**BÁO CÁO BÀI TẬP 5**

**WEBSITE TÌM KIẾM - BINGLE**

# Thông Tin Nhóm:

|  |  |  |
| --- | --- | --- |
| **Mssv** | **Họ tên** | **Email** |
| 0812515 | **Phan Nhật Tiến** | [phntien@gmail.com](mailto:phntien@gmail.com) |
| 0812527 | **Huỳnh Công Toàn** | [7hanatos13@gmail.com](mailto:7hanatos13@gmail.com) |
| 0812642 | **Nguyễn Đức Xuân** | [abcxyz2357@gmail.com](mailto:abcxyz2357@gmail.com) |

**GVLT: TS. Nguyễn Văn Vũ**

**Lớp:** Các chủ đề nâng cao trong CNPM TH2008/3

# Môi trường thực thi:

* Tool: Eclipse Indigo (để chạy AspectJ).
* Web server: **Apache Tomcat 7.0.**
* Language: **Java.**
* Công nghệ: **JSP** + **JSTL** + **JSON** + **Search API (Google + Bing)** + **AspectJ + Bing Translate**.
* Yêu cầu khác: JDK 6.0 (để chạy AspectJ), có kết nối internet.

# Giới Thiệu:

## Bingle đáp ứng được các yêu cầu chức năng:

* Tìm kiếm dựa trên từ khóa của người dùng nhập vào.
* Người dùng sẽ nhìn thấy danh sách các kết quả dựa vào câu truy vấn.
* Người dùng sẽ nhìn thấy chi tiết khi click vào link kết quả.
* Mỗi kết quả bao gồm: Tiêu đề, link, mô tả ngắn, nguồn.

## Bên cạnh đó còn đáp ứng được yêu cầu phi chức năng:

* Sử dụng Aspects để log tất cả các exceptions và từ khóa tìm kiếm. File log được lưu ở D:/DangHoc/Bingle.log.
* Cho kết quả chính xác nhất có thể. Làm tăng hiệu năng của trang web.
* Giao diện thân thiện, dễ dùng, dễ thao tác. Các nội dung của mỗi kết quả được phân biệt với nhau rõ ràng.
* Kết quả tìm kiếm sẽ phân bố từ trên xuống như sau: các kết quả chung của Google và Bing, các kết quả riêng của Google, cuối cùng là các kết quả riêng của Bing.
* Có phân trang kết quả.
* Loại bỏ quảng cáo không cần thiết.
* Loại bỏ kết quả trùng nhau.

## Đáp ứng yêu cầu tìm kiếm bằng ngôn ngữ tiếng Việt.

* Có thể tìm kiếm bằng từ khóa không dấu lẫn có dấu.
* Các tiêu đề của tất cả các kết quả được chuyển ngữ sang tiếng việt.

# Giao Diện:

## Giao diện lần đầu:



## Giao diện khi xem About



## Giao diện khi tìm kiếm:



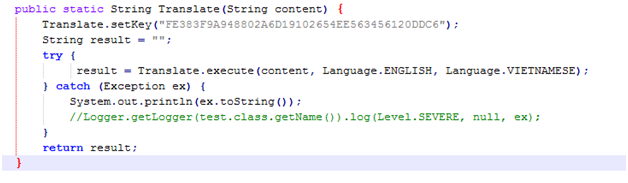
# Diễn giải:

## Thuật toán tìm kiếm:

Sau khi nhận được từ khóa, chương trình sẽ mã hóa (encode) từ khóa. Sau đó nhận kết quả dạng JSON từ Google và Bing sau khi gửi request dạng RESTful. Sau khi sắp xếp các kết quả, chương trình sẽ chuyển ngữ toàn bộ tiêu đề dùng Bing Translate API

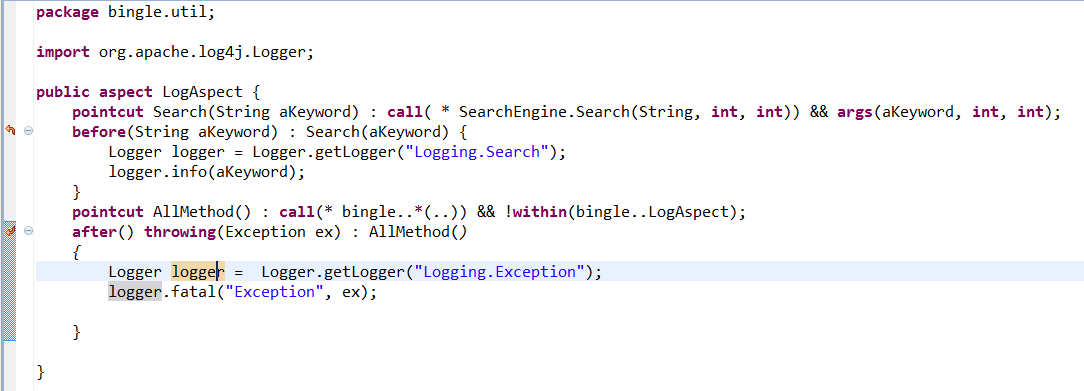
## Chuyển ngữ:

Việc chuyển ngữ sẽ áp dụng thư viện mở của Bing Translate.



## Logging

Chương trình sẽ log lại tất cả các từ khóa tìm kiếm của người dùng và các exception trong chương trình, sử dụng AspectJ và log4j. Có thể cấu hình log4j trong file log4j.properties.



# Ghi chú

* Số kết quả tối đa mỗi lần request của Bing là 50.
* Số kết quả tối đa mỗi lần request của Google là 10.
* Số kết quả tối đa mỗi từ khóa của Google là 100.

# Khuyết điểm

* Do hạn chế về số lượng kết quả của Google (mỗi từ khóa chỉ có tối đa 100 kết quả) nên trang web chỉ hiển thị 10 trang kết quả. Mỗi trang từ 10 đến 20 kết quả từ cả Google và Bing.
* Trang web không có chức năng tìm kiếm nâng cao, chức năng nhắc từ khóa.
* Do hạn chế của Bing translate nên một số từ dịch không chính xác với ngữ cảnh.

# Tài liệu tham khảo:

## Về Bing:

* <http://www.bing.com/developers/s/APIBasics.html>
* <http://code.google.com/p/bing-search-java-sdk/>
* <http://msdn.microsoft.com/en-us/library/dd250846.aspx>

## Về Google:

* <http://code.google.com/apis/customsearch/v1/performance.html>

## Về JSON

* <http://www.json.org/java/>
* <http://answers.oreilly.com/topic/257-how-to-parse-json-in-java/>
* <http://mkyong.com/tutorials/java-json-tutorials/>
* <http://viralpatel.net/blogs/2009/02/creating-parsing-json-data-with-java-servlet-struts-jsp-json.html>

## Về AOP

* <http://www.andrewewhite.net/wordpress/2010/03/17/aspectj-annotation-tutorial/>
* <http://eclipse.org/aspectj/doc/released/faq.php#q:whatisaj>
* <http://plugins.netbeans.org/plugin/4015/aspectj-ajde-for-netbeans>
* <http://blog.espenberntsen.net/2010/03/18/aspectj-examples-with-pointcuts-based-on-annotations/>

## Log4j

* <http://logging.apache.org/log4j/1.2/manual.html>

## Về Bing translate

* <http://code.google.com/p/microsoft-translator-java-api/>