1. Проект PyPy реализован для воспроизведения чистого дизайна, расшифровывается «Python, реализованный на Python», хотя насампере он реализован на опилках Python, которое називаэться RPython. Точнее PyPy представляет собой среду, где можно подключить любой язык программирования. Чистый дизайн языка PyPy позволяет создавать очень эффективные низкоуровневые оптимизаторы. В частности, в PyPy входит just-in-time (JIT)-компилятор. Это та же технология, которая произвела революцию производительности Java-программ в форме виртуальной машины HotSpot. С помощью проэкта PyPy наблюдается значительное ускорение програмы, по сравнению с CPython.

С помощью даного проэкта можна выполнять сбор ссылок на Web-страничке, анализировать html код, и, например, Извлечение списка статей с Web-страницы.

Ограничения:

Опасно работать с HTML напрямую, потому как этот язык слегка замысловат и нужен специальный анализатор. Для этого лучше использовать html5lib. Тем не менее, прочесывание Web ― обычно специализированная операция по извлечению информации в некоей конкретной ситуации. Для такого ограниченного использования pyparsing хорош и предоставляет некоторые тонкие возможности.

В ситуациях, когда синтаксический анализ применяется в сочетании с другими операциями, подход гринлетов позволяет структурировать обработку наподобие конвейеров командной строки UNIX. В тех случаях, когда задача осложняется работой с несколькими исходными страницами, возникает еще одна проблема: операции urllib2 не асинхронны, и вся программа блокируется всякий раз, когда обращается к Web-странице.