

## Hướng dẫn sử dụng chi tiết

Trước khi chạy chương trình, người dùng cần down 2 folder là *corpus* và *metadata* từ link bên dưới và đưa chúng vào folder *release* trong bài nộp để chương trình chạy bình thường.

[https://drive.google.com/file/d/1KjhcqRbtIcS5ranqiDyKQV-DMkpsy\\_eO/view?fbclid=IwAR2zEcCgRkJLc-MpQKP6aIwjNzStZqOo59suVM9Ak0eoC7TtyTFn1kFFSdc](https://drive.google.com/file/d/1KjhcqRbtIcS5ranqiDyKQV-DMkpsy_eO/view?fbclid=IwAR2zEcCgRkJLc-MpQKP6aIwjNzStZqOo59suVM9Ak0eoC7TtyTFn1kFFSdc)

Khi người dùng bắt đầu chạy chương trình, màn hình console sẽ hiện lên một menu, có 3 lựa chọn ứng với 3 phần của chương trình và 1 lựa chọn thoát khỏi chương trình. Người dùng sẽ nhập lựa chọn của mình ở phía dưới menu.

```
***** VNSE SEARCH ENGINE *****
*                               *
*                               *
*                               *
* 1   *   Index                 *
* 2   *   Search                 *
* 3   *   Modify corpus          *
* 4   *   Refresh config         *
* 5   *   Exit                   *
*                               *
Your option: _
```

1. Khi người dùng lựa chọn số 1, tức là lựa chọn chức năng Index của chương trình, tại vì quá trình index file corpus này tốn nhiều thời gian, nên sẽ có lựa chọn cho người dùng.

```
***** VNSE SEARCH ENGINE *****
*                               *
*                               *
*                               *
* 1   *   Index                 *
* 2   *   Search                 *
* 3   *   Modify corpus          *
* 4   *   Refresh config         *
* 5   *   Exit                   *
*                               *
Your option: 1
WARNING: Indexing process will takes a long time to complete.
It may affect the performance
Do you really want to index the corpus? (Y/N): _
```

- Nếu người dùng muốn thực hiện index file corpus thì nhập vào “Y”, Lúc đó chương trình sẽ bắt đầu thực hiện index.
- Ngược lại, nếu người dùng ko muốn thực hiện index, thì nhập vào “N”, lúc đó chương trình sẽ quay lại menu ban đầu.

2. Nếu người dùng muốn tìm kiếm, thì nhập vào “2”, lúc này chương trình sẽ yêu cầu người dùng nhập vào thứ muốn tìm kiếm.

```
***** VNSE SEARCH ENGINE *****
*                               *
*                               MENU                               *
*                               *
*****
* 1 * Index *
* 2 * Search *
* 3 * Modify corpus *
* 4 * Refresh config *
* 5 * Exit *
*****
Your option: 2
===== VNSE Search Engine =====
The more meaningful keywords are, the faster the response is
=====++=====
Your query: 
```

- Tới đây, người dùng sẽ nhập vào thứ muốn tìm kiếm và chọn **Enter**, chương trình sẽ bắt đầu quá trình tìm kiếm, sau khoảng 15s màn hình sẽ hiện lên một dãy file document tìm được, sắp xếp theo chiều tốt nhất giảm dần từ trên xuống.

```
*****
*      VNSE SEARCH ENGINE - Deliver the best quality      *
*      Sorry for our long time response                    *
*      You may get better performance for meaningful keywords *
*****
Top 20 files:
1.      [0.138256] corpus\Am nhac\AN_VNE_ (256).txt
2.      [0.110860] corpus\Am nhac\AN_VNE_ (601).txt
3.      [0.105072] corpus\Am nhac\AN_VNE_ (186).txt
4.      [0.102185] corpus\Am nhac\AN_VNE_ (187).txt
5.      [0.100962] corpus\Am nhac\AN_TN_ (892).txt
6.      [0.096330] corpus\Am nhac\AN_VNE_ (610).txt
7.      [0.092298] corpus\Am nhac\AN_VNE_ (629).txt
8.      [0.089689] corpus\Am nhac\AN_VNE_ (177).txt
9.      [0.080270] corpus\My thuat\MT_VNE_ (22).txt
10.     [0.075314] corpus\Khong gian song\KGS_VNE_ (46).txt
11.     [0.072581] corpus\Giao duc\GDu_NLD_ (622).txt
12.     [0.071691] corpus\Am nhac\AN_VNE_ (606).txt
13.     [0.064221] corpus\Am nhac\AN_VNE_ (612).txt
14.     [0.056992] corpus\Am nhac\AN_VNE_ (811).txt
15.     [0.054749] corpus\Am nhac\AN_VNE_ (594).txt
16.     [0.049454] corpus\Chung khoan\CK_VNE_ (324).txt
17.     [0.049231] corpus\Am nhac\AN_VNE_ (345).txt
18.     [0.046300] corpus\Am nhac\AN_VNE_ (548).txt
19.     [0.045929] corpus\Am nhac\AN_VNE_ (575).txt
20.     [0.045340] corpus\Am nhac\AN_TN_ (880).txt
Select one of the following options:
      1. Read a document
      2. Back to menu
Your option: _
```

- Đến đây, chương trình sẽ đưa ra 2 lựa chọn cho người dùng
  - + Nếu người dùng muốn đọc nội dung của document nào đó, thì nhập “1”, tiếp theo người dùng sẽ nhập vào số thứ tự của document muốn đọc dựa trên danh sách document trước đó. Sau đó nội dung của document được chọn sẽ hiện ra để người dùng đọc, cùng với đó sẽ quay lại bước lựa chọn trước đó.
  - + Nếu người dùng muốn quay lại menu ban đầu thì nhập “2”, chương trình sẽ quay lại menu ban đầu.
- 3. Nếu người dùng muốn chỉnh sửa corpus thì nhập vào “3”, tới đây chương trình sẽ đưa ra 3 lựa chọn cho người dùng.

```

***** VNSE SEARCH ENGINE *****
*                                  *
*                      MENU                      *
*                                  *
*****
* 1   *   Index                                  *
* 2   *   Search                                *
* 3   *   Modify corpus                         *
* 4   *   Refresh config                       *
* 5   *   Exit                                  *
*****
Your option: 3

===== Modify corpus =====
1. Index new file
2. Remove a file
3. Back to menu
Your option: █

```

- Lựa chọn 1 là index 1 file mới vào corpus. Nếu muốn thực hiện chức năng này, người dùng nhập vào “1”, sau đó chương trình sẽ yêu cầu người dùng nhập vào đường dẫn đến file đó.

Lưu ý: file cần index vào corpus phải là file văn bản có encoding là UTF16-LE.

- + Nếu màn hình in ra “Successfully index new file” thì file đã được index thành công và sẽ trả về các lựa chọn trước đó.
- + Ngược lại thì index thất bại và vẫn trả về các lựa chọn trước đó.
- Nếu người dùng muốn xóa 1 file nào đó khỏi corpus thì nhập vào “2”, sau đó nhập vào đường dẫn của file đó.

Tương tự như phần index a file thì phần này cũng sẽ thành công nếu màn hình hiện ra “Successfully” và thất bại nếu ngược lại.

- Muốn thoát khỏi chức năng “Modify corpus” này thì người dùng nhập vào “3”, chương trình sẽ trở về menu ban đầu.

4. Nếu người dùng muốn chạy lại dữ liệu file config thì nhập “4”, sau khi nhập thì chương trình sẽ tự động cập nhật lại file config đã được chỉnh sửa và in ra số liệu hiện tại của file này ra màn hình để người dùng kiểm tra cùng với menu ban đầu để người dùng tiếp tục sử dụng.
5. Nếu người dùng muốn thoát khỏi chương trình thì nhập vào “5”, chương trình sẽ tự động thoát ra ngoài.