Keylink Data Structure

Manual

1. What is Keylink Data Structure?

We often encounter projects and experiments where keywords are associated to the web links. For each of those keywords present in the list, one or more links are collected to gather and process the data related to it. Keylink data structure maintains a web link to every keywords present in the list. The number of links maintained is dependent on the project and the user requirement.

Keylink data structure supports the following operations:

- **Add a new key** – adds a new key to the list of keylinks
- **Collect the link** – gets the link from the web for the keys present in the keylink which do not have the associated links yet
- **Remove key** – removes a key and associated link from the keylink
- **Show** – prints the keylink with loaded keys and available links. Prints NULL if no links are collected for the key yet.

2. Assumptions and Applicability

The data structures works in an environment where system is connected to web. The data structure can be used in the projects like

- Crawling
- Parsing
- Data analysis
- Data Mining
- Etc.

3. Python Implementation

A python implementation of the concept can be found along with this manual. The requirements to run the code are:

- Python 3+ (Preferably 3.4)
- BeautifulSoup
- System connected to internet

The code collects only one link for the given key. The code can be easily extended to add any number of required links to the key.

The main() looks the following way:

```
def main():
    while True:
        print("Menu")
        print("--------------")
        print("1-Show\n2-Add Key\n3-Get Links\n4-Remove Key\n5-Exit")
        print("--------------")
        print("Enter your choice\n")
        choice = input()

        if choice == '1':
            show()
        elif choice == '2':
            add_key()
        elif choice == '3':
            get_links()
        elif choice == '4':
            remove_key()
        else:
            print("Program Terminates")
            break
```

Prakash B. Hegade

prakash.hegade@gmail.com