**Проект solver**

1. На определенной стадии развития проекта, рассмотреть пополнение онтологической базы путем выявления новых характеристик слов из толковых словарей и “проверенных” источников, таких как Wikipedia. Предполагается выявлять новые характеристики пропуская их через семантическое пространство…
2. В дальнейшем, для увеличения производительности: разделить словари по признаку частотности – наиболее употребляемые слова разместить в словарь, который будет проверяться в первую очередь и т.д. (Есть особенность: в любом случае для семантического дерева необходимо наличие всех слов). Для этого предполагается делать пометки для омонимов (одинаковых слов, фраз) в предыдущих словарях. Пометки будут указывать на следующий словарь, где необходимо искать такой же омоним, но с другими характеристиками.
3. Сделать универсальными все элементы системы (алфавиты, правила, разделители, направление текста и т.д.).
4. Настроить на web-сервере ограничение по IP-адресу на определенное время. Рассмотреть вариант capctha. Рассмотреть вариант пользователей, групп пользователей.
5. (<http://tomcat.apache.org/tomcat-7.0-doc/config/filter.html#Remote_Address_Filter>). Написать doFilter для контроля запросов (<http://www.jvmhost.com/articles/block-ip-address-apache-tomcat-filter>).
6. Касательно направления текста в арабском, китайском, японском языках – в текстах всегда есть начало и конец, поэтому направление не так важно.
7. Анализ текстов – это грубо говоря анализ последовательностей символов алфавита построенных по определенным правилам.
8. Могут быть фамилии точно такие как и слова, поэтому необходимо предусмотреть обработку и Слов и Фамилий на одном уровне – без мгновенного исключения из диапазона.