

# Final Project Progress Report

Ryan Baker (captain; NetID: rtb2) -- [rtb2@illinois.edu](mailto:rtb2@illinois.edu)

Which tasks have been completed? Which tasks are pending?

For convenience in reference, the original proposal includes the following tasks:

Task	Expected Time
Scrape the Blizzard Hearthstone API and build a simple DB of the cards	COMPLETE (4 hours)
Build an inverted index with information from the card titles and text from the card body	5 hours
Build an API (and a UI if time permits) through which a user can submit their current cards, enter search keywords and ask for recommendations	5-10 hours
Build an evaluator with known existing decks	5 hours

The first task to scrape the Blizzard Hearthstone API and build a simple DB of the cards has been completed. As part of this, I created two different python scripts, both checked into the Github project repository. One script handles the download of cards from the Blizzard Hearthstone API. This required signing up for the developer API and creating an OAuth client to handle requests with the Blizzard API. The cards can then be queried and are returned as paginated results, which must be collected with multiple queries to the Blizzard API. The second script utilizes the first script to gather a list of cards. It then creates a SQLite DB, with a simple table for the cards that includes all text based fields from the card details in their raw form. It populates this DB with all of the cards returned from the Blizzard Hearthstone API.

The remaining steps are yet to be completed.

What challenges are being faced?

There are currently no significant challenges being faced. Some of the work from the first step did reveal some extra details that will need to be handled. The next step will be to create an inverted index through which a user could search through the cards based on keywords. Based on preliminary investigation into the text fields, and documentation from the Blizzard Hearthstone API, I will need to combine several different text based fields associated with each card into a single "document" per card. Then, I will also need to do some pre-processing of each card to remove some markdown and HTML based tagging that adds cruft to the card texts.