

**HƯỚNG DẪN LỌC HỘI THOẠI**

**MỤC LỤC**

[1.1. KHÁI NIỆM 3](#_Toc470711592)

[1.2. ĐĂNG NHẬP 3](#_Toc470711593)

[1.3. MÀN HÌNH CHỦ ĐỀ QUAN TÂM 3](#_Toc470711594)

[1.4. TẠO DỰ ÁN 5](#_Toc470711595)

[1.5. TẠO CHỦ ĐỀ 5](#_Toc470711596)

[1.6. HƯỚNG DẪN CÚ PHÁP TẠO KEYWORDS 8](#_Toc470711597)

[1.6.1. Tạo topic tracking mới 8](#_Toc470711598)

[1.6.2. Tìm kiếm mention sử dụng ô Search Phrase 9](#_Toc470711599)

[1.6.3. Lọc dữ liệu theo Platform 4](#_Toc470711600)

[1.6.4. Lọc dữ liệu theo nguồn (Source) 5](#_Toc470711601)

[1.6.5. Lọc dữ liệu theo nguồn Tác Giả (Author) có hội thoại trong hệ thống 7](#_Toc470711602)

[1.6.5.1. Tìm Author ID 7](#_Toc470711603)

[1.6.5.2. Gõ Cú Pháp Lọc Author ID: 9](#_Toc470711604)

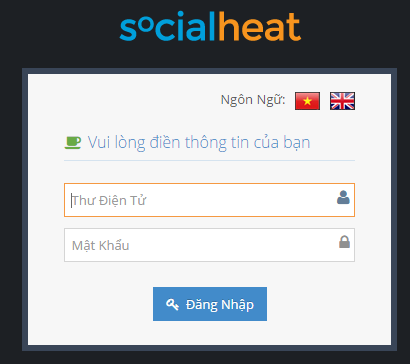
### KHÁI NIỆM

Dưới đây liệt kê các khái niệm được dùng trong SocialHeat

* **Thảo luận (mention):** Là bài viết, videos, post trên Facebook, Forums, Online News, Youtube, Instagram.
* **Chủ đề (Chủ đề):** Là tập hợp các thảo luận
* **Dự án (Project):** Là tập hợp các chủ đề
* **Tài khoản Khách Hàng (Account):** Mỗi khách hàng dùng hệ thống được tạo một tài khoản khách hàng để chứa các dự án và chủ đề. Tài khoản khách hàng khác tài khoản người dùng hệ thống
* **Câu truy vấn (Query):** Là cú pháp dùng kết hợp các phép toán AND , OR, NEAR. Exclude cùng với từ khóa (keywords) để lọc các thảo luận

### ĐĂNG NHẬP

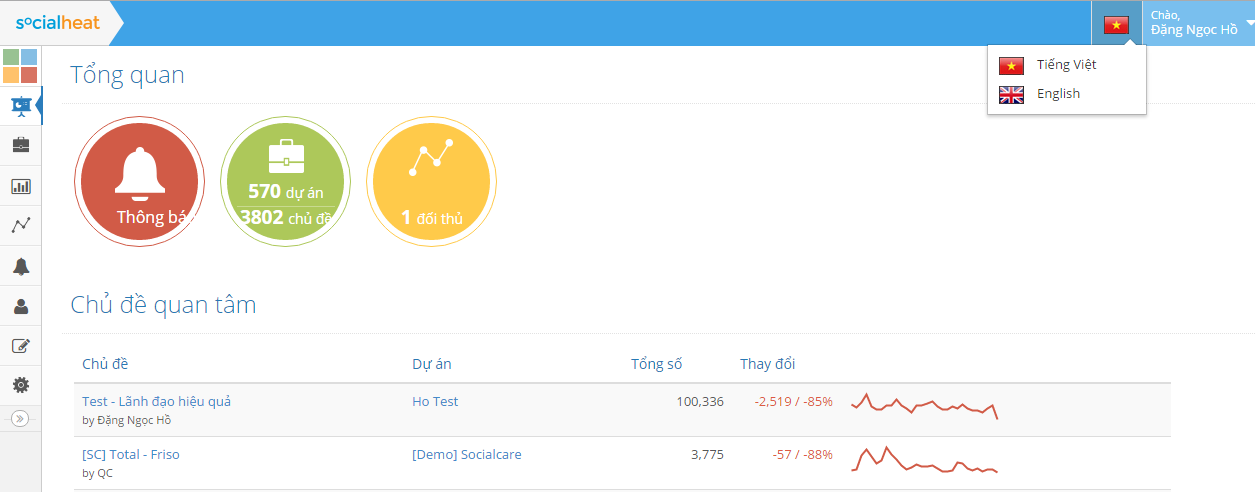
1. Vào địa chỉ: <http://socialheat.vn>
2. Chọn lá cờ Việt Nam cho ngôn ngữ tiếng Việt.
3. Đăng nhập dùng tài khoản được cung cấp (thư điện tử, mật khẩu).

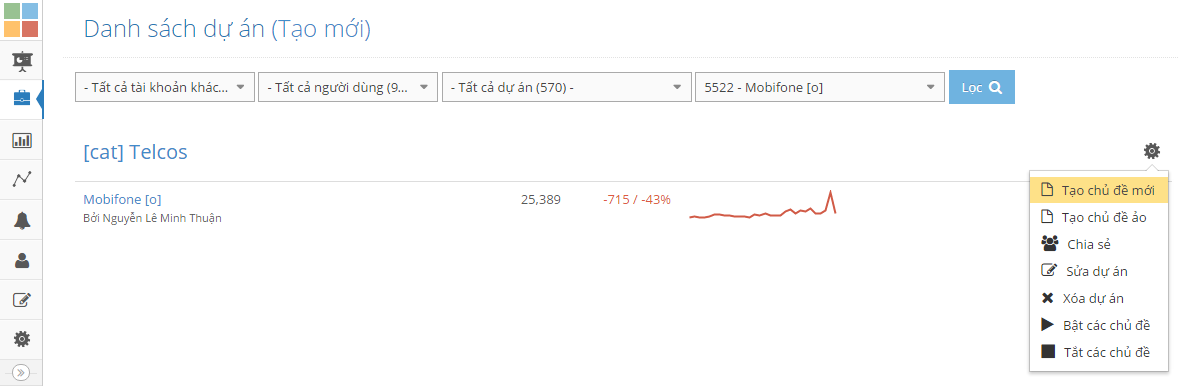


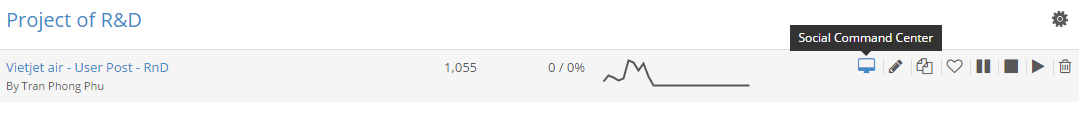
### MÀN HÌNH CHỦ ĐỀ QUAN TÂM

Sau khi đăng nhập, anh sẽ thấy màn hình tổng quan chứa danh sách các chủ đề đang được quan tâm, số lượng dự án, số thông báo, và số báo cáo về đối thủ. Trên góc trái có hình lá cờ để thay đổi ngôn ngữ.

Anh có thể click vào tên Dự Án để vào xem toàn bộ Chủ Đề





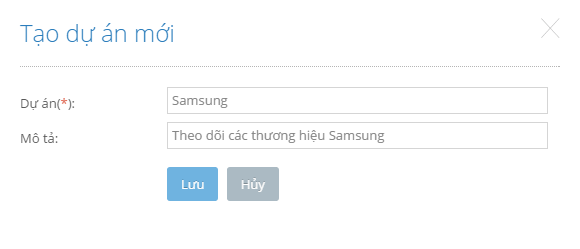


Các thông tin của từng Chủ Đề:

* Tên chủ đề
* Tổng bài tin, post, comment trên tất cả kênh trực tuyến có liên quan (từ khoá của chủ đề) trong 30-ngày gần đây nhất.
* Các chức năng ngay menu của Dự Án.
  + Tạo chủ đề mới
  + Tạo chủ đề ảo: Chủ đề Ảo là chủ đề chỉ dùng để đọc, không thể chỉnh sửa thảo luận. Chủ đề Ảo sẽ có nội dung là của chủ đề Gốc
  + Chia sẽ: Để chia sẽ chủ đề cho người dùng khác trong hệ thống.
  + Sửa dự án: Dùng để sửa thông tin dự án (tên, mô tả)
  + Xóa dự án: Xóa dự án sẽ xóa luôn các chủ đề trong dự án
  + Bật các chủ đề: Dùng để khởi động lại các chủ đề trong dự án đang bị tạm ngừng thu thập dữ liệu
  + Bật các chủ đề: Dùng tạm ngừng các chủ đề trong dự án

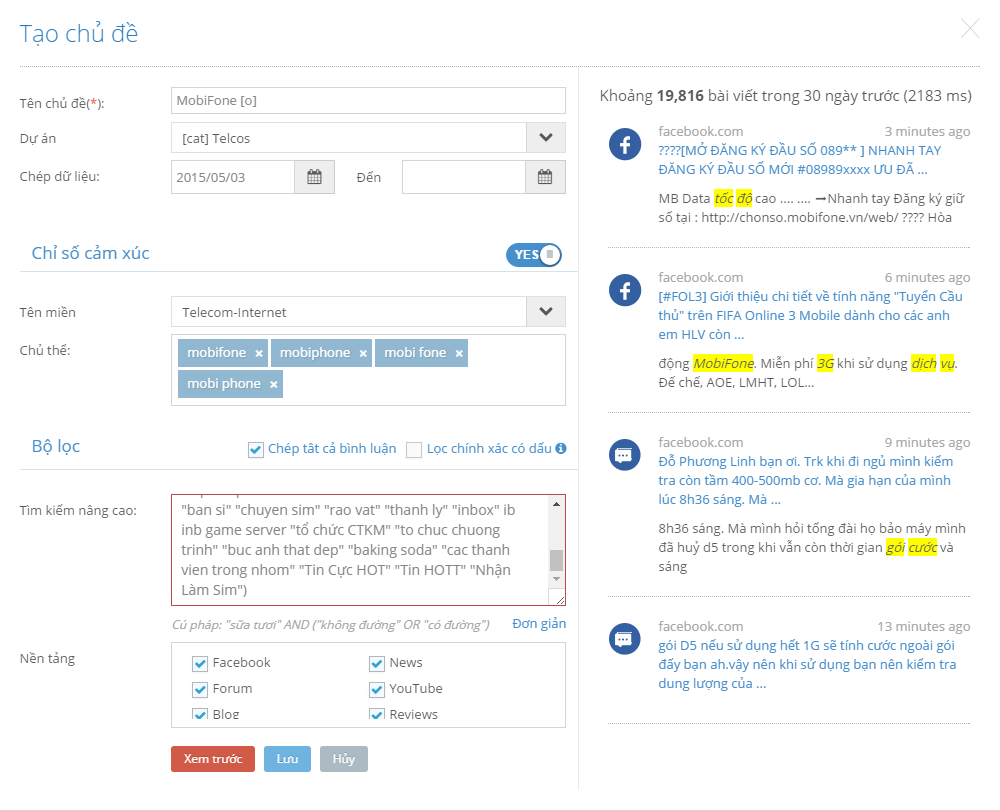
### TẠO DỰ ÁN

Chọn Tạo Mới ở Danh Sách Dự Án.



### TẠO CHỦ ĐỀ

Anh/ chị bấm nút Tạo Chủ Đề Mới ngay nút Răng Cưa của từng Dự Án để tạo Chủ Đề theo dõi dữ liệu trong trong thời gian thực



1. **Tên chủ đề**: Điền tên chủ đề
2. **Dự án**: Chọn dự án chứa chủ đề sau khi tạo
3. **Chép dữ liệu**: Dữ liệu sẽ được chép vào Chủ Đề trong khoảng thời gian được chọn. Khi ô Đến để trống thì dữ liệu được chép vào liên tục. Khi ô Đến có thời hạn xác định thì qua thời điểm đó dữ liệu không chép vào Chủ Đề nữa.
4. **Chỉ số cảm xúc**:
   1. **Yes/No:** Chọn "YES" nếu anh/ chị muốn hệ thống đánh giá chỉ số cảm xúc của từng thảo luận dựa vào các chủ thể bên dưới. Chọn “NO” hệ thống sẽ không đánh giá chỉ số cảm xúc trên từng thảo luận, việc này sẽ làm dữ liệu được đưa vào Chủ Đề nhanh hơn.
   2. **Tên miền**: Đây là những ngành hàng liên quan đến sản phẩm, dịch vụ, hoặc công ty chứa bộ tự điển, thuật toán cho việc đánh chỉ số cảm xúc. Anh/ chị hãy chọn ngành hàng liên quan đến Chủ Đề của anh/ chị để hệ thống đánh chính xác chỉ số cảm xúc. Nếu không có sẵn ngành hàng, anh chọn General. Anh/ chị có thể liên hệ với bên cung cấp hệ thống để tạo thêm ngành hàng mới.
   3. **Chủ Thể**: Anh/ chị nhập vào các biến thể của thương hiệu, chủ đề mà người dùng đề cập. Ví dụ, anh/ chị muốn biết người dùng đánh giá thế nào về Mobi Fone, anh/ chị nhập vào “MobiFone, Mobiphone, Mobi Fone” (dùng dấu , hoặc phím Enter để cách từng từ). Anh/ chị càng có nhiều biến thể đúng thì càng chính xác, bao gồm những từ sai chính tả có thể có.
5. **Bộ lọc**:

Anh/ chị tạo các câu truy vấn bằng cách kết hợp các toán tử AND, OR, NEAR, và EXCLUDE với các từ khóa liên quan đến chủ đề để thu thập các thảo luận. Sử dụng dấu phẩy để phân cách các từ.

* 1. "Chứa tất cả từ khóa": thảo luận chứa tất cả từ khóa trong ô này.
  2. "Chỉ chứa vài từ khóa:": thảo luận chỉ chứa một hoặc một và từ khóa trong ô này
  3. "Loại trừ": loại trừ các thảo luận chứa những từ khóa này.
  4. Nâng cao: dành cho những người sử dụng viết câu truy vấn phức tạp

Vui lòng xem phần kế tiếp để có hướng dẫn chi tiết hơn về cách sử dụng bộ lọc và câu truy vấn

1. **Nền tảng**: Các nền tảng là nguôn thu thập dữ liệu
2. **Tùy chọn:**
   1. **Tìm chính xác từ có dấu :** Tìm chính xác những thảo luận có chứa từ khóa có dấu. Ví dụ “điện thoại” thì chỉ tìm thảo luận chứa “điện thoại”, thảo luận chứa từ “dien thoai” không được thu thập
   2. **Chép tất cả bình luận**: thu thập tất cả các bình luận có từ khóa hoặc không có từ khóa của mỗi thảo luận.

Sau đó, nhấn nút Xem Trước để xem kết quả trả và điều chỉnh câu truy vấn cho phù hợp (nếu cần). Khi Ok thì bấm nút Lưu.

### HƯỚNG DẪN LỌC DỮ LIỆU

#### Tìm kiếm mention sử dụng ô Search Phrase



Cách viết câu truy vấn thông thường

Ví dụ câu truy vấn như sau:

**( (**"iphone 6" **AND** "chất lượng"**)** **OR** **(**"galaxy s6" **AND** "chất lượng"**) )** **AND -(**"thanh lý" "mua bán"**)**

* **Diễn dịch**: Tìm các bài viết có **(**chứa **CẢ** từ "iphone 6" **VÀ** từ "chất lượng"**)** **HOẶC** bài viết có **(**chứa **CẢ** từ "galaxy s6" **VÀ** từ "chất lượng"**).** **VÀ** tất cả các bài viết này **KHÔNG CÓ CHƯA BẤT CỨ MỘT TRONG NHỮNG** từ "thanh lý" **HOẶC** "mua bán".
* Cụm dấu **"**abcxyz **"** dùng để xác định CHÍNH XÁC MỘT CỤM TỪ và bắt buộc phải đúng thứ tự đã viết.
  + Lưu ý là dấu **“ ”** trong Microsoft Words khác với dấu **" "** trong hệ thống keywords. Khuyến khích sử dụng phần mềm hỗ trợ Notepad++ ([download tại đây](https://notepad-plus-plus.org/download/v6.8.3.html)) khi tạo keywords
* Từ khóa **OR, AND** là phải **VIÊT HOA** để không lấy bài viết tiếng Anh.
* Từ khóa **AND**
  + dùng để CHỈ LẤY bài viết có CHỨA CÙNG LÚC từ khóa này VÀ từ khóa khác – không cần phải đúng theo thứ tự
  + dùng để CHỈ LẤY kết quả của tập họp các câu truy vấn này GIAO Với tập họp các câu truy vấn khác.
* Từ khóa **OR**
  + dùng để lấy bài viết có chứa từ khóa này HOẶC từ khóa khác.
  + dùng để lấy kết quả của tập họp các câu truy vấn này HOẶC tập họp các câu truy vấn khác.
* Kết hợp dấu tổ hợp từ khóa **AND -(**“abc” “xyz”**)** để loại trừ các bài viết chứa các từ khóa không liên quan.
* Dấu - phải viết sát dấu (
* Lưu ý dấu - trong Microsoft Words có thể bị thay đôi do chế độ auto correct. Khuyến khích sử dụng phần mềm hỗ trợ Notepad++ ([download tại đây](https://notepad-plus-plus.org/download/v6.8.3.html)) khi tạo keywords

| **Ví dụ câu truy vấn** | **Giải thích ý nghĩa** |
| --- | --- |
| (("iphone 6" AND "chất lượng") OR ("galaxy s6" AND "chất lượng")) AND -("thanh lý" "mua bán") | Lấy các bài viết có chứa từ "iphone 6" và từ "chất lượng" HOẶC bài viết có từ "galaxy s6" và từ "chất lượng".Và tất cả các bài viết này không có chứa từ "thanh lý" hoặc "mua bán". |
| ("galaxy s6" AND "chất lượng") OR (id\_source:fb\_322004096814) | Lấy các bài viết có chứa từ "galaxy s6" và từ "chất lượng" HOẶC các bài viết nằm trong Source có ID là 322004096814 . |

**Lưu ý cần kiểm tra**

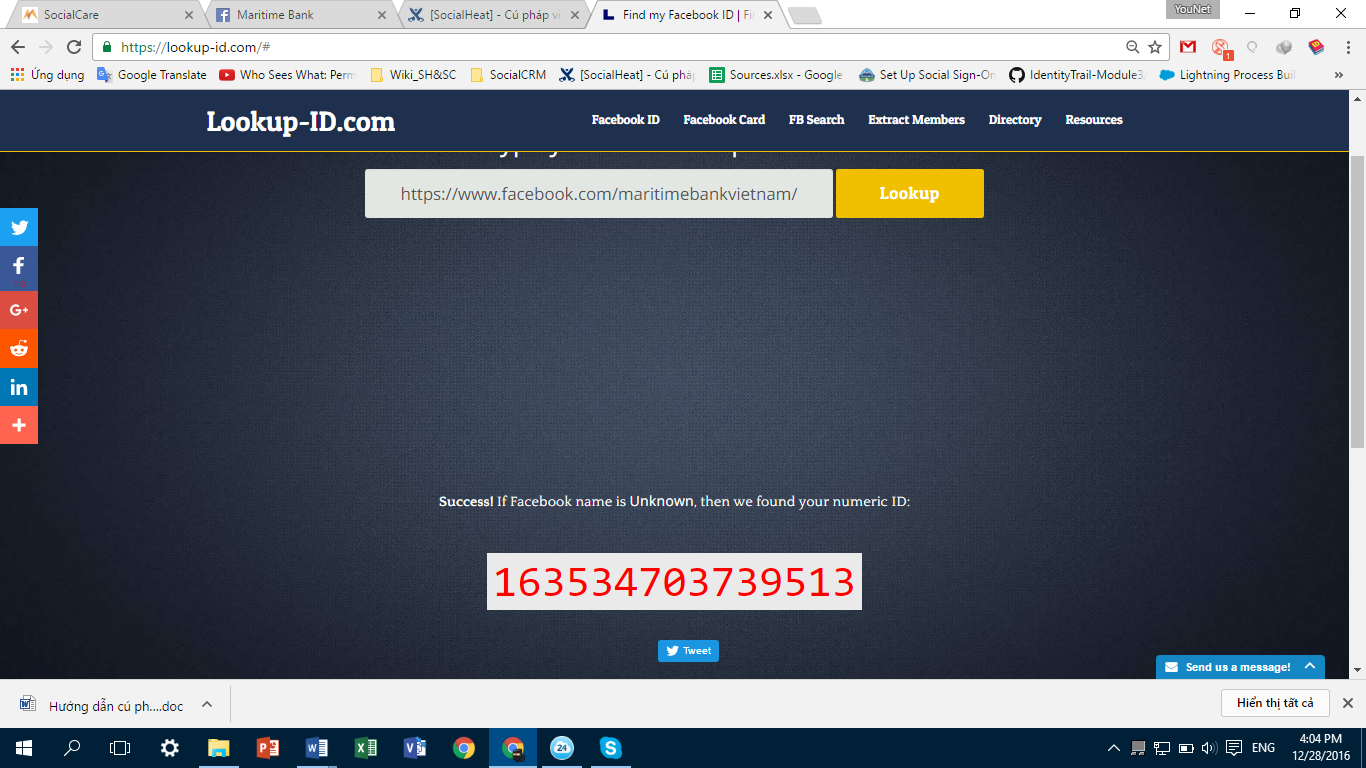
1. Kiểm tra các dấu ( ), " " phải đầy đủ, đúng thứ tự
2. Dấu nháy kép dùng để lấy đây đủ các từ
3. Khoảng trắng giữa các từ được hiểu là HOẶC.
4. Để Topic có dữ liệu trong Source nào đó thì dùng cú pháp **id\_source**(ví dụ chỉ muốn xem bài viết đến từ 1 group trên Facebook)

Cách tìm từ khóa theo kí tự đại diện

Thường dùng trong các trường hợp cần thu thập các link phát tán từ cùng 1 web domain.

#### Lọc dữ liệu theo nguồn (Source)

1. Tìm Source ID (id\_source)
   * Ví dụ Facebook:
     + Ví dụ ID là phần đỏ: <https://www.facebook.com/groups/1517161771885681/>
     + Nếu đường link không hiển thị rõ ràng ID của page (vd: https://www.facebook.com/maritimebankvietnam/) thì có thể dùng website [lookup-id.com](https://lookup-id.com/) để truy vấn.



1. Trong ô Câu Truy Vấn (Search Phase) gõ

id source của Facebook / Youtube có format là:

* Facebook: <fb\_id của source> . Có thể dùng [https://lookup-id.com/](https://lookup-id.com/%20) để tìm id
* Youtube: <id của source> . id source các bạn lấy trực tiếp từ đường link, chính là đoàn ký tự sau /channel/

Ví dụ câu query hoàn chỉnh để lấy mention từ fanpage YouNet Media sẽ là: id\_source:fb\_301428359876367

id source của News / Forum / Ecommerce / Review / Blog có format chính là domain của website luôn.

Ví dụ câu query hoàn chỉnh để lấy mention từ tinhte sẽ là: id\_source:tinhte.vn