week. If you want to know more about fees read our charges regulation.

30 U: how is the library organised

31 B: I will ask to get updated with this information.

32 U: what is the infrastructure of the library

33 B: Well, in the traditional sense a library is a place where books are stored. But there is a lot more behind libraries...

34 U: what exactly - can you give more details

35 B: There are not only books but also lots of electronic resources. You can find a broad offer of electronic information resources at our Library. Either visit our website or pop in for a chat with the librarians.

36 U: how can I chat with them

37 B: I guess I would be allergic to all sorts of animals because me and my electronic parts are rather sensitive. But still I like watching those cute little puppies.

38 ...

**BoB's answers.** It is evident from this example dialogue that the type of answers from BoB have little resemblance to those of simple factoid QA systems, where the answers consist of few words. BoB's answers consist of one to several sentences, and often over-answer the user question. In this sense, BoB can be compared to other QA systems that return Frequently Asked Questions (FAQs) type answers [5, 6]; the particular style of answers has been referred to as belonging to *encyclopedic QA* [8]. BoB's answers span the whole topic spectrum covered by

a typical library information help-desk. Its answer repository might contain overlapping (with respect to their information content) answers, or answer duplicates. BoB also covers a good amount of typical small-talk topics (example dialogue, Line 37 — although in this case it is retrieved due to the user having mistakenly entered “cat” instead of “chat” on Line 36), which is regularly put to use in the collected dialogues. Finally, a small feature of BoB’s answers is the possibility for them to contain clickable links (see passages marked in square brackets on Lines 9, 16 and 18), which lead to the described web page being opened in a new browser window.

**System-initiated clarification sub-dialogues.** Upon certain user questions, BoB initiates a clarification sub-dialogue by presenting the user with a system answer that contains a small number of suggestions for the user’s next input (see the words in angle brackets in our example dialogue on Line 21). The user normally answers such questions with an utterance containing one of the proposed topics; the BoB interface presents these topic terms, which were marked accordingly by our library domain experts, as clickable links, which lead to the contained words to be submitted to BoB as the next user input. On Line 22, the user typed in — or clicked on — one of the suggested topics, thus continuing the clarification sub-dialogue.

**Users continue or shift between topics.** Within a single dialogue, users might shift between different unrelated topics (example dialogue, Line 30), or they might ask topic continuation FU Qs (Lines 17 and 26). We consider rephrasing FU Qs a sub-type of topic continuation FU Qs; users typically ask these after a system answer which did not answer their original question (Lines 28 and 32). Another special case of topic continuation FU Qs are those containing context-dependent discourse phenomena. In our example dialogue, we find two cases of such context dependency of the FU Q: a case of (semantic) ellipsis on Line 13, and an anaphoric pronoun on Line 36.

**Unconstrained user input.** A final important property of BoB IQA dialogues is the fact that user input is completely unconstrained. This leads to user input consisting merely of keywords (rather rarely), or much more frequently to user input containing typos (Lines 8, 30 and 36) or ungrammatical sentence or question form. Together with the fact that user question types are often complex (as opposed to simple, factoid QA), this renders a simplistic classification of the question type based on wh-words moot.

### 3.2 Meta information from BoB dialogue management

When BoB interacts with a user, it keeps log files of the IQA dialogue. First of all, these logs include a timed protocol of user input and BoB’s responses: the user and system utterances are the literal part of the information. On the other hand, BoB also logs several dimensions of meta information, which are all based on BoB’s internal status of its *dialogue management* routine.

As described in Section 2.3, there are two special, dialogue-related situations in which BoB’s search algorithm does not do a full search through all of BoB’s question patterns: context-dependent FU Qs, and system-initiated sub-dialogues.

**Context-dependent FU Qs (“local” question patterns)** For each answer it returns to the user, BoB keeps track of whether this answer was part of a question pattern marked as context-dependent (“local”).

**System-initiated sub-dialogues** For each user input that BoB receives when it is in sub-dialogue mode, it keeps track of whether the input follows the pre-defined sub-dialogue structure. One of the following two cases applies:

1. *Continue*. The user input matches one of the regular expression patterns specifically designed for matching the expected answers to the particular system clarification question. This means that the user is following along the route suggested by the system, having chosen one of the proposed branches to continue the sub-dialogue
2. *Break*. The user decides to leave the sub-dialogue, without providing one of the expected answers to BoB’s clarification question.

**Apology message (user question not understood)** There is another piece of meta information related to BoB’s internal dialogue management status that we keep track of: whether a particular system answer was an apology message telling the user that BoB did not understand their previous input. Internally, this message is chosen as a system response whenever the previous user utterance did not match any of the hundreds of regular expression patterns defined in BoB’s repository. One of a small number of apology responses<sup>4</sup> is given at random, to add some variation. If after having issued one of these messages the immediately succeeding user input again does not match any of BoB’s question patterns, the system assumes that the user is lost, and issues a different apology message<sup>5</sup> that provides a generic overview of what the user can ask BoB about (again marking the system answer’s special status as part of the answer’s meta information).

### 3.3 Meta information from human annotation

We now turn to the meta information in the BoB Dialogue corpus that stems from post-hoc human annotation. For a small portion of the logged dialogues we have been adding different levels of manual annotations.<sup>6</sup> Note that while the currently available human annotation is in no way complete nor consistent, and thus only of limited use, we encourage the community to help us gather more annotations, or devise a more consistent annotation scheme.

The following paragraphs explain the notions behind the individual levels of human annotation. We begin with the distinction of FU Q types. Our particular choices of types proposed to the annotator are due to previous work in the field of Interactive QA: from [2] and [9] we adopt the distinction between Topic Shift and Topic Continuation, while from [1] we use the notions of rephrases and context dependency.

---

<sup>4</sup>The actual set of BoB’s responses:

- Sorry, I don’t understand.
- Could you rephrase your question, please.
- I will ask to get updated with this information.
- Gosh, I didn’t quite get that. Strange, since I am actually pretty intelligent. Well, I will ponder upon that further and do some research, so the issue will make more sense to me on your next visit.
- I don’t understand that, but it’s not a biggie. My creators check on me regularly, and I will ask them to resolve this issue. They will surely take care of it.

<sup>5</sup>BoB’s actual response:

Let me try to help you out here. I can tell you something about the organization of the University Library. Borrowing, reserving, returning, renewing items. Using the OPAC. Searching for books, journals, and articles. Using the Interlibrary Loan Service.

<sup>6</sup>All annotations of FU Q type and question relevance were performed either by the author a colleague computational linguist, while all annotations of answer correctness were contributed by one of our library domain experts.

**Topic shifts:** mark a FU Q as a *topic shift* or a *topic continuation* based on your intuitive notion of whether the FU Q “switches to something completely different”, or “is in some way related to the previous dialogue”.

**Rephrase:** mark whether you consider the FU Q to be an attempt to re-formulate the same question. The FU Q could be a literal repetition of the previous question, or it could be a rephrasing.

**Context-dependent FU Q:** tag a FU Q as *context-dependent* if you believe it needs to be considered along with some information provided by *the dialogue context* in order to be correctly understood.

**Fully-specified FU Q** tag a FU Q as *fully specified* if you believe it does not need any information from the *the dialogue context* in order to be correctly understood.

**Question relevance:** classify a user question by its well-formedness and its relevance for the library help-desk domain, following these annotation instructions until one of the three criteria matches:

- if the user question is problematic, e.g. typo, strange syntax, nonsense, keyword only, information-seeking about some non-library topic  
→ mark it as *irrelevant*
- if the user question is information-seeking *and* about a library topic  
→ mark it as *relevant*
- if the user question *not* information-seeking  
→ mark it as *semi-relevant*

**Answer correctness:** mark whether you consider the system answer (which could also be an apology message stating that BoB did not understand the user’s question) OK for the given question; if you find BoB’s answer was not correct in the first place, provide the correct answer BoB should have returned instead.

### 3.4 Description of columns in BoB log file tables

Table 2 gives a complete overview of all fields/columns comprising the BoB Dialogue files.<sup>7</sup> Please confer to [bob.iqa-dialogues.net](http://bob.iqa-dialogues.net) for examples of actual spreadsheet tables containing BoB dialogue data, using the table schema described here.

### 3.5 Basic statistics of BoB Dialogue Data files

Table 3 gives a complete overview of the basic statistics of the actual BoB Dialogue Data files we provide. Note how the languages German and Italian were introduced with some months’ delay after the English version of BoB had already gone online. The IDs in column one refer to the file names of the Excel and .csv files.

---

<sup>7</sup>These files are provided both as Microsoft Excel tables, and as plain UTF-8-encoded text files, with the first line containing the column descriptor, and the “section” symbol as field separator.

## References

- [1] N. Bertomeu, H. Uszkoreit, A. Frank, H.-U. Krieger, and B. Jörg. Contextual phenomena and thematic relations in database QA dialogues. In *Proc. of the Interactive Question Answering Workshop at HLT-NAACL 2006*, pages 1–8, New York, NY, 2006.
- [2] M. De Boni and S. Manandhar. Implementing clarification dialogues in open domain question answering. *Journal of Natural Language Engineering*, 11(4):343–361, 2005.
- [3] L. T. Dinh. Question and answer classifier for closed domain interactive question answering. Master’s thesis, Charles University Prague, 2009.
- [4] M. Kirschner. *The Structure of Real User-System Dialogues in Interactive Question Answering*. PhD thesis, Free University of Bozen-Bolzano, Italy, 2010.
- [5] S. L. Lytinen and N. Tomuro. The use of question types to match questions in faqfinder. In *Papers from the 2002 AAAI Spring Symposium on Mining Answers from Texts and Knowledge Bases*, Stanford, CA, 2002.
- [6] S. J. Mlynarczyk and S. L. Lytinen. Faqfinder question answering improvements using question/answer matching. In *Proc. of L&T-2005 - Human Language Technologies as a Challenge for Computer Science and Linguistics*, Poznan, Poland, 2005.
- [7] Z. Ratkovic. Deep analysis in iqa: evaluation on real users’ dialogues. Master’s thesis, European Masters Program in Language and Communication Technologies, 2009.
- [8] B. van Schooten, R. op den Akker, R. Rosset, O. Galibert, A. Max, and G. Illouz. Follow-up question handling in the IMIX and Ritel systems: A comparative study. *Journal of Natural Language Engineering*, 15(1):97–118, 2009.
- [9] F. Yang, J. Feng, and G. Di Fabrizio. A data driven approach to relevancy recognition for contextual question answering. In *Proc. of the Interactive Question Answering Workshop at HLT-NAACL 2006*, pages 33–40, New York City, NY, 2006.

Col.	Type	Applies to	Description
A	automatic	Q, A	<a href="#">Dialogue ID</a> . Unique ID for each user’s web session
B	automatic	Q, A	<a href="#">Date</a> . Date that web session was initialized
C	manual <sup>a</sup>	Q	<a href="#">Question relevance</a> . 0 = we might want to ignore this Q (typo, strange syntax, test/keywords, Out-of-domain, SubDialogue continuation of ‘0’); 1 = interesting Q (or SubDialogue continuation of ‘1’); 2 = semi-interesting Q, with no relevance for library domain (greeting, thanks, smalltalk, SubDialogue continuation of ‘2’)
D	manual	A	<a href="#">Answer correctness</a> . 1 = current Answer (can be ‘no pattern matched’ apology message!) is correct; 2 = corrected answer either to be found in column J, or answer has still to be added to BoBs question/answer repository
E	manual	Q	<a href="#">FU Q type</a> . 1 = FUQ; 10 = FUQ semantically breaking Subdialogue; 2 = Topic Shift Q; 20 = Topic Shift Q semantically breaking Subdialogue; 4 = Q semantically follows SubDialogue (possible only IF previous answer was marked ‘correct’); 8 = Q Re phrases previous Q; 80 = Q re phrases and semantically breaks Subdialogue
F	manual	Q	<a href="#">FU Q sub-type</a> . 1 = context-dependent FUQ; 0 = fully specified FUQ
G	automatic	Q, A	<a href="#">Language</a> . EN, DE or IT
H	automatic	Q, A	<a href="#">Status</a> . The utterance’s speaker: BoB or User
I	automatic	Q, A	<a href="#">Message</a> . The utterance proper. In case of a BoB utterance, each utterance starts with a number (01-10) in square brackets. This number encodes an “emotion” that BoB’s graphical representation adopts when issuing that utterance.
J	manual	A	<a href="#">Corrected answer</a> . Hand-corrected system answer. Provided for some rows where column D == 2.
K	automatic	Q, A	<a href="#">[dialogueId]</a> . Dialogue counter within the same user’s web session (increased only if the user hits the “reset dialogue” button)
L	automatic	Q, A	<a href="#">[uttId]</a> . Counter for utterances within a dialogue
M	automatic	A	<a href="#">[aIsLocal]</a> . ‘1’ iff the answer was retrieved from a question pattern flagged as “context-dependent FU Q” / “local” <sup>b</sup>
N	automatic	A	<a href="#">[aIsError]</a> . ‘1’ iff the answer is an apology message that BoB did not understand the previous user input <sup>b</sup>
O	automatic	A	<a href="#">[answerRerankerUsed]</a> . ‘1’ iff there were several answer candidates, and the actual answer had to be picked by the machine-learning-based answer reranker
P	automatic	Q	<a href="#">[qIsFirst]</a> . ‘1’ iff this is the first question in the dialogue
Q	automatic	Q	<a href="#">[qTechnicallyContinuesSD]</a> . ‘1’ iff this question was issued when BoB was in sub-dialogue mode, and the question <b>continues the sub-dialogue</b> by following one of the proposed sub-dialogue paths
R	automatic	Q	<a href="#">[qTechnicallyBreaksSD]</a> . ‘1’ iff this question was issued when BoB was in sub-dialogue mode, and the question <b>breaks the sub-dialogue</b> by NOT following any of the proposed sub-dialogue paths

<sup>a</sup> Note that all fields of type “manual” are currently far from complete, nor consistent.

<sup>b</sup> Cf. Section 3.2 for a description of dialogue-related meta-information

Table 2: Column descriptions of BoB Dialogue Data

ID	Start date	End date	Dialogues (all lang.)	User messages			Hand annotations (only partial!):				
				EN	DE	IT	Q relevance (Col C)	FU Q class (Cols E,F)	A correct. (Col D)		
bob001	2008-09-29	2008-10-02	296	605	0	0	-	-	-	-	-
bob002	2008-10-02	2008-10-10	615	1234	0	0	yes	yes	yes	yes	yes
bob003	2008-10-10	2008-10-21	296	382	0	0	yes	yes	yes	yes	yes
bob004	2008-10-21	2008-11-17	440	615	0	0	-	-	-	-	-
bob005	2008-11-17	2009-01-09	371	349	0	0	-	-	-	-	-
bob006	2009-01-28	2009-03-01	385	364	509	0	yes	-	-	yes	yes
bob007	2009-03-01	2009-03-04	129	92	255	0	yes	-	-	yes	yes
bob008 <sup>a</sup>	2009-03-16	2009-06-15	847	1287	785	0	-	-	-	-	-
bob009 <sup>a</sup>	2009-06-16	2009-09-24	647	629	653	0	yes	-	-	yes	yes
bob010 <sup>a</sup>	2009-09-24	2009-10-15	422	109	319	725	yes	-	-	yes	yes
bob011	2009-12-03	2010-01-15	267	197	303	189	-	-	-	-	-
bob012	2010-01-15	2010-02-01	155	112	168	92	-	-	-	-	-
bob013	2010-02-01	2010-06-07	532	398	530	318	-	-	-	-	-
<i>Sum</i>			5402	6373	3522	1324					

<sup>a</sup> Problem in file: special characters (German umlauts, Italian accents) coded as ‘?’

Table 3: Basic statistics of BoB Dialogue Data files