The suggestions for improvement of the web crawler

1. It could execute different processes for websites with http and https link, so have a match pattern block when it gets the link
2. Have a container or external database to store all of information it parsed from the website
3. Have multiple threads working on parsing simultaneously.