Семантическая аннотация в корпусных исследованиях

**Семантическая аннотирование (разметка)** - явное описание семантики контента ресурса при помощи понятий семантической модели (словаря, тезауруса, онтологии). Такое явное описание семантики выполняется указанием четкого соответствия между определенной частью контента ресурса и его семантикой, описанной в семантической модели. Аннотирование при этом базируется на RDF.

# Виды аннотаций

1. Неформальные описывается на естественном языке и поэтому обычно не обрабатывается с помощью программ.
2. Формальные аннотации составляются с использованием специальных языков (таких как XML и RDF), что позволяет выполнять их программную обработку.
3. Онтологические составляется на основе некоторой семантической модели (онтологии), описывающей основные понятия и отношения той предметной области, к которой относится описываемый ресурс, то она называется.

# Методы семантического аннотирования документов

При **ручном** способе аннотирования составлением метаописания документа занимается специалист. Однако программная система может оказывать ему помощь, например, показывая онтологии и имеющиеся экземпляры, выполняя проверку их правильности.

При **полуавтоматическом** методе аннотирования система автоматически составляет начальный вариант аннотации, а затем специалист может проверять и выполнить ее корректировку.

При **автоматическом** аннотировании все операции выполняет система без участия человека.

# Аннотации в GATE

* Универсальный способ представления информации о тексте
* Всегда соответствует некоторый отрезок текста
* Документ может иметь неограниченное количество аннотаций
* Всегда типизирована
* Может иметь атрибуты, набор которых расширяем

# Lancaster USAS Semantic Tagger Framework

Lancaster Semantic annotation tagger – традиционная система аннотирования для текстов. Разработан в 2004 году в центре UCREL Ланкастерского университета для английского языка (Rayson et al., 2004). Реализованы программные компоненты для работы с 10-ю языками. Разработан web-API для индексации текстов (http://ucrel.lancs.ac.uk/usas/).

# Ссылки

1. Новицкий А.В. Обзор некоторых направлений интеграции гетерогенных ресурсов в электронных библиотеках Петрозаводск:, 2009. 350–356 с.
2. Хоай Л., Тузовский А.Ф. Семантическое аннотирование документов в электронных библиотеках // Интеллектуальные системы. 2012. C. 157–164.
3. Oren E., Hinnerk M`ller K., Scerri S., Handschuh S., Sintek M. What are Semantic Annotations? URL: http://cites- eerx.ist.psu.edu/viewdoc/download? doi=10.1.1.97.7985&rep= rep1&type=pdf