

The Design and Implementation of iCodeSearch

Tong Liu

Nanjing University of Aeronautics and Astronautics

0. Contents

Background and motivation

Features of iCodeSearch

System design

Implementation

1. Background and Motivation

Why code search?

Promote code reuse

Developers can quickly build apps using open source code

Find solutions by example

Most code examples (snippets) provide solutions to specific problems

I. Background and Motivation

Existing code search tools

• **Google Codesearch, Codase, Koders, etc.**

• **Limitations**

- * Hard to find useful examples within numerous result items
- * Lack of interaction with user, not customizable

2. Features of iCodeSearch

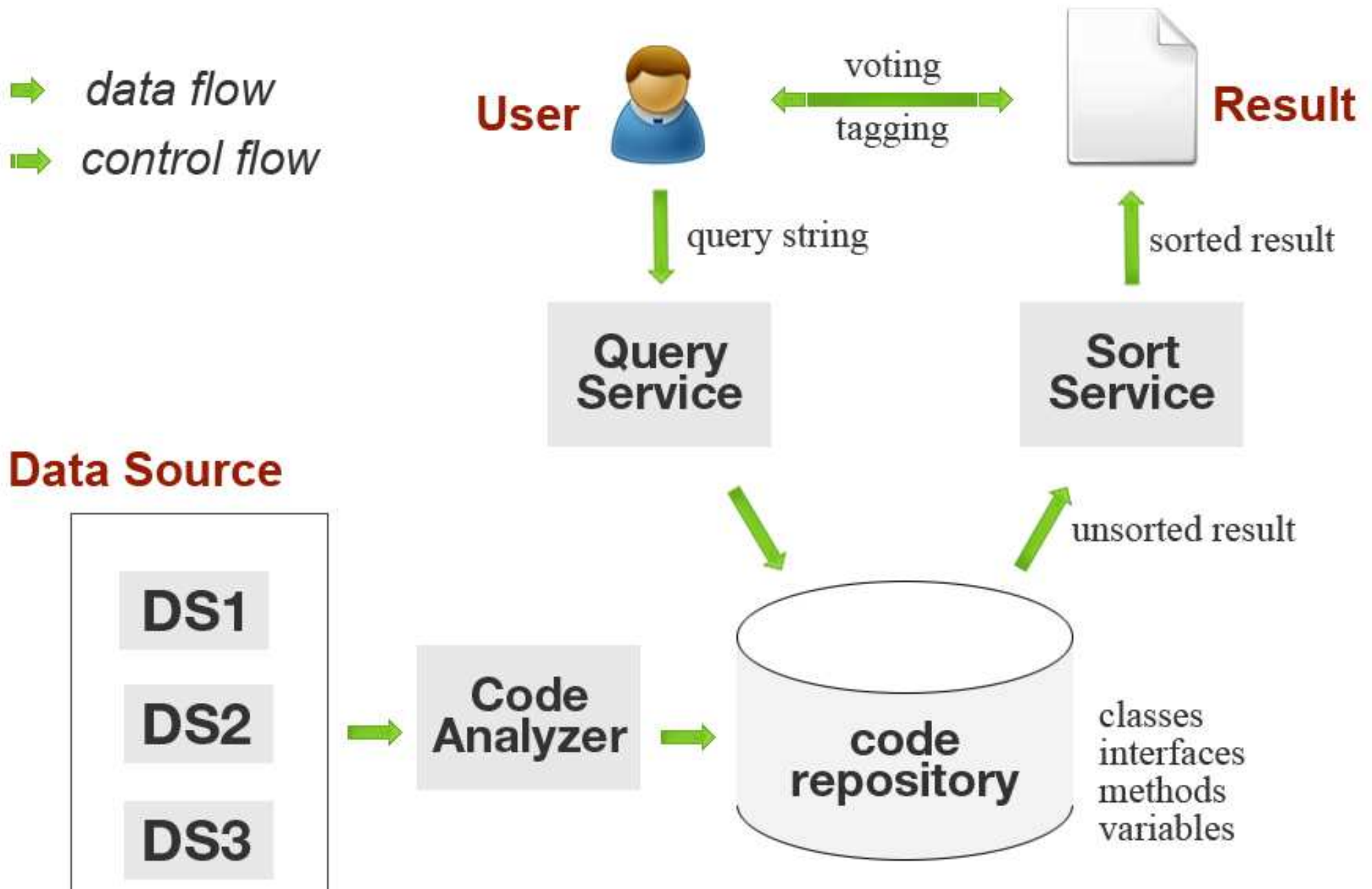
Multiple data source

Metric-based result sorting

Social features such as voting and tagging

System Design Overview

➔ *data flow*
➔ *control flow*



3. System design

Multiple data source



files from web crawler



user submitted examples



other search engines

3. System design

Example extraction

What makes a good code example?

- * Well-documented, i.e. commented
- * Contains every keyword in the query string
- * Neither too long, nor too short

3. System design

Interactive search result

A web 2.0 style result page

- * Users can vote for a result item, like they do at Digg.com
- * Result items can be tagged

3. System design

Sorting based on votes / tags

Let users decide whether a search result deserves high priority

Add semantic tags to search results so that they can be browsed by category

Results with higher votes / more tags appear on top

4. Implementation

Data source module

- **Make search calls to Google Codesearch and/or Koders, then fetch the top results**
- **Use JoBo to crawl source files from open source repositories such as sourceforge.net**
- **Accept user-submitted code snippets**

4. Implementation

Code analyzer module

JDT (Java Development Tools) is used to navigate a java source file and extract information of code entity such as classes, methods, and variables.

4. Implementation

Finding code examples

- Assign a float value (score) to each result item
- The value is a weighted add of several metrics:
 - * *Comment ratio*
 - * *Code length*
 - * *Keywords cohesion*
- Result items with high score are treated as “code examples” and will appear prior to normal results

4. Implementation

Voting

Code Search

Search

You searched for "ActionSet" which produced 14 result(s).

CodeSearch result

[t\jedit\ActionSet.java](#)

class

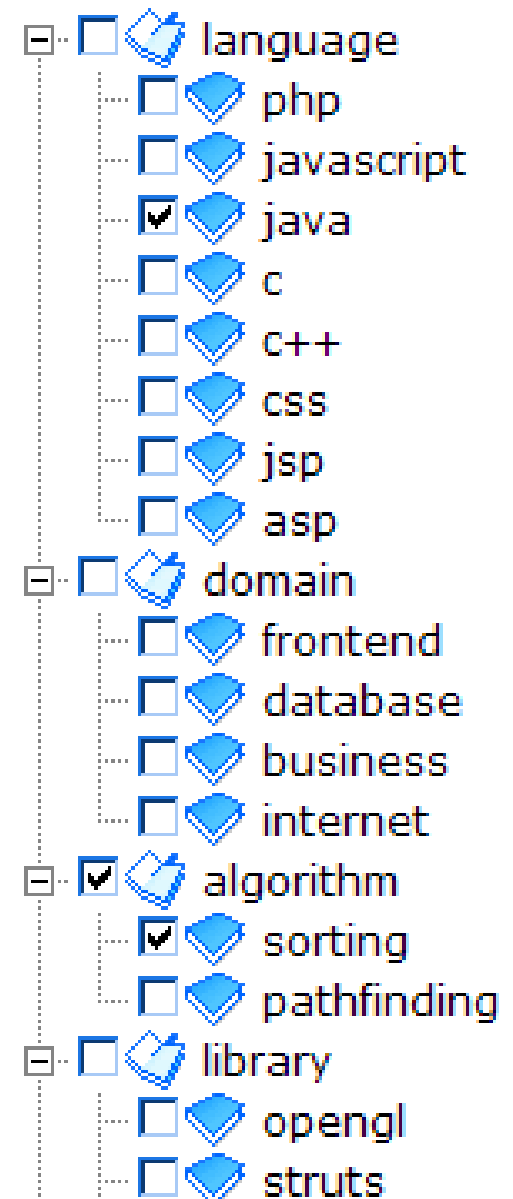
```
/**
 * A set of actions, either loaded from an XML file, or constructed at runtime
 * by a plugin.
 *
 * Action sets loaded from XML files
 *
 * Action sets are read from these files inside the plugin JAR:
 *
```

4. Implementation

Tagging

Select tags from tag tree

Sub-categorized



4. Implementation

Home page

Login

Reg

iCode Search

[search codes](#) - [search examples](#)

Search

Google

Koders

Examples

Search for a single entity

Search for multiple entities

Optionally enable results from Google Codesearch and Koders

"Action", "array", "move"

"move array", "map index"

Links

» [Submit code snippet](#)

End

Thanks!