ANTswers: an interactive library FAQ

Daniele 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 built for an academic library with 2 main libraries, a study center, Special Collections & Archives, and a medical library.

Our goal is to create a library chatbot with:
- Enhance reference and 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 simulate a human patron
- Provide UCI specific information

ANTswers was in development for a year prior to going live. A focus group of 13 downloaded to one of our library maintained servers.

AIML stands for Artificial Intelligence Markup Language and is an XML compliant following criteria:

- General Information:
  - Data | Hour | Quarter | Week | Day of Week
- Determining Answer Percentage:
  - Topic of Questions
  - # of Library questions – asked | answered | answer rate
  - # of general questions – asked | answered | answer rate
- What is Being Asked:
  - Questions like – books | leads | articles | etc.

Statistics are downloaded weekly into an Access Database which is linked to an Excel spreadsheet. At the beginning of each quarter a spreadsheet is created using pivot charts and by refreshing the data weekly statistics can be quickly and easily shared.

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 allows for tagging, 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:
  - Data | Hour | Quarter | Week | Day of Week
- Determining Answer Percentage:
  - Topic of Questions
  - # of Library questions – asked | answered | answer rate
  - # of general questions – asked | answered | answer rate
- What is Being Asked:
  - Questions like – books | leads | articles | etc.

Statistics are downloaded weekly into an Access Database which is linked to an Excel spreadsheet. At the beginning of each quarter a spreadsheet is created using pivot charts and by refreshing the data weekly statistics can be quickly and easily shared.

IMPLEMNENTATION

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.

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>QUESTION</pattern><template>ANSWER</template></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.

TRANSCRIPTS

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

MAINTENANCE

Ants spray and chew their way through the wood. They eat the soft inner part of the wood. The wood may be weakened and this can lead to a decrease in strength.

Revises the AIML for Borrowing & Fines we can find the category for RENEW. We can then add a new category for the phrase, "How soon can I renew?" Next time this question is asked the library patron will receive the correct response.

The file will then be saved, uploaded, and the older versions will be deleted from Program-O (http://blog.program-o.com/) which is written down to one of our library maintained servers.

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)

LANGUAGE ANALYSIS

http://www.online-utility.org/text/analyse.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 is used to refine the ANTswers database but is also of interest as the UCI Libraries works to redesign our library website.

FUTURE PLANS

- Redesign of ANTswers’ library related files due to a changing library web presence.
- Generate a refined database using knowledge gained from the previous transcripts – goal is to eventually reach a 95% answer rate for library questions. ANTswers will then move out ofbeta.
- A research project is currently underway to compare how questions are asked in an in-person reference, QuestionPoint (24/7), and ANTswers.

BUILD YOUR OWN CHATBOT

This 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

http://lib.uci.edu/ALAPoster

Want to Know More?

http://ifs.uci.edu/ANTanswers