

## Aufgabe 1 Parsen von Webseiten

### Lernziele:

- Reguläre Ausdrücke verstehen und auf umfangreiche und komplexe Daten anwenden.
- Generische verschachtelte Java Hashtabellen anwenden.
- Generische Java Listen anwenden.
- Verwendung der Klasse Scanner vertiefen.

**Die Aufgabe muss zu Beginn des Praktikums vollständig bearbeitet sein!**

### Aufgabe

Gegeben eine Webseite, die Comicverfilmungen als Tabelle aufbereitet. Hier ein kleiner Auszug:

Asterix	Asterix und Obelix gegen Caesar (1999)
	Asterix & Obelix: Mission Kleopatra (2002)
	Asterix bei den Olympischen Spielen (2008)
	Asterix & Obelix – Im Auftrag Ihrer Majestät (2012)
The Atom	Legends of Tomorrow (Fernsehserie, seit 2016)
The Coldest City	Atomic Blonde (2017)
Atomik Circus – Le Retour de James Bataille (von <a href="#">Thierry Poiraud</a> )	Atomik Circus – Le Retour de James Bataille (2004)
Attack on Titan	Attack on Titan (2015)
Menschenblut	Aufbruch der Blutcrew (Kurzfilm, 1984)
Außer Kontrolle ( <i>Il Gioco / Click</i> , von <a href="#">Manara</a> )	Entfesselte Lust ( <i>Le Déclat</i> , 1985)
	Außer Kontrolle (Fernsehserie, 1997)
The Avengers ( <i>Die ruhmreichen Rächer</i> )	Marvel's The Avengers (2012)
	Avengers: Age of Ultron (2015)
	Avengers: Infinity War (2018)
	Avengers: Endgame (2019)
Avivato (von Lino Palacio)	Avivato (1949)
Azumi	Azumi (2003)
	Azumi 2 – Death or Love (2005)
<b>B</b> [ <a href="#">Bearbeiten</a>   <a href="#">Quelltext bearbeiten</a> ]	
Bäche und Flüsse (von Pascal Rabaté)	Les Petits Ruisseaux (2010)
Bad Kids Go to Hell (von Barry Wernick und Matthew Spradlin)	Bad Kids Go to Hell (2012)
	Bad Kids of Crestview Academy (2017)
Il brigante Grossi e la sua miserabile banda (von Michele Petrucci)	La banda Grossi (2018)
Barbarella	Barbarella (1968)
Barb Wire ( <a href="#">Dark Horse</a> )	Barb Wire (1996)

Parsen Sie den Inhalt der Webseite und wandeln Sie diesen in eine Java-Tabelle um, die unter den Comicnamen Tabellen enthält, die Jahresintervalle auf Listen von Verfilmungen (nur Name der Verfilmungen) abbilden.

**Beispiel:** Aus dem nachfolgenden Tabelleneintrag der Webseite

The Addams Family (Zeitungscartoons)	Die Addams Family (Fernsehserie, 1964–1966)
	Halloween with the New Addams Family (1977)
	Addams Family (1991)
	Die Addams Family in verrückter Tradition (1993)
	Addams Family – Und die lieben Verwandten (1998)
	Die neue Addams Familie (Fernsehserie, 1998–1999)

wird die der Java Tabelleneintrag:

```
{The Addams Family (Zeitungscartoons)=  
  { 1977=[ Halloween with the New Addams Family],  
    1991=[Addams Family],  
    1993=[Die Addams Family in verrückter Tradition],  
    1998=[Addams Family – Und die lieben Verwandten],  
    (1998,1999)=[Die neue Addams Familie (Fernsehserie)],  
    (1964,1966)=[Die Addams Family Fernsehserie]  
  }  
}
```

```
Map<String, Map<YearInterval, List<String>>> comicMap
```

### Hinweise:

- Die lokale Kopie der Webseite liegt im Projektverzeichnis. Wenn Sie das Programm starten, wird nach der Datei im Projektverzeichnis gesucht.
- Die Klasse für die Darstellung von Jahresintervallen (`YearInterval`) ist im Projekt enthalten.
- Zum Einlesen der lokalen Kopie der Webseite studieren Sie bitte das Beispiel aus der Vorlesung.
- Mit `Year.of(1993)` erzeugen Sie das Jahr 1993.
- Die Tabelle der Comic-Verfilmungen ist eine geschachtelte Tabelle mit List-wertigen Einträgen auf der untersten Ebene. Sie hat den Typ:

```
Map<String, Map<YearInterval, List<String>>> comicMap
```