Работа с регулярными выражениями

|  |  |  |
| --- | --- | --- |
| Название этапа | Листинг | Комментарии |
| Что такое регулярные выражения? | Нет листинга | Говоря простым языком, регулярное выражение — это последовательность символов, используемая для поиска и замены текста в строке или файле.  Регулярные выражения имеют богатую и не самую логичную историю. Изначально одна часть синтаксиса регулярных выражений была разработана учеными в 60-70 года с целью анализа текстов (можно было не руками посчитать, сколько раз в Войне и Мир было сказано слово Наташа в различных склонения) . Кроме того, в регулярных выражения существуют еще и символы, которые пришли из Unix – систем (наш горячо любимый альтернативный синтаксис). Кроме того, регулярные выражения испытали влияние замечательного языка программирования – Perl, который на заре развития Интернета широко применялся для обработки форм и проверки при помощи регулярных выражений. |
|  |  |  |
|  |  |  |
| Первые шаги в регулярных выражениях | https://regex101.com/r/Gh51p8/1 | Давайте посмотрим на просто слова, которые мы можем искать в тексте, если будет задавать просто символы. |
| Работа с точками | https://regex101.com/r/cJEyz0/1 | Работаем с точакми, к примеру, мы можем нам найти все as |
| Поиск всех окончаний и экранирование символов точки и слеша | <https://regex101.com/r/TAvJLq/1/> |  |
| Поиск какой-то определенной группы символов | https://regex101.com/r/wFEF9B/1 | Используем квадратные скобки для поиска определенной группы символы |
| Отдельный вариант с поиском окончаний одного и тоже слова | https://regex101.com/r/t4tWNO/1 |  |
| Регулярное выражение почты | https://regex101.com/r/03g5b7/1 | Делаем регулярное выражение, которое может обрабатывать почту  [\w.][+@\w+\.\w](mailto:+@\w+\.\w)+  Проагрейдженный вариант |
| Регулярное выражение, которое обрабатывает телефоны | <https://regex101.com/r/uUyUnw/1/> | Очень примитивный вариант на самом деле, можно немного улучшить, если есть желание |
| Работа с beautiful soup | | |
| Теория BS4 | Нет листинга | Прекрасный суп – это библиотека для Python, которая достает необходимую для вас информацию из файлов в формате HTML или XML. Это можно было бы достать регулярными выражениями, но куда проще воспользоваться готовой библиотекой. Она, к счастью или сожалению, требует некоторых познаний в HTML5, которые мы постараемся нагнать за сегодня. |
| HTML5 | Bs4/index.html | Пишем простую веб страницу с любыми тегами классами и id |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |