**Можливості пакетів json, xml та html**

**JSON** це текстовий формат обміну даних між комп'ютерами. JSON базується на тексті, може бути прочитаним людиною. Формат дозволяє описувати різні структури даних головним чином використовується для передачі структурованої інформації через мережу. Часто використовується, як заміна XML, через те, що займає мало місця і дозволяє утворювати різні складні структури. Створений для JavaScript проте вдало використовується і на інших мовах програмування, наприклад на Python, з певними змінами в типах даних.

| **Python** | **JSON** |
| --- | --- |
| Dict | object |
| list, tuple | array |
| Str | string |
| int, float, int | number |
| True | true |
| False | false |
| None | null |

 Має багато методів і багатофункіональний, що уможливлює якісний доступ до інформації і її обробку.

**Lxml** – досить швидка і корисна бібліотека для обробки, парсингу XML и HTML на Python. Часто використовується спільно з бібліотекою pandas для якіснішої роботи. Крім того дозволяє розкладати елементи у дерева. Особливістю бібліотеки є її швидкодія. Модуль lxml містить модуль objectify, який перетворює документи XML в об’єкти Python.

**Html parser.** Python - це одна з мов, яка широко використовується для зняття даних з веб-сторінок. Це дуже простий спосіб збирати інформацію. Наприклад, це може бути дуже корисним для швидкого вилучення всіх посилань на веб-сторінці та перевірки їхньої дійсності. HTML Parser, аналізує HTML / XHTML веб-сторінки та надає потрібну нам інформацію. Це клас, який визначається різними методами, які можуть бути переоцінені відповідно до наших вимог. Для використання HTML Parser, веб-сторінка повинна бути знайдена. З цієї причини, HTML Parser часто використовується з urllib2. Також містить багато методів для кращої обробки інформації.