## Job Offer: Web Content Analysis & Data Retrieval using Language Models

### Objective

Analyze a database of 1,500 links provided in a CSV format to extract specific details from articles linked therein.

### Tasks

1. Determine mentions of specified subjects (Paul Phua or Wei Seng Phua). Identify it is not relevant to him.

2. Identify and categorize:

- Judicial cases mentioned

- Mention if he belongs to a criminal group (14k, Triads, gang)

- Mentions of the acquittal of cases.

3. Extract the following:

- Language of the article.

- Date of the publication.

- Publisher's name.

- Publisher's specialty (particularly whether it is related to gambling).

- Publisher's country.

- Publisher's global contact email

### Preferred Technology

Bard API (or equivalent language model with web content understanding capabilities).

### Proposal Requirements

1 Outline of the Procedure: Detailed step-by-step plan to accomplish the tasks.

2.Technologies & Tools: List and justify the technologies you intend to employ.

3.Estimated Time: Provide an estimate of the number of hours required for the entire operation.

4.Alternative Solutions: If you believe there's a more efficient or accurate method than the Bard API, please specify.