

# 2023 DSAN scholarship competition

---

Each year the DSAN department holds a "scholar-ship" competition for students, which will start this week.

## Data-set description

---

- This dataset is the outcome of a web-crawling exercise aimed at identifying employment opportunities that could potentially interest DSAN students.
- There are roughly 85 searches, each yielding up to 10 job postings, for a total of around 850 jobs, which are currently active online, as of 04/14/2023.
- The postings were obtained using the following search query terms:
  - `["data scientist","data analyst","neural networks","big data and cloud computing","machine learning", 'reinforcement learning','deep learning', "time series","block chain","natural language processing"]`
- The searches are divided into two categories: nationwide and local searches for the United States as a whole and DC, respectively. It is worth noting that these searches are not mutually exclusive, and the files may contain duplicate job postings. These duplicates should be removed during the cleaning process.
- Furthermore, the DC search result extend outside of the specified search area. However, they likely contain a higher sampling of jobs in the DC area.
- The search results are stored in multiple JSON files, with the file name representing the search term. with each file containing the results of a single search

## Instructions

---

- Your job is to extract as much meaningful information from this dataset as possible and construct a data-driven narrative story about the current nature of the job market.
- This is a very open ended project and there are many directions in which you can take it.
- **Warning:** Do not use large-language model tools, such as chat-gpt, to create any of your content. Doing so will disqualify you from the competition.