

# A Web-Based Platform for Java Specifications

---

Arik Coats

Evan Dye

Robert Kloster

# Specifications

## What are specifications?

- Formal contract between the software (provider) and the user (client)
- Defines code behavior: preconditions, postconditions, invariants, etc.
- Similar to formatted pseudo code documentation

## Benefits

- Avoid more errors while coding
- Faster coding and debugging.
- Reliable, testable code.

# Problem Statement

Despite their usefulness, there is a severe lack of formal specifications for common software libraries. There are several reasons for this:

- In order to specify your software, all of the code the software relies on must be specified as well.
- Formal specifications are very difficult to write and require uncommon knowledge.
- Writing formal specifications is time consuming, so many companies do not consider them worthwhile or cost effective.

# Purpose of project

The implementation and development of a Q/A web forum dedicated to the creation, discussion, and refinement of formal behavioral specifications for commonly used Java libraries.

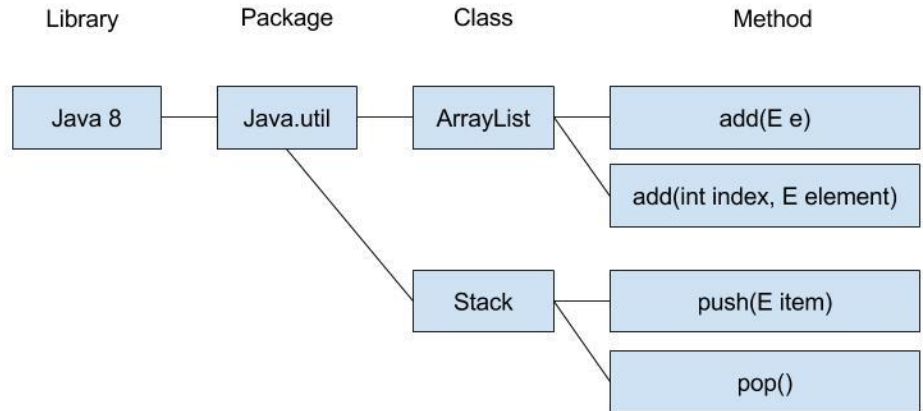
# Solution Plan for CprE/SE 491

Construct a forum to view, discuss, create, and approve specifications, and meets the following requirements:

1. The forum will contain the source code for classes and methods of commonly used Java libraries for easy reference.
2. The forum will provide an API for easy access of data by 3<sup>rd</sup> party systems.
3. The forum will offer users meaningful, related posts from Stack Overflow.

# MyBB Forum Structure

- Forum tree structure for supported libraries.
- For all class and method subforums display:
  - Discussion threads.
  - Source code
  - Related discussions



# Java Libraries: Goals

The final version of the website should have a database and forum trees corresponding to all supported common Java libraries.

- Develop system to extract source code information and make it available for use by the forum.
- Allow for more meaningful queries of source code.
- Develop system to automatically generate forum trees from the extracted libraries.

# Java Libraries: Development and Challenges

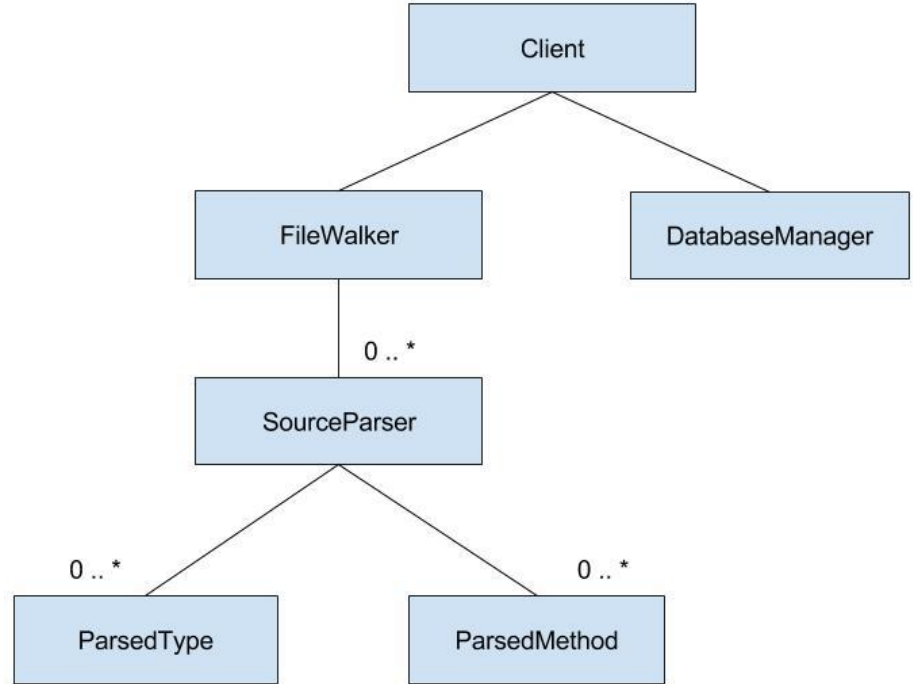
Researched many possible solutions to extract source code information.

- Unused solution: GrepCode API
  - Limited API can not extract sufficient data.
  - Cannot be used in an automated system.
- Unused solution: Java Reflection API
  - Limited available data for extraction.
  - Tedious manual parsing.
  - Difficult to implement for multiple java files.
- Chosen solution: use third party parser
  - Eclipse JDT, Abstract Syntax Tree
  - JDBC, MySQL



# ExtractSource: Diagram

- **Client**
  - Runs FileWalker.
  - Sends results to DatabaseManager.
- **FileWalker**
  - Collects Library and Package data
  - Searches directory for Java files
  - Creates a SourceParser for each Java file found.
- **SourceParser**
  - Extracts and parses Java source code using Eclipse JDT.
  - Creates ParsedType and ParsedMethod objects for all classes and methods, respectively.
- **DatabaseManager**
  - Creates and populates local database with extracted source code data using JDBC and MySQL.



# Java Libraries: Progress

Complete:

- Successfully extracted source information of multiple Java libraries and stored it in a local MySQL database.
- Tested on the libraries Java 8 and Apache Commons Collections 4.

Yet to be done:

- Import extracted data to the database on the project server.
- Develop automated system to generate forum trees from the extracted libraries.

# API Implementation: Goals

- Final version of the website must be able to provide users with data regarding
  - questions,
  - answers,
  - tags,
  - users,
  - etc.as JSON objects via RESTful API calls.
- Format and functionality should be highly similar to Stack Exchange's API system.

# API Implementation: Solutions

- Assess primary use cases of API system, plan API calls
- Implement custom API system within MyBB
- Focus primarily on query calls for data; posts, users, etc.

# API Implementation: Progress

Created Custom API system to return data for

- answers      /answers, /answers/{ids}
- questions    /questions, /questions/{ids}, /questions/{ids}/answers
- users        /users,    /users/{ids}

## For CprE/SE 492...

- Re-evaluate API system requirements, adjust accordingly
- Coordinate with team to restructure MyBB's database
- Change MyBB source code to properly edit new additions to the database

# Related Posts: Goal

For each class/method:

- Generate a list of links to related Stack Overflow posts.

Purpose:

- Provide resources to research the class/method under discussion.

The screenshot shows the May1639 Forums website. The main content area displays a forum thread titled "add(E e)" under the path "Forums > Libraries > Sample Data Structure Library > java.util > ArrayList". The thread is by user "endye" and has 1 reply and 22 views. Below the thread information is a "Source Code" section. To the right, a "Related Discussions" sidebar lists several links to other forum posts, including "how to add an object to arraylist in java", "How to use ArrayList.addAll()", "How to add listener on ArrayList in java", "Java Add Object to ArrayList error", "How do I synchronize access to an array list accessed by multiple threads?", "Type mismatch: cannot convert from boolean to int java", "How does using add() to one ArrayList interfere the size of the other one?", "How do I add a new Object to an ArrayList?", "While and for loops for adding to two dimensional array", and "(Java) Internal work of add(obj o) method in ArrayList<obj>". A red line connects the "add(E e)" text in the forum title to the "Related Discussions" sidebar.

May1639 Forums

Hello There, Guest! Login Register

Portal Search Member List Calendar Help

Forums > Libraries > Sample Data Structure Library > java.util > ArrayList

add(E e)

Users browsing this forum: 1 Guest(s)

Mark this forum read | Subscribe to this forum

Thread / Author	Replies	Views	Rating	Last Post [asc]
Specification Proposal endye	1	22	★☆☆☆☆	12-06-2015, 09:32 PM Last Post: adcoats

Sort by: Last Post Order: Descending From: The Beginning Go

Source Code

New Posts No New Posts  
Hot Thread (New) \* Contains Posts by You  
Hot Thread (No New) Locked Thread

Search this Forum: Go

Forum Jump: ----- add(E e) Go

Contact Us May1639 Home Page Return to Top Lite (Archive) Mode RSS Syndication

Powered By MyBB, © 2002 - 2015 MyBB Group.

Current time: 12-08-2015, 03:56 PM

# Related Posts: Stack Exchange API

## Step 1: Query Stack Overflow

```
jQuery.getJSON( "http://api.stackexchange.com/2.2/search/  
  advanced?order=asc&sort=relevance&accepted=True&  
  tagged=add;arraylist&site=stackoverflow" );
```

Returns JSON:

Repeat while the "has\_more" field is true.

## Step 2: Rank Results (Current Method)

Consider:

- R - Relevant Tags (i.e. Java, Post-condition)
- N - Number of Views
- S - Score/Rating

$$\begin{aligned}\text{Rank} &= C_1 * \text{Normalize}(R) \\ &+ C_2 * \text{Normalize}(N) \\ &+ C_3 * \text{Normalize}(S)\end{aligned}$$

Where  $C_i$  is the weight of the term. Currently, all weights are 1.

```
{  
  "items": [30]  
  0:  
  {  
    "tags": [4]  
      0: "java"  
      1: "arraylist"  
      2: "add"  
      3: "character"  
    "owner":  
    {  
      "reputation": 313  
      "user_id": 455946  
      "user_type": "registered"  
      "accept_rate": 61  
      "profile_image": "https://www.gravatar.com/avatar/....  
      "display_name": "masb"  
      "link": "http://stackoverflow.com/users/455946/masb"  
    }  
    "is_answered": true  
    "view_count": 51965  
    "accepted_answer_id": 8985476  
    "answer_count": 6  
    "score": 11  
    "last_activity_date": 1386344446  
    "creation_date": 1327401846  
    "last_edit_date": 1327402019  
    "question_id": 8985432  
    "link": "http://stackoverflow.com/questions/8985432/how-to-use-arraylist-addall"  
    "title": "How to use ArrayList.addAll()?"  
  }  
  ...  
  "has_more": true  
  "quota_max": 300  
  "quota_remaining": 290  
}
```

# Related Posts: Stack Overflow Data Dump

Limitations of StackExchange API:

- 300 queries per day limit per IP (10,000 with Key).
- 30 requests per second limit per IP.
- Expensive to parse all interesting content in real time.

New Solution: Stack Overflow Data Dump

- The Stack Exchange Network openly provides quarterly data dumps for all sites.
- Provides all posts, users, and other data in easily parsable XML format.
- Related Posts will be ranked using the previously outlined method.
- Related posts may be ranked in advance and stored in the database.

This improved solution will be implemented



# Looking Ahead

We have a solid foundation for moving forward in CprE/SE 492. Here are our plans for the next semester.

- Complete and/or expand on currently implemented systems.
- Implement new features
  - Proposing and approving new specifications.
  - Record more varieties of metadata to better implement an API consistent with Stack Exchange.
- Merge project with team May1620.

# Questions?

# Resources

1. Rajan, Hridayesh, Tien N. Nguyen, Gary T. Leavens, and Robert Dyer. "Inferring Behavioral Specifications from Large-scale Repositories by Leveraging Collective Intelligence." "37th International Conference on Software Engineering: NIER Track" May 2015, ICSE'15, Florence, Italy.