**NHÓM 01**

**ĐỒ ÁN:**

**MẠNG XÃ HỘI CHIA SẺ ĐỊA ĐIỂM THÀNH PHỐ HỒ CHÍ MINH**

**HIGH LEVEL ARCHITEXTURE**

**Bảng ghi nhận thay đổi tài liệu**

|  |  |  |  |
| --- | --- | --- | --- |
| **Ngày** | **Phiên bản** | **Mô tả** | **Tác giả** |
| 09/10/2010 | 1.0 | Khởi tạo bản High Level Architecture | Cả nhóm |
|  |  |  |  |
|  |  |  |  |

**MỤC LỤC**

[1. Giới thiệu 3](#_Toc464003322)

[**1.1** **Mục đích tài liệu** 3](#_Toc464003323)

[**1.2** **Đối tượng của tài liệu** 3](#_Toc464003324)

[**1.3** **Tác giả** 3](#_Toc464003325)

[**1.4** **Thuật ngữ sử dụng** 3](#_Toc464003326)

[**1.5** **Cấu trúc tài liệu** 3](#_Toc464003327)

[2. Sơ đồ CSDL 4](#_Toc464003328)

[3. Mô tả CSDL 4](#_Toc464003329)

[**3.1** **USER** 4](#_Toc464003330)

[**3.2** **DIADIEM** 5](#_Toc464003331)

[**3.3** **HINHANH** 5](#_Toc464003332)

[**3.4** **COMMENT** 6](#_Toc464003333)

[**3.5** **RATING** 6](#_Toc464003334)

[4. Phân hoạch dữ liệu 6](#_Toc464003335)

[**4.1** **Dữ liệu thô** 6](#_Toc464003336)

[**4.2** **Phân hoạch dữ liệu thô** 7](#_Toc464003337)

[**4.3** **Tạo ra các bảng dữ liệu từ dữ liệu đã phân hoạch được** 8](#_Toc464003338)

[5. Chứng minh tính khả thi 8](#_Toc464003339)

[**5.1** **Sơ đồ quá trình truy xuất dữ liệu** 9](#_Toc464003340)

[**5.2** **Giảng giải sơ đồ** 9](#_Toc464003341)

[6. Kết luận 9](#_Toc464003342)

1. **Giới thiệu**
   1. **Mục đích tài liệu**

Tài liệu High Level Architexture bao gồm mô hình quan hệ chứa các bảng trong CSDL và mối liên hệ giữa các bảng với nhau để xây dựng nên trang web tìm kiếm địa điểm có tích hợp mạng xã hội.

* 1. **Đối tượng của tài liệu**
* Khách hàng: Tất cả người sử dụng dịch vụ của trang web tìm kiếm này
* Các thành viên trong nhóm
  1. **Tác giả**

Tác giả của tài liệu này thuộc về nhóm 1 bộ môn QLDAPM/2013 do Thầy Ngô Huy Biên phụ trách.

* 1. **Thuật ngữ sử dụng**

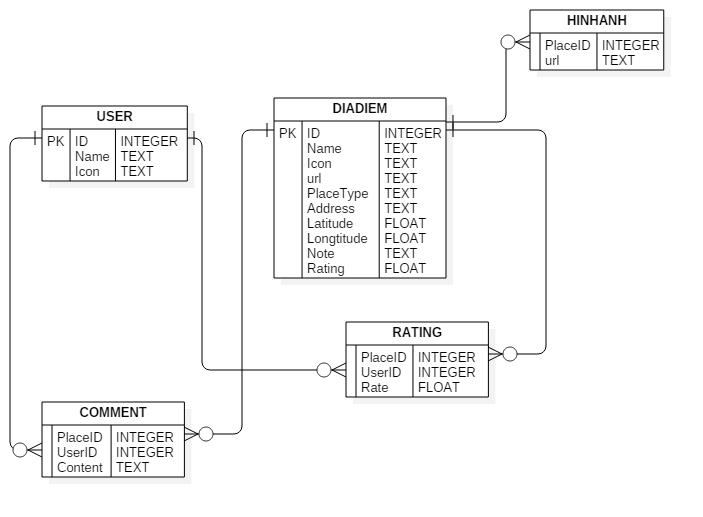
|  |  |
| --- | --- |
| Thuật ngữ | Ý nghĩa |
| QLDAPM | Quản lí dự án phần mềm |
| CSDL | Cơ sở dữ liệu |
|  |  |

* 1. **Cấu trúc tài liệu**

Tài liệu bao gồm 4 phần:

* Sơ đồ CSDL
* Mô tả chi tiết về các bảng trong CSDL
* Phân hoạch dữ liệu: Chuyển đổi dữ liệu từ trang vnfinder.com thành cơ sở dữ liệu của mình.
* Chứng minh tính khả thi: Trình bày hoạt động của sơ đồ cơ sở dữ liệu để đạt hiệu quả cao trong việc tìm kiếm dữ liệu cung như lưu trữ CSDL

1. **Sơ đồ CSDL**



1. **Mô tả CSDL**
   1. **USER**

* Ý nghĩa: Chứa thông tin về một người dùng.

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ngữ nghĩa |
| 1 | ID | Integer | Khoá của user |
| 2 | Name | Text | Tên của user |
| 3 | Icon | Text | Avartar của user đó |

* 1. **DIADIEM**
* Ý nghĩa: Chứa thông tin về những địa điểm

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ngữ nghĩa |
| 1 | ID | Integer | Khoá của địa điểm |
| 2 | Name | Text | Tên của địa điểm |
| 3 | Icon | Text | Ảnh đại diện của địa điểm |
| 4 | url | Text | Đường dẫn đến chi tiết địa điểm |
| 5 | PlaceType | Text | Loại địa điểm có thể là trường học, bệnh viện, atm, v.v… |
| 6 | Address | Text | Địa chỉ của địa điểm |
| 7 | Longtitude | Float | Kinh độ của địa điểm này |
| 8 | Latitude | Float | Vĩ độ của địa điểm này |
| 9 | Note | Text | Ghi chú về địa điểm |
| 10 | Rating | Float | Xếp hạng về địa điểm này |

* 1. **HINHANH**
* Ý nghĩa: Chứa thông tin về hình ảnh của từng địa điểm

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ngữ nghĩa |
| 1 | PlaceID | Integer | ID của địa điểm |
| 2 | url | Text | url của hình ảnh này |

* 1. **COMMENT**
* Ý nghĩa: Comment về địa điểm

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ngữ nghĩa |
| 1 | PlaceID | Integer | ID của địa điểm |
| 2 | UserID | Integer | ID của người comment |
| 3 | Content | Text | Nội dung comment |

* 1. **RATING**
* Ý nghĩa: Rating về địa điểm

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ngữ nghĩa |
| 1 | PlaceID | Integer | ID của địa điểm |
| 2 | UserID | Integer | ID của người comment |
| 3 | Rate | Float | Xếp hạng cho địa điểm |

1. **Phân hoạch dữ liệu**
   1. **Dữ liệu thô**

Dữ liệu thô là dữ liệu có được sau khi thu thập thông tin về các địa điểm trong thành phố Hồ Chí Minh, gồm có: tên, địa chỉ, loại dịch vụ, tọa độ, chú thích.

**Đại học Khoa học tự nhiên**

227 Nguyễn Văn Cừ, P.4, Q.5, Hồ Chí Minh

10.762605, 106.682351

**Galaxy Tân Bình**

246 Nguyễn Hồng Đào, P.14, Q.Tân Bình, Hồ Chí Minh

10.790997, 106.640603

**KFC Điện Biên Phủ**

311 Điện Biên Phủ, P.4, Q.3, Hồ Chí Minh

10.776610, 106.683698

**Urban Station Coffee**

90 CMT8, P.11, Q.3, Hồ Chí Minh

10.776107, 106.685380

* 1. **Phân hoạch dữ liệu thô**

Dữ liệu thô ban đầu sẽ được định dạng lại theo dạng: tên địa điểm, loại địa điểm, địa chỉ, vĩ độ, kinh độ, chú thích.

|  |
| --- |
| Dữ liệu thô được chuẩn hóa |
| Đại học Khoa học tự nhiên; Trường học; 227 Nguyễn Văn Cừ, P.4, Q.5, Hồ Chí Minh; 10.762605; 106.682351 |
| Galaxy Tân Bình; Rạp chiếu phim; 246 Nguyễn Hồng Đào, P.14, Q.Tân Bình, Hồ Chí Minh; 10.790997; 106.640603; Giá vé: 50k – 85k |
| KFC ĐIện Biên Phủ; Quán ăn; 311 Điện Biên Phủ, P.4, Q.3, Hồ Chí Minh; 10.776610; 106.683698; Chuyên về gà rán |
| Urban Station Coffee; Quán Coffee; 90 CMT8, P.11, Q.3, Hồ Chí Minh; 10.776107; 106.685380 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tên địa điểm | Loại địa điểm | Địa chỉ | Vĩ độ | Kinh độ | Chú thích |
| Đại học Khoa học tự nhiên | Trường học | 227 Nguyễn Văn Cừ, P.4, Q.5, Hồ Chí Minh | 10.762605 | 106.682351 |  |
| Galaxy Tân Bình | Rạp chiếu phim | 246 Nguyễn Hồng Đào, P.14, Q.Tân Bình, Hồ Chí Minh | 10.790997 | 106.640603 | Giá vé: 50k – 85k |
| KFC ĐIện Biên Phủ | Quán ăn | 311 Điện Biên Phủ, P.4, Q.3, Hồ Chí Minh; 10.776610 | 10.776610 | 106.683698 | Chuyên về gà rán |
| Urban Station Coffee | Quán Coffee | 90 CMT8, P.11, Q.3, Hồ Chí Minh | 10.776107 | 106.685380 |  |

* 1. **Tạo ra các bảng dữ liệu từ dữ liệu đã phân hoạch được**

Dữ liệu sau khi phân hoạch được đưa vào bảng: DIADIEM

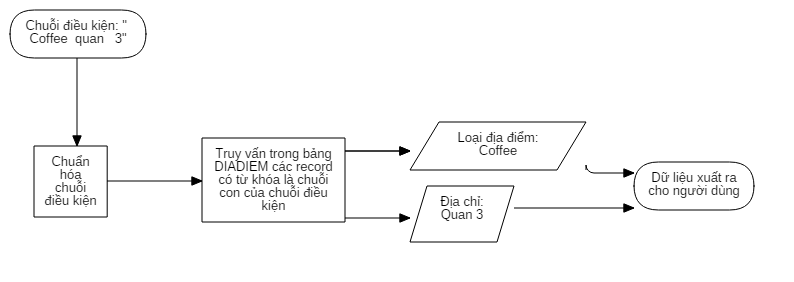
1. **Chứng minh tính khả thi**

Khi người dùng nhập vào thông tin cần tìm kiếm:

C:\Users\Lionheart\AppData\Local\Microsoft\Windows\INetCacheContent.Word\ttt.png

Ta có được chuỗi tìm kiếm là: “ Coffee quan 3”

* 1. **Sơ đồ quá trình truy xuất dữ liệu**



* 1. **Giảng giải sơ đồ**
* Chuẩn hóa chuỗi dữ liệu tìm kiếm:
  + Ta có chuỗi người dùng nhập vào là: “ Coffee quan 3”. Sau đó ta sẽ xử lý chuỗi bằng cách loại bỏ khoảng trắng đầu dòng, giữa dòng và cuối dòng, ta được chuỗi “Coffee quan 3”
* Truy vấn trong bảng DIADIEM các record có từ khóa là chuỗi con của chuỗi điều kiện:
  + Ta duyệt trong bảng DIADIEM xem các từ khóa trong bảng có là chuỗi con của chuỗi điều kiện hay không để rút ra thông tin tìm kiếm.
  + Ta có chuỗi “Coffee” là chuỗi con của chuỗi điều kiện và thuộc cột PlaceType và chuỗi “quan 3” thuộc cột Address, từ đó xuất ra dữ liệu cần thiết cho người dùng.

1. **Kết luận**

Như mô hình CSDL đã phân tích ở trên, chúng ta có thể thấy rằng dữ liệu chúng ta được tách ra thành 5 bảng riêng biệt (HINHANH, DIADIEM, User, RATING, COMMENT). Như vậy chúng ta có thể dễ dàng quản lí về danh sách địa điểm cũng như lúc tìm kiếm. Chúng ta sẽ quản lí user ở một bảng riêng biệt và việc tách bảng rating và comment ra riêng đảm bảo cho việc lưu trữ dữ liệu cũng như khi truy vấn làm cho chúng ta cảm thấy dễ dàng hơn.