ANTswers: an interactive library FAQ

Danielle Kane, Research Librarian for Emerging Technologies and Service Innovation, UCI Libraries

BACKGROUND

ANTswers is an AIML chat robot (chatbot) built to answer questions about the UCI Libraries but to also respond to general conversation. ANTswers is available from the library homepage (http://www.lib.uci.edu/) all conversations are anonymous, and no login is required.

In March of 2014 for beta testing. ANTswers is built for an academic library with 2 main libraries, a study center, Special Collections & Archives, and a medical library.

Our goal in creating a library chatbot was:
• Enhance reference services
• Serve a large number of people at any one-time of the day
• Serve a wide range of library patrons
• Use it to remind staff patrons
• Provide UCI specific information

IMPLEMENTATION

ANTswers was in development for a year prior to going live. A focus group of 13 members of library staff was created to discuss the personality, look and feel, and scope of the chatbot. The majority of the AIML programming was completed by 1 staff member (botmaster) with support provided by a programmer in the Libraries’ IT department.

- 9/23/13 – 12/9/13
  • Focus: Focus group determines scope and personality
- 12/9/13 – 1/28/14
  • Focus: Development of AIML files and testing
- 1/28/14 – 3/28/14
  • Transcripts reviewed and revision to database made by botmaster.
- 3/25/14 – current
  • ANTswers goes live

BASICS OF AIML

AIML stands for Artificial Intelligence Markup Language and is an XML compliant language. Each question / answer pair is considered a unit of knowledge and is represented as:

<category><pattern>CONTACT YOU</pattern></category>

Where *category* defines the unit of knowledge, *pattern* defines the keyword, phrase, or question, and <template> defines the chatbots response. Recursion can be used to map secondary categories to the primary category.

<category><pattern>CONTACT YOU *</pattern></category>

<template><srai>CONTACT YOU</srai></template></category>

EVALUATION

While Program-O contains transcript logs they are just viewable. Initially in the testing phases a Google form was used to track statistics on each conversation but the connection between the statistics and the conversation was lost. Moving forward the best solution was a statistics database built using MySQL, which pulls transcripts from Program-O and allow for logging, reviewing (statistics), and forwarding transcripts to other Library staff.

Each transcript is reviewed and statistics are submitted, statistics are kept on the following criteria:
• General Information:
  • Date | Hour | Quarter | Week | Day of Week
• Determining Answer Percentage:
  • # of Questions
  • # of Library questions – asked | answered | answer rate
  • # of General questions – asked | answered | answer rate
  • What is being asked:
  • Questions type – books | journals | articles | etc.
  • Statistics are downloaded weekly onto an Access Database which is linked to an Excel spreadsheet. At the beginning of each quarter a spreadsheet is created using pivot tables and by refreshing the data weekly statistics can be quickly and easily shared.

Transcripts are reviewed on a daily basis (M-F) in our Lib Web Admin System, transcripts are exported and run through a text analyzer on a quarterly and yearly basis. Information on how library users are asking questions are used to refine the ANTswers database but is also of interest as the UCI Libraries works to redesign our library website.

MANTENANCE

Transcripts are reviewed on a daily basis (M-F) in our Lib Web Admin System, chatbots responses are evaluated and changes are made to the AIML code if needed.

BUILD YOUR OWN CHATBOT

The collection includes approximately 130 AIML files that are available for download and for the use in creating other library chatbots. You will need to review and edit each of the files for your particular institution. To implement your own library chatbot you will need a server, chatbot software (we used Program-O), editing software (we used Notepad++) and at least minimal experience with HTML, CSS, JavaScript and AIML.

http://escholarship.org/uc/uci_libs_antswers

WANT TO KNOW MORE?

http://www.online-utility.org/text/analyzer.jsp

Chatlog transcripts are exported and run through a text analyzer on a quarterly and yearly basis. Information on how library users are asking questions are used to refine the ANTswers database but is also of interest as the UCI Libraries works to redesign our library website.

Number of characters (including spaces) : 120547
Number of characters (without spaces) : 77343
Number of words : 14048

RESULTS

Data is from 3/25/14 – 6/9/15
• Total number of conversations: 715
• Total number of questions: 3,486
• Total number of Library questions: 1,918 (55% of total)
• Total number of General questions: 1,554 (45% of total)

Some top phrases containing 4 words (without punctuation marks)

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Occurrences</th>
<th>Phrase</th>
<th>Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>How do I</td>
<td>200</td>
<td>What time is it</td>
<td>21</td>
</tr>
<tr>
<td>Where is it</td>
<td>161</td>
<td>I need to</td>
<td>67</td>
</tr>
<tr>
<td>What is the</td>
<td>67</td>
<td>I want articles on</td>
<td>27</td>
</tr>
<tr>
<td>Where can I</td>
<td>82</td>
<td>How can I</td>
<td>48</td>
</tr>
<tr>
<td>What is your</td>
<td>51</td>
<td>Do I get</td>
<td>23</td>
</tr>
<tr>
<td>Who is the</td>
<td>27</td>
<td>How do I</td>
<td>49</td>
</tr>
<tr>
<td>How do</td>
<td>23</td>
<td>How can</td>
<td>36</td>
</tr>
<tr>
<td>Do you have</td>
<td>32</td>
<td>How do I</td>
<td>49</td>
</tr>
<tr>
<td>I need to</td>
<td>67</td>
<td>What is the</td>
<td>67</td>
</tr>
<tr>
<td>What is your name</td>
<td>21</td>
<td>I want articles on</td>
<td>27</td>
</tr>
<tr>
<td>What is the library</td>
<td>36</td>
<td>What do you need</td>
<td>67</td>
</tr>
</tbody>
</table>

Some top phrases containing 3 words (without punctuation marks)

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Occurrences</th>
<th>Phrase</th>
<th>Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is your name</td>
<td>21</td>
<td>I want articles on</td>
<td>27</td>
</tr>
<tr>
<td>What is the library</td>
<td>36</td>
<td>What do you need</td>
<td>67</td>
</tr>
<tr>
<td>What is your name</td>
<td>21</td>
<td>I want articles on</td>
<td>27</td>
</tr>
<tr>
<td>What is the library</td>
<td>36</td>
<td>What do you need</td>
<td>67</td>
</tr>
<tr>
<td>What is your name</td>
<td>21</td>
<td>I want articles on</td>
<td>27</td>
</tr>
<tr>
<td>What is the library</td>
<td>36</td>
<td>What do you need</td>
<td>67</td>
</tr>
<tr>
<td>What is your name</td>
<td>21</td>
<td>I want articles on</td>
<td>27</td>
</tr>
<tr>
<td>What is the library</td>
<td>36</td>
<td>What do you need</td>
<td>67</td>
</tr>
</tbody>
</table>

FUTURE PLANS

• Redesign of ANTswers’ library related files due to a changing library web presence.
• Create a refined database using knowledge gained from the transcript logs – goal is to eventually reach a 90% answer rate for library questions. ANTswers will then move out of beta.
• A research project is currently underway to compare how questions are asked in in-person reference, QuestionPoint (24/7), and ANTswers.