



# ASSIGNMENT 2: ANNOTATION

## KEY POINTS

- ▶ Task1 – Relevance Annotation
- ▶ Task2 – Open end discussion on the process of Task1.

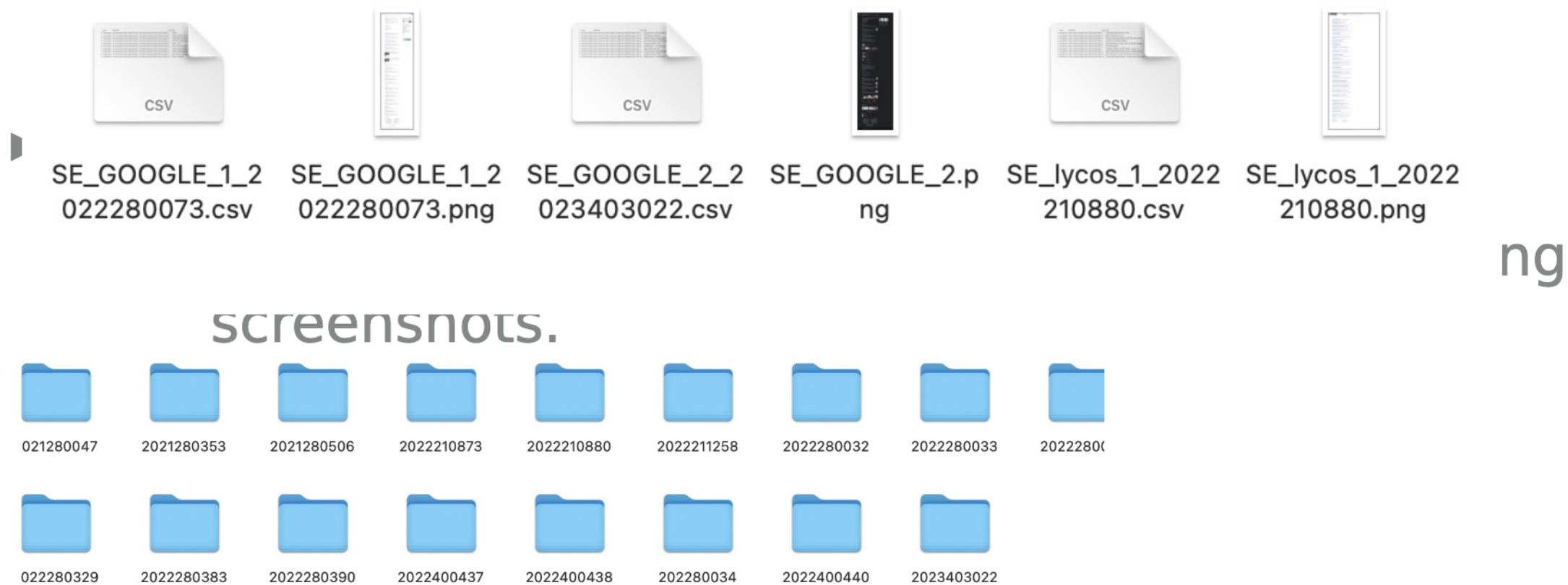
# REQUIREMENTS

- 1. Relevance Annotation.
  - Is the result relevant to the query when you only see the snippet?
  - Is the result relevant to the query when you see the landing page?
- 2. Submit Document (PDF) which include open end discussion on the process of Task1.
  - Is it difficult to annotate relevance of a result?
  - Is relevance annotation subjective or objective?
  - What are the factors that affect your relevance annotation for a page?
  - What are the differences when you annotate relevance of a result based on snippet or landing page?
  - Any other thoughts

# DETAILS FOR ANNOTATION

## STEPS FOR ANNOTATION

- ▶ Download data from Review( 批阅 ) of Assignment2-Data in WebLearning
- ▶ Check whether the folder name is yourStudentID



The screenshot displays a file explorer window with two rows of files and folders. The top row contains six files: three CSV files (SE\_GOOGLE\_1\_2\_022280073.csv, SE\_GOOGLE\_2\_2\_023403022.csv, SE\_lycos\_1\_2022\_210880.csv) and three PNG files (SE\_GOOGLE\_1\_2\_022280073.png, SE\_GOOGLE\_2.png, SE\_lycos\_1\_2022\_210880.png). The bottom row shows two rows of folders, each labeled with a student ID. The first row of folders includes 021280047, 2021280353, 2021280506, 2022210873, 2022210880, 2022211258, 2022280032, 2022280033, and 2022280034. The second row of folders includes 022280329, 2022280383, 2022280390, 2022400437, 2022400438, 202280034, 2022400440, and 2023403022. The text 'screenshots.' is visible in the middle of the image, and the text 'ng' appears on the right side.

# ANNOTATION DETAILS:

A

B

C

D

E

F

G

H

Anotate these columns

SE\_ BING\_2\_2021280506

1

2

3

4

5

6

7

8

9

10

11

	query	description	rank	title	url	title_score	url_score
0	beijing distance to inner mongolia	I want to travel to i	1	Soju - Wikipedia	https://en.wik	0	0
1	beijing distance to inner mongolia	I want to travel to i	2	Soju alcohol content: What is it? (percentage, proof and ...	https://drinkp	0	0
2	beijing distance to inner mongolia	I want to travel to i	3	Soju: Everything You Need to Know! Updated Guide (2022)	https://inepa	0	0
3	beijing distance to inner mongolia	I want to travel to i	4	Soju - Wikipedia	https://en.wik	0	0
4	beijing distance to inner mongolia	I want to travel to i	5	What is Soju? A Guide to Korea's National Drink   Man of ...	https://manof	0	0
5	beijing distance to inner mongolia	I want to travel to i	6	The Differences Between Soju, Shochu, and Sake, ...	https://inepa	0	0
6	beijing distance to inner mongolia	I want to travel to i	7	What Is Soju Alcohol Content And About Its History	https://alcohoc	0	0
7	beijing distance to inner mongolia	I want to travel to i	8	Why You Should Be Drinking Korean Soju Right Now	https://www.l	0	0
8	beijing distance to inner mongolia	I want to travel to i	9	How Much Soju To Get Drunk? - DrinkStack	https://drinks	0	0
9	beijing distance to inner mongolia	I want to travel to i	10	What Is Soju? - How To Drink Soju, the National Spirit of ...	https://www.t	0	0

## •Annotation

- Annotation should based on both the query and the description.
- When annotate the 'title\_score':
  - Scoring based on title and screenshot
  - Do not refer to the url content
  - If there is no corresponding content in the screenshot, you can score based on the title only
- When annotate the 'url\_score':
  - Scoring based on the url content(landing page)
  - Do not refer to the title and screenshot

## ANNOTATION DETAILS:

► Perfect > Good > Fair > Not relevant > Spam/Page not found

3 > 2 > 1 > -1 > -2

Under normal circumstances, you **should not** annotate as 0. 0 is used as the initial value, and used for bad data(for example, duplicate and wrong rows).

► Each student should annotate results of 3 queries.



Open CSV file in MS Excel, Numbers .....

After finishing task save as CSV (keep the CSV setting)

Keep the original file name



# ANNOTATION DETAILS: A TOOL FOR OPEN URLS

## ► Recommendation

- An Extension in Chrome – “Copy All Urls”
- Open several urls simultaneously in Chrome

扩展程序

更多扩展程序

已添加



**Copy All URLs**  
for Google Chrome

### Copy All Urls

提供方: Vincent Paré

Copy tabs URL to clipboard (formats : text, HTML, JSON or custom). Paste to open

★★★★★ 340 生产工具

评分



## TASK2 – OPEN END DISCUSSION

- Submit Document which include open end discussion on the process of Task1.
  - Is it difficult to annotate relevance of a result?
  - Is relevance annotation subjective or objective?
  - What are the factors that affect your relevance annotation for a page?
  - What are the differences when you annotate relevance of a result based on snippet or landing page?
  - Any other thoughts

► short answers is ok.

## SUMMARY

- ▶ HW2\_studentID.zip
  - ▶ 3 csv with original file name.
  - ▶ Discussion.txt(.md .pdf .word...)
  - ▶ Screenshots are **not needed**.

## **DEADLINE:**

**24:00(UTC+08:00, Beijing Time) Mar. 12,  
Sunday**

Submit your homework to web learning platform of thu:

[Http://learn.Tsinghua.edu.cn](http://learn.Tsinghua.edu.cn) our course section  
“Assignment”( 课程作业 )

# QUESTIONS ?