


Problem Statement 1

There are millions of events hosted every year. We want to track new events launched or rehosted from available websites links.

Task for this test:

- We need to create a model that identifies upcoming event (start and end date) from the event websites
- Attaching - sample websites to run the test on -  Expired Event Data
- Note:
 - data could be on the main page, sub-pages or on an image
 - Date could be in different language (country)

Submission Guidelines:

- Include a document explaining the algorithm and implementation.
- Submit the source code that can be tested on another sample set

Problem Statement 2

There are millions of websites with list of companies attending an event (called exhibitors).

Problems statement: build solution to extract exhibitor lists from this websites, which may or may not be structured

- Assume a website contains a list of companies (exhibitors) with their names and additional fields like website url, location.
- Task is to extract company (exhibitor) names along with the websites
- You can store other information about the company (optional)

Some website url for extracting the data:

✚ Exhibitor Data 1 ; [Sample Listing URL 1](#)

✚ Exhibitor Data 2 ; [Sample Listing URL 2](#)

✚ Exhibitor Data 3 ; [Sample Listing URL 3](#)

Note:

- Goal is not to scrape the page on-time. Rather create a script that can run on other sites
- We would test your code by running it on another sample url to check its effectiveness

Submission Guidelines:

- Submit the source code for the exhibitor information extraction logic (Optional).
- Provide a document explaining the logic and methodologies employed / test cases demonstrating the effectiveness of your logic.
- Use Python or PHP for developing code.