Seminar Review

Project Guide : Mrs. Varsha Powar

Detection of country of origin of an application

- Goal of a system is to return country of origin of an application

  Automatically

Previous Systems

- They used manual data entry
- Not scalable
- Need manual intervention to change the country
- Number of applications info is limited

Detection of country of origin of an application

1. Web Mining/Scraping
2. Search key terms using regex (regular expression)
3. Using saved featured data

   I.city names

   II.companies

Web Mining/Scraping

❖ Technology used
  ➢ Puppeteer library
  ➢ Chromium driver
  ➢ Nodejs
  ➢ Regex

❖ Search terms Used
  ➢ App's name
  ➢ App's headquarters
  ➢ App's parent company name

```
var d = {
    "scrappedData": [
        {
            "key": "tiktok",
            "Result": [
                "TikTok is the destination for short-form mobile videos. Our mission is to capture and presen
                "TikTok/Douyin is a Chinese video-sharing social networking service owned by ByteDance, a Bei
            ],
        },
        {
            "key": "tiktok country",
            "Result": [
                "TikTok/Douyin (and formerly Musical.ly) users use the app largely to create, share, and view
                "Bytedance's Douyin app now has 400m daily active users in China. Douyin is the Chinese name
                "Jump to TikTok Usage by Country - Tiktok Usage By Country. The majority of TikTok's users ar
            ],
        },
        {
            "key": "tiktok headquarters country",
            "Result": [
                "TikTok is headquartered in Culver City, CA and has 3 office locations across 3 countries. Vi
                "Jul 9, 2020 - ByteDance may move TikTok headquarters out of China amid intense scrutiny. Chi
            ],
        },
```

Country Detector

❖ Use of featured data
  ➢ Companies
  ➢ Cities
  ➢ Countries
❖ Check which featured data matches the paragraph key words
  ➢ Mainly in address,parent company,headquarter
❖ Will return country code in every paragraph and priority
❖ Will map all country code and return final country code

```js
var d = {
    "scrappedData": [
        {
            "key": "tiktok",
            "Result": [
                "TikTok is the destination for short-form mobile videos. Our mission is to capture and presen
                "TikTok/Douyin is a Chinese video-sharing social networking service owned by ByteDance, a Bei
            ],
            "priority": 0.3,
            "CountryCode": "CN"
        },
        {
            "key": "tiktok country",
            "Result": [
                "TikTok/Douyin (and formerly Musical.ly) users use the app largely to create, share, and view
                "Bytedance's Douyin app now has 400m daily active users in China. Douyin is the Chinese name
                "Jump to TikTok Usage by Country — Tiktok Usage By Country. The majority of TikTok's users ar
            ],
            "priority": 0.3,
            "CountryCode": "CN"
        },
        {
            "key": "tiktok headquarters country",
            "Result": [
                "TikTok is headquartered in Culver City, CA and has 3 office locations across 3 countries. Vi
                "Jul 9, 2020 — ByteDance may move TikTok headquarters out of China amid intense scrutiny. Chi
            ],
            "priority": 0.4,
            "CountryCode": "CN"
        },
```