

**«ЭВК-2010»**

**14 сентября 2010**

## **Круглый стол «Полнотекстовый поиск»**

**Салов А.А.**, руководитель департамента корпоративных Интернет-решений компании Softline

### ***«Решения для организации полнотекстового и параметрического поиска для электронных архивов и библиотек»***

В докладе – история создания и внедрения проекта построения полнотекстового поиска по электронной библиотеке в РГБ. Пилотный проект выполнен в срок, и в августе проведена успешная презентация результатов. На слайдах презентации был продемонстрирован профессиональный поиск vs проекта, показан интерфейс поисковой системы в рамках пилотного проекта РГБ, интерфейс альтернативной ПС по данным пилотного проекта РГБ, экономическое обоснование выбранной системы, сравнение базового функционала GSA и Fast. Представлена Google Search Appliance как пример универсального поиска. Докладчик рассказал также о Центре поисковых решений Softline.

**Подобайло Н.Н.**, директор департамента ОАО «Армада»

### ***«Библиотечная поисковая система. Практический опыт создания решения на базе промышленной СУБД в РГБ»***

Постановка задачи звучала достаточно просто - создать информационную систему, которая:

- включит в себя весь существующий объем электронных библиографических описаний, которые есть и появятся со временем в РГБ;
- обеспечит эффективное управление этими записями;
- обеспечит эффективный поиск по массиву этих записей, в том числе полнотекстовый поиск по содержанию электронных документов, которые соответствуют библиографическим описаниям.

Выбор технологической платформы пал на ведущего программного вендора – ORACLE. Полнотекстовый поиск с учетом коллекций и подсчетом статистики потребовал применить фирменный механизм ORACLE – партицирование по коллекциям, который позволил наконец добиться приемлемых результатов по показателям производительности, в наиболее часто встречающихся запросах (поиск с учетом коллекции, и переход из коллекции в коллекцию). Представлены различные возможности и сервисы системы.

**Авраменко А. В.**, начальник департамента программирования ООО «ИТ Системы»

### ***«Реализация системы полнотекстового поиска по электронным коллекциям РГБ на базе поискового сервера SPHINX»***

Докладчик продемонстрировал способы доступа к информации и использование электронных ресурсов в библиотеке «от прошлого к будущему», оцифровку документов, систему поиска документов как с точки зрения читателей, так и требований библиотеки к системе поиска. Перечислены особенности систем полнотекстового поиска для

электронных библиотек и преимущества решения разработчика. Продемонстрированы варианты простого и расширенного поиска.

**Шварцман М.Е.**, начальник отдела исследования компьютерных систем РГБ  
*«Кто ищет лучше? Кто видит дальше?»*

В докладе дана сравнительная оценка 3-х полнотекстовых систем: Армада, ИТ системы и Softline.

Было предложено сравнение по параметрам: соответствие ТЗ, скорость загрузки (дозагрузки) данных, полнота и точность поиска, правильность отображения найденного, удобство интерфейса, набор операторов; на разных массивах документов; на различном железе: **ИТ Системы** - Dual Core Xeon E5140; 16GB RAM HP Smart Array E200i RAID1, массив RAID6 на EMC; **Softline** - Dual Core Xeon 5148LV, 16GB RAM, 5x250GB RAID 5; **Армада** – 4 процессора, 24 GB RAM, 2x250 GB RAID 0. Сравнивались производительность и функциональность, морфология и операторы; поиск по атрибутам, фильтр результатов по коллекциям и т.д. В итоге были определены плюсы и минусы каждой из рассматриваемых систем.