Net-Centric Lab 6

Student: Nguyen Duc Toan

ID: ITCSIU21112

Crawler.go

```go
package main

import (
    "bytes"
    "encoding/json"
    "fmt"
    "io"
    "net/http"
    "os"

    "golang.org/x/net/html"
)

type Manga struct {
    Title string `json:"title"`
    Link  string `json:"link"`
}

type Genre struct {
    Name  string  `json:"name"`
    Manga []Manga `json:"manga"`
}

func main() {
    genres := []string{"drama", "fantasy", "comedy", "action", "slice_of_life", "romance", "super_hero", "sf", "thriller", "supernatural"}

    var results []Genre

    for _, genre := range genres {
        mangaList := fetchMangaTitles(genre)
        results = append(results, Genre{Name: genre, Manga: mangaList})
    }

    file, _ := json.MarshalIndent(results, "", " ")
    _ = os.WriteFile("manga_titles.json", file, 0644)
}

func fetchMangaTitles(genre string) []Manga {
    url := fmt.Sprintf("https://www.webtoons.com/en/%s", genre)
    resp, err := http.Get(url)
    if err != nil {
        fmt.Println("Error fetching genre page:", err)
        return nil
    }
    defer resp.Body.Close()

    body, err := io.ReadAll(resp.Body)
    if err != nil {
        fmt.Println("Error reading response body:", err)
        return nil
```

```go
        doc, err := html.Parse(bytes.NewReader(body))
        if err != nil {
            fmt.Println("Error parsing HTML:", err)
            return nil
        }

        return findMangaTitles(doc)
}

func findMangaTitles(n *html.Node) []Manga {
        var mangaList []Manga
        var walk func(*html.Node)
        count := 0

        walk = func(n *html.Node) {
            if count >= 10 {
                return
            }
            if n.Type == html.ElementNode && n.Data == "a" {
                var link, title string
                for _, attr := range n.Attr {
                    if attr.Key == "href" {
                        link = attr.Val
                    }
                }
                for c := n.FirstChild; c != nil; c = c.NextSibling {
                    if c.Type == html.ElementNode && c.Data == "div" {
                        for gc := c.FirstChild; gc != nil; gc = gc.NextSibling {
                            if gc.Type == html.ElementNode && gc.Data == "p" && hasClass(gc, "subj") {
                                if gc.FirstChild != nil && gc.FirstChild.Type == html.TextNode {
                                    title = gc.FirstChild.Data
                                }
                            }
                        }
                    }
                }
                if link != "" && title != "" {
                    mangaList = append(mangaList, Manga{Title: title, Link: link})
                    count++
                }
            }
            for c := n.FirstChild; c != nil; c = c.NextSibling {
                walk(c)
            }
        }

        walk(n)
        return mangaList
}

func hasClass(n *html.Node, class string) bool {
        for _, attr := range n.Attr {
            if attr.Key == "class" && attr.Val == class {
                return true
            }
        }
        return false
}
```

Manga_titles.json

```json
[
  {
    "name": "drama",
    "manga": [
      {
        "title": "Love 4 a Walk",
        "link": "https://www.webtoons.com/en/romance/love-4-a-walk/list?title_no=6278"
      },
      {
        "title": "The Remarried Empress",
        "link": "https://www.webtoons.com/en/fantasy/the-remarried-empress/list?title_no=2135"
      },
      {
        "title": "To Whom It No Longer Concerns",
        "link": "https://www.webtoons.com/en/drama/to-whom-it-no-longer-concerns/list?title_no=6476"
      },
      {
        "title": "Marionetta",
        "link": "https://www.webtoons.com/en/fantasy/marionetta/list?title_no=4425"
      },
      {
        "title": "Teenage Mercenary",
        "link": "https://www.webtoons.com/en/action/teenage-mercenary/list?title_no=2677"
      },
      {
        "title": "Trapped in a Soap Opera",
        "link": "https://www.webtoons.com/en/drama/trapped-in-a-soap-opera/list?title_no=6751"
      },
      {
        "title": "The Greatest Estate Developer",
        "link": "https://www.webtoons.com/en/fantasy/the-greatest-estate-developer/list?title_no=3596"
      },
      {
        "title": "I'm the Queen in This Life",
        "link": "https://www.webtoons.com/en/fantasy/im-the-queen-in-this-life/list?title_no=4886"
      },
      {
        "title": "LUMINE",
        "link": "https://www.webtoons.com/en/fantasy/lumine/list?title_no=1022"
      },
      {
        "title": "Selfish Romance",
        "link": "https://www.webtoons.com/en/romance/selfish-romance/list?title_no=6555"
      }
    ]
  },
```

```json
      "name": "fantasy",
      "manga": [
        {
          "title": "The Remarried Empress",
          "link": "https://www.webtoons.com/en/fantasy/the-remarried-empress/list?title_no=2135"
        },
        {
          "title": "To Whom It No Longer Concerns",
          "link": "https://www.webtoons.com/en/drama/to-whom-it-no-longer-concerns/list?title_no=6476"
        },
        {
          "title": "Not Your Typical Reincarnation Story",
          "link": "https://www.webtoons.com/en/romance/not-your-typical-reincarnation-story/list?title_no=5556"
        },
        {
          "title": "Osora",
          "link": "https://www.webtoons.com/en/fantasy/osora/list?title_no=6202"
        },
        {
          "title": "Marionetta",
          "link": "https://www.webtoons.com/en/fantasy/marionetta/list?title_no=4425"
        },
        {
          "title": "The Dragon King's Bride",
          "link": "https://www.webtoons.com/en/romance/the-dragon-kings-bride/list?title_no=5517"
        },
        {
          "title": "Omniscient Reader",
          "link": "https://www.webtoons.com/en/action/omniscient-reader/list?title_no=2154"
        },
        {
          "title": "My In-Laws are Obsessed with Me ",
          "link": "https://www.webtoons.com/en/fantasy/my-inlaws-are-obsessed-with-me/list?title_no=3918"
        },
        {
          "title": "A Savage Proposal",
          "link": "https://www.webtoons.com/en/romance/a-savage-proposal/list?title_no=7121"
        },
        {
          "title": "I Thought My Time Was Up!",
          "link": "https://www.webtoons.com/en/romance/i-thought-my-time-was-up/list?title_no=5171"
        }
      ]
    },
```

...