|  |  |
| --- | --- |
| Kiệt | - Cào dữ liệu trực tiếp từ web (II) sử dụng **scrapy**  - Nghiên cứu cào 5 loại điểm của 1 user bất kỳ vì Element không có |
| Huy |
| Lượng | - Cào dữ liệu trực tiếp từ web (II) sử dụng **selenium,** |
| Khải | - Cào dữ liệu bằng API (I)  - Hiểu ý nghĩa từng trường, nhất là những trường có đánh dấu “?”  - Cân nhắc lấy thêm/bỏ đi các thuộc tính khác nếu cần, trong bản word này chỉ là sơ bộ  - Lưu trữ dưới dạng file csv (chia file cho fù hợp, đảm bảo tính lkết giữa các file) |
| Minh |
| Khánh | - Thử tìm kiếm API cho việc cào comment (có thể dựa vào I.2, phần “get comment”). Nếu nửa tuần chưa ra thì chung task với Lượng :v  - Cào comment trên app di động |

1. API
2. list các shop trong 1 địa điểm (fải lướt xem thêm ltục): [https://www.foody.vn/\_\_get/Place/HomeListPlace?page={i}&count={#shops}&districtId=&cateId=&cuisineId=&isReputation=&type=1](https://www.foody.vn/__get/Place/HomeListPlace?page=%7bi%7d&count=%7b#shops}&districtId=&cateId=&cuisineId=&isReputation=&type=1)

VD: <https://www.foody.vn/__get/Place/HomeListPlace?page=1&count=10&districtId=&cateId=&cuisineId=&type=1>

- Cào các trường sau cho 1 shop:

+ thông tin cơ bản: name, addr, CityId, RestaurantId, RestaurantStatus

+ avgrating

+ TotalPictures, TotalSaves, TotalReviews, TotalFavourites, TotalViews, TotalCheckIns

+ CreatedOn //ngày shop dc tạo đầu tiên trên foody?

+ Latitude & Longitude //toạ độ

+ ResUrlReviews //URL dẫn tới trang chứa bình luận của shop đó

1. info 1 shop:

* get delivery dishes: [https://gappapi.deliverynow.vn/api/dish/get\_delivery\_dishes?request\_id={shop\_id}&id\_type=1](https://gappapi.deliverynow.vn/api/dish/get_delivery_dishes?request_id=%7bshop_id%7d&id_type=1)

VD: <https://gappapi.deliverynow.vn/api/dish/get_delivery_dishes?request_id=786678&id_type=1>

+ dish\_type\_name //Loại đồ ăn: 1 loại đồ ăn gồm nhiều món ăn, VD: bún thì có bún cá, bún chả

+ total\_like //tổng lượt like loại đồ ăn đó

+ is\_available //loại đồ ăn đó còn món ko

+ dish\_name //món ăn thuộc loại đồ ăn đó. VD: bún cá thập cẩm

+ dish\_price //giá 1 món ăn thuộc 1 loại đồ ăn

+ options: // có option nào cho món ăn đó. VD: có option chọn thêm topping như rau, cua,...

[option\_1: name, price;

option\_2: name, price;

....

]

+ option\_min\_select, option\_max\_select

- get detail: [https://gappapi.deliverynow.vn/api/delivery/get\_detail?request\_id={shop\_id}&id\_type=1](https://gappapi.deliverynow.vn/api/delivery/get_detail?request_id=%7bshop_id%7d&id_type=1)

VD: <https://gappapi.deliverynow.vn/api/delivery/get_detail?request_id=1077197&id_type=1>

+ phone

+ district\_id

+ category //loại hình quán ăn?

+ cuisines //quán ăn thuộc phân khúc món nào? món Việt, Âu,...?

+ service\_fee //fí dịch vụ

+ avg\_price

+ min\_order\_value //fải mua tối thiểu bnhiêu tiền?

+ promotion // trong này nhiều trường nhỏ, tự nghiên cứu thêm

+ min\_charge //fí fụ thu nhỏ nhất

+ week\_days, operating //thời gian mở cửa

+ min\_shipping\_fee

+ is\_food\_delivery

+ price\_range

- get comments (chỉ get dc 10 cmt, ko get dc hết cmt, nghiên cứu thêm nếu có thể. Ko dc thì Crawl web thuần): <https://www.foody.vn/__get/Review/ResLoadMore?ResId=1077197&LastId=16625000&Count=10&Type=1&fromOwner=&isLatest=true&ExcludeIds=>

VD: <https://www.foody.vn/__get/Review/ResLoadMore?ResId=1077197&LastId=16625000&Count=20&Type=1&fromOwner=&isLatest=true&ExcludeIds=>

1. Crawl

Làm trước với 1 cửa hàng (vì url đi tới trang bình luận đã cào dc ở API nên chỉ cần thiết lập sẵn code đi tới đường url đó là dc)

- Điểm số:

+ location\_score //điểm vị trí

+ space\_score

+ quality\_score

+ serving\_score

+ price\_score

- review\_text:

excellence, good, average, bad

- Với từng user:

+ username

+ description //comment

+ date //ngày cmt

+ user\_location\_score, user\_space\_score, user\_quality\_score, user\_serving\_score, user\_price\_score

+ is\_return //có quay lại ko

+ n\_people //khách đi mấy người

* API: <https://gappapi.deliverynow.vn/api/meta/get_metadata>
  + Phương thức thanh toán (paynow)