

И. В. Ромашкевич

**Информационно-поисковые языки
и поиск в Интернете**

Помимо естественных языков, в человеческом обществе используется большое число искусственных языков. Информационные языки предназначены для оптимизации выполнения основной функции языка – коммуникативной. Они служат для однозначной записи различного рода информации с целью ее дальнейшего накопления и поиска. Оптимизация достигается путем устранения вариантности выражения смысла средствами языка, а также применением простого аппарата грамматики. Проблема построения информационных языков в наше время приобрела особое значение в связи с ростом потока научно-технической информации. Для удобства обработки информации с помощью компьютеров она (информация) должна записываться в компактной и однозначной форме средства-

ми того или иного информационного языка. На сегодняшний день существует немалое количество информационных языков. С точки зрения типологии информационно-поисковых языков браузеры в Интернете работают на основе унитарно-языков, т.е. языков, использующих ключевые слова, которыми в данном случае могут быть абсолютно любые слова, входящие в документ (вплоть до союзов, предлогов и т.д.). У таких языков есть большой недостаток: они не учитывают смысловые отношения между ключевыми словами. Это приводит к увеличению выдачи документов (т.е. большее количество документов может быть выдано в ответ на искомый запрос), но вместе с тем и увеличивается так называемый поисковый шум – возрастает количество нерелевантных документов, не соответствующих желаемой теме запроса. Кроме того, при поиске в Интернете большое значение имеет так называемый критерий организации интерфейса той или иной поисковой системы, а именно: предлагается ли нам под каждой найденной ссылкой просмотреть часть документа, содержащую те предложения, в которых есть интересующие нас слова, или нет. Наличие таких предложений значительно сокращает время поиска и одновременно увеличивает его комфортность. Немаловажным фактом является также то, что разные поисковые системы имеют свою «ориентацию», т.е. в ответ на наш запрос по какой-либо теме одна система выдает более полное множество документов, чем другая. Разные «поисковики» различаются также своими языками запросов. Исходя из всего вышесказанного, мы приходим к необходимости иметь более систематизированные знания в области информационно-поисковых систем и оценки работы существующих браузеров, что и является целью нашей работы. В данной работе проводится анализ различных поисковых систем, языков запросов, сравниваются и оцениваются результаты поиска по теме «компьютерная лингвистика» для четырех различных поисковиков (Yandex, Google, Yahoo, I.ua), после чего даются рекомендации по их использованию в области заданной темы (более подробно результаты эксперимента описаны в докладе).