

По цели создания различают тематические, общетематические, имиджевые и бизнес-блоги. Тематические ресурсы посвящены только одной теме: музыке, политике, путешествиям и т.д. Тематическими являются и сплоги (сайт-блоги или спам-блоги), созданные для раскручивания других ресурсов или осуществления спам-рассылок. Полной противоположностью этого вида блогов являются общетематические дневники, содержащие тексты на разные темы и ориентированные на среднего пользователя. Имиджевые блоги ведут как отдельные люди, так и целые компании для поиска новых бизнес-партнеров, клиентов или раскрутки собственного бренда.

По содержанию выделяют текстовые блоги, видеоблоги (влоги), подкасты, фотоблоги. Самым распространенным видом является текстовый блог, основу которого составляют статьи, хотя его автор может использовать видео и фотоснимки. Подкасты создаются для широкого круга читателей и представляют собой блоги, посетители которых могут загружать постоянно обновляемые аудиофайлы. Сайты, на которых реализуются подкасты, похожи на ток-шоу. Видеоблогов в настоящее время много, поскольку видеоконтент удобен для восприятия, а записать видео не является сложной проблемой. Довольно часто успешные влоги на канале Youtube привлекают внимание компаний, которые могут продвигать свою продукцию или услуги, спонсируя прямую или скрытую рекламу. Фотоблоги также очень оригинальный и популярный вид блога, часто реализуемый в социальной сети Instagram, где главное не тексты, а фотоснимки и небольшие заголовки или комментарии, сопровождающие фото. Существуют и более редкие разновидности блогов – тлог и микроблог. На тлоге публикуют различные виды информации. Но для одной записи используется только один формат: видео, текст либо ссылки. Первым бесплатным сервисом тлогов является Tumblr. В микроблоге публикуют небольшие сообщения, ссылки или видео. Самым известным сервисом микроблогов признана социальная сеть Twitter.

Таким образом, как гипертекстовый документ и одновременно интерактивный инструмент речевого взаимодействия в форме полилога, блог обслуживает неофициальные публичные коммуникативные ситуации. При этом ситуация речевого взаимодействия автора и комментаторов блога имеет много общего с реальным разговором, который реализуется посредством письменного текста. В докладе приводятся фрагменты франкоязычных блогов разных типов, иллюстрирующих приведенные выше положения.

А. Кацубинская

ПРЕДСТАВЛЕНИЕ ИНФОРМАЦИИ В АВТОМАТИЧЕСКОМ ПЕРЕВОДНОМ СЛОВАРЕ

Электронный текст имеет нелинейную структуру, включает несколько параллельных слоев информации. С одной стороны, текст состоит из языковых элементов различного уровня (слова, фразы, предложения), с другой стороны, он состоит из структурных сегментов различных типов (заголовки,

сноски, ремарки, цитаты, таблицы, формулы). При анализе текстов учитывается как языковое, так и структурное членение текста. Многие структурные элементы текста имеют яркие языковые особенности, которые требуют специальной формы представления. Разметка – это расстановка в тексте документа специальных маркеров (тегов), которые эксплицируют «скрытые» элементы информации, присутствующие в тексте. Разметка широко используется в корпусах текстов, электронных словарях и т.д. Электронный словарь подразумевает определенный способ выборки словника, множество текстов, из которых делается эта выборка, дополнительную информацию, которая сопровождает каждое слово.

В настоящее время лексикография находится под сильным воздействием новых методов обработки информации. Структура словарной статьи в разных электронных словарях различна. Словарная статья содержит разнотипную лингвистическую информацию. Словари включают в свои словарные статьи следующие сведения: заглавное слово, его тематическую принадлежность, грамматическую информацию, неформализованные толкования или стандартизированные дефиниции, лексическую сочетаемость, семантические иерархические связи заглавного слова, стилистическую окрашенность, лингво-географические ограничения, контексты, различную служебную информацию.

В нашей работе материалом исследования служат англоязычные тексты по экономике, взятые с Интернет-сайта <http://marketwatch.com>. По всем анализируемым текстам были получены алфавитно-частотные словари с помощью программы DIST. Практическое теггирование полученных словарей по проанализированным англоязычным текстам осуществлялось на основе системы кодирования, разработанной на кафедре информатики и прикладной лингвистики. Так, например, согласно кодировочной таблице существительное *economy* кодируется следующим образом: *ECONOMY* – NO1S42; *GLOBAL* – AQPOOO21. Нами были закодированы все тексты из отобранного массива англоязычных текстов.

Для пополнения лингвистической информационной базы данных нами были использованы программы, которые позволяют накладывать закодированный словарь на текст. Для получения закодированного текста между словом и тегом был проставлен символ «_» без пробела, каждое слово вводилось с новой строки. Тем самым, каждое словоупотребление текста получает соответствующую информацию о различных лингвистических признаках.

В. Керножицкая

ПОДХОДЫ К АВТОМАТИЧЕСКОЙ НОРМАЛИЗАЦИИ ЛЕКСИЧЕСКИХ ЕДИНИЦ В СООБЩЕНИЯХ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ

Участники сетевого общения зачастую используют формы выражения мыслей, далекие от традиционных лексических норм. Так, в текстах интернет-коммуникации встречаются аббревиатуры, отражающие, по сути,