

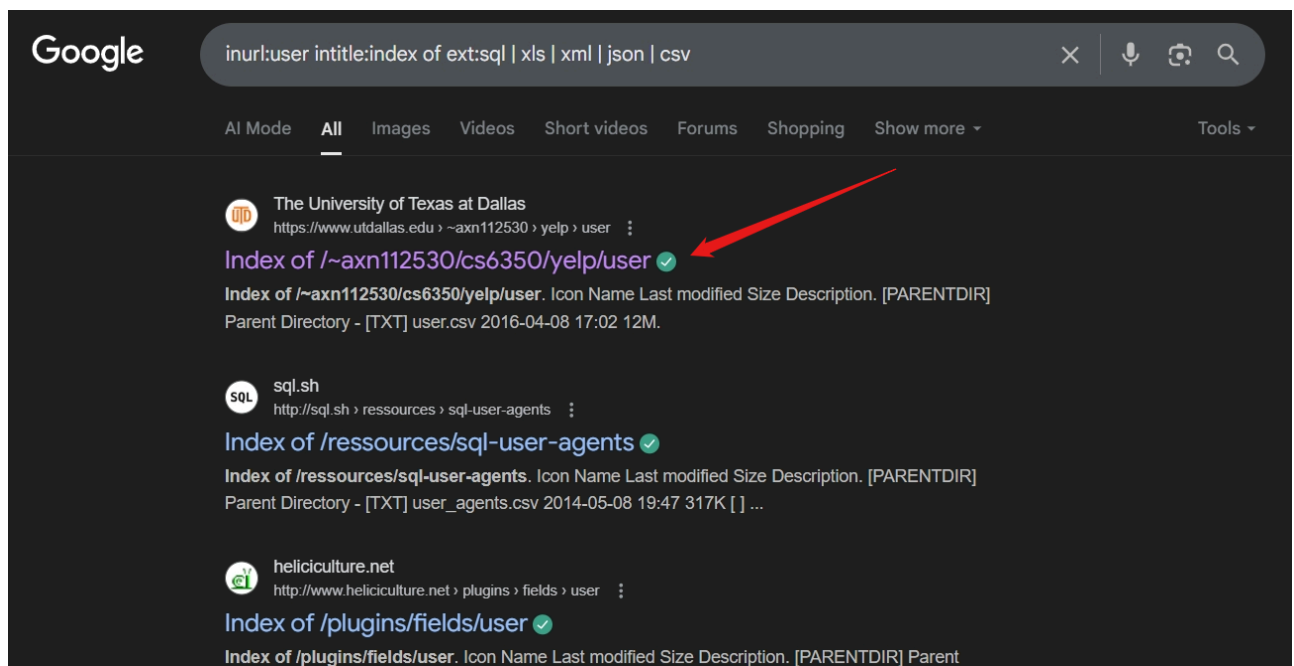
Google Dorking

This document demonstrates how Google Dorking uncovered unintentionally exposed user data in a CSV file on a university server, highlighting a significant cybersecurity oversight.

Step 1: Crafting the Dork Query

I started by entering the following Google dork into the search bar:

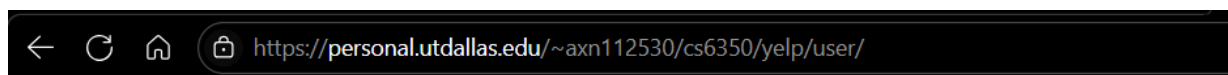
`inurl:user intitle:index of ext:sql | xls | xml | json | csv`



From the results, one particular link stood out:

[Index of /~axn112530/cs6350/yelp/user](https://www.utdallas.edu/~axn112530/cs6350/yelp/user)

The domain belonged to **The University of Texas at Dallas**, indicating a student or course-based directory.



Index of /~axn112530/cs6350/yelp/user

Icon	Name	Last modified	Size	Description
[PARENTDIR]	Parent Directory			-
[TXT]	user.csv	2016-04-08 17:02	12M	

Step 2: Visiting the Open Directory

I clicked on the link, which led me to a raw **index page** showing a list of files.

[Index of /~axn112530/cs6350/yelp/user](#)

There was a file named:

user.csv

Index of /~axn112530/cs6350/yelp/user

Icon	Name	Last modified	Size	Description
	[PARENTDIR] Parent Directory			-
	[TXT] user.csv	2016-04-08 17:02	12M	

The page had no login or protection — it was a simple file listing with metadata like size (12M) and date (2016-04-08).

Step 3: Observing the File Structure

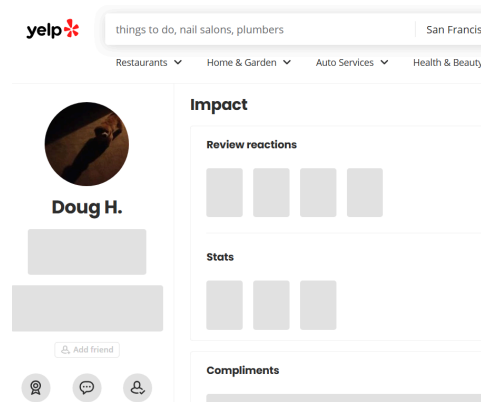
I downloaded and opened the CSV file using a spreadsheet viewer.

[illegible]

The file had multiple columns, including unique IDs, user names, and what looked like Yelp profile links.

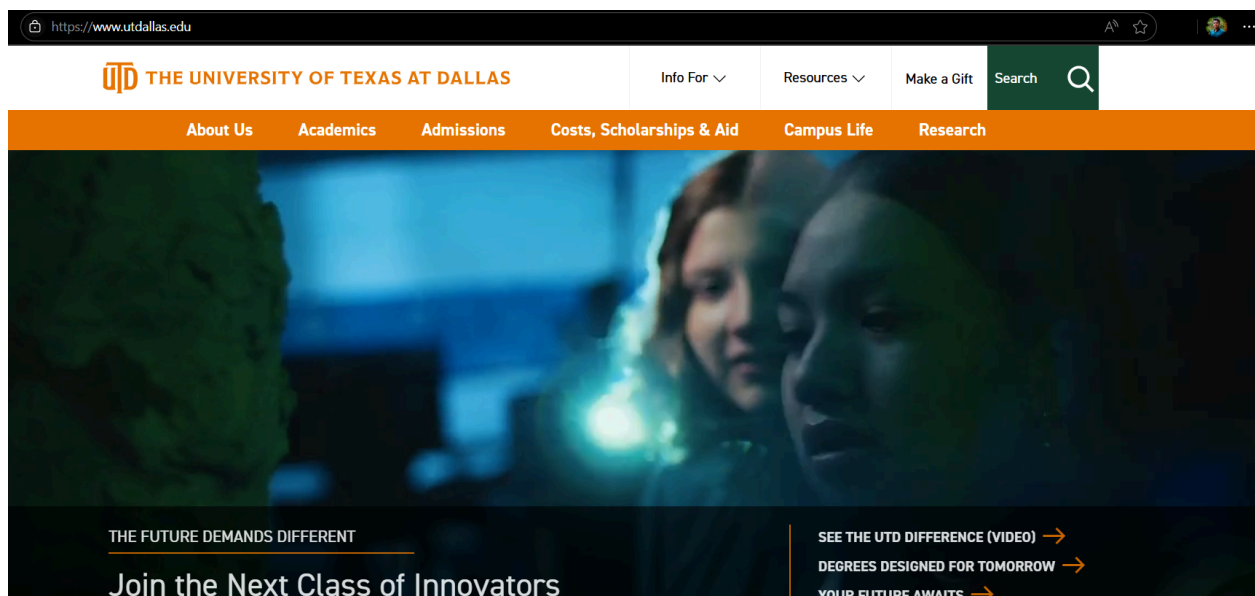
Step 4: Verifying a Single Entry

To test if the entries were real, I copy-pasted one of the profile links into a browser — specifically for **Doug H.** The page loaded successfully, matching the data from the CSV file.



Step 5: Checking the Source Domain

Finally, I visited the home page of the domain: <https://www.utdallas.edu> to confirm that the file was hosted under a university domain.



This verified that the source was an academic server.