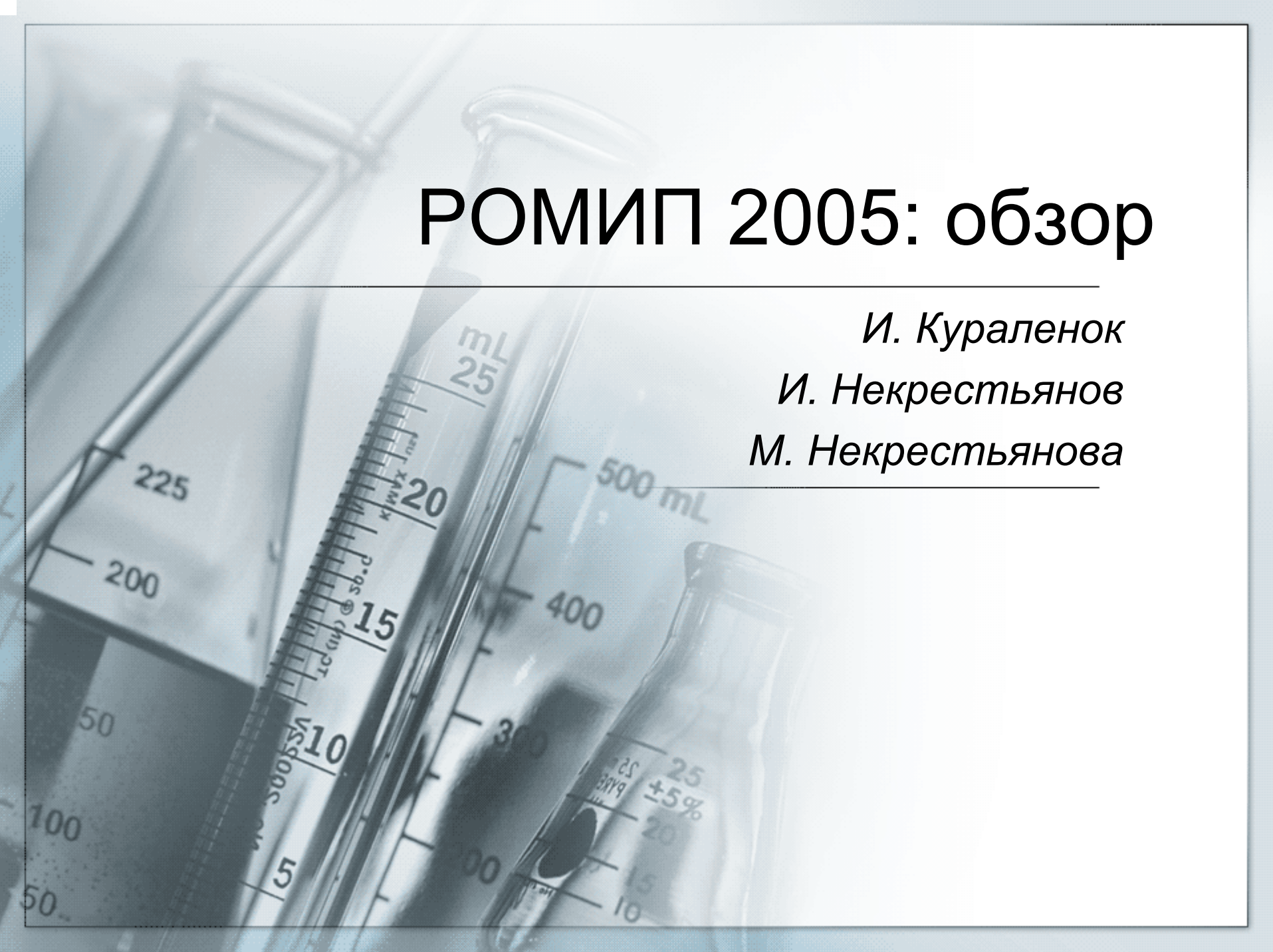




РОМИП 2005: обзор

И. Кураленок

И. Некрестьянов

М. Некрестьянова

Что такое РОМИП?

Российский семинар по оценке методов информационного поиска

- Русскоязычные задания
- Крупные коллекции
- Использование апробированных подходов к оценке
- Равноправие и анонимность участников
- Использование независимых экспертов для оценки результатов поиска
- Возможность повторного использования результатов

Что требуется от участника?

- Подать заявку
- Участвовать в формировании правил проведения дорожек
- Выполнить полученные задания и сдать ответы в оргкомитет
- Проанализировать результаты оценки и подготовить статью
- Компенсировать часть затрат на проведение оценки и организацию семинара
- Сделать доклад на очном семинаре

Как производится оценка?

- Задания для оценки отбираются оргкомитетом ПОСЛЕ сбора ответов участников
- Участник может забраковать до 10% предложенных оргкомитетом заданий
- Оценка на основе сравнения с «эталоном»
- «Эталон» обычно строится с помощью ассессоров, которые оценивают ответы на правильность
- При повторении дорожки часть заданий оценивается повторно
- По результатам вычисляются некоторые стандартные метрики

Построение «эталона»

- Оцениваются все *уникальные* ответы
- Ассессор НЕ знает какой системы это ответ
- Один ассессор оценивает ответы разных систем
- Порядок выдачи заданий для оценки случаен
- Обычно минимум два ассессора оценивает одно и тоже задание
- Многозначная шкала (4 значения)
- Использование расширенных описаний
(Цель – упразднить неоднозначность трактовки)

Коллекции

	Состав	Размер	Предоставлена
Narod.Ru	Веб-сайты из домена narod.ru	728.000 док. 22 000 сайтов.	Яндекс
Legal	Законодательство РФ	60.000 док.	Кодекс
DMOZ	Веб-сайты из русской части DMOZ	300.000 док. 2087 сайтов	Рамблер
News	Все новости за три периода из 17 источников	24.000 док. 50 Мб.	Яндекс

Дорожки по поиску

	Заданий	Оценивалось	Объем оценки	Согласие асессоров
Поиск по Веб	24166	75 (25)	500	0.89
Поиск по норм. коллекции	12925	83 (30)	200*	-
Поиск по смешанной коллекции	37091	158	-	0.89
Поиск по образцу	8205	58	-	-

Дорожки по поиску (2)

- «Поиск по образцу»
 - Задание: запрос + документ
 - Запросы из оценивавшихся в 2004 году
 - Образцы – релевантные документы (AND)
 - Задача ассессора: проверить документы на соответствие *запросу*
- Совместная оценка
- «Расширенные описания» от ассессоров
- Использование экспертов для нормативных запросов (но только одна оценка)

Тематическая классификация

	Категорий	Обучающее множество	Оцени- валось	Объем оценки	Согласие асессоров
Веб сайтов	247	2116 сайтов	26 кат. 6971	240	0.91
Веб докум.	247	2116 сайтов	26 кат. 9734	240	0.89
Нормат . докум.	183	13772 док.	40 кат.	0	-

- Выборочная оценка Веб страниц (10%)

Аннотирование

Задание: запрос + документ

(21872 задания на 158 запросах)

Ответ:

фрагмент текста < 300 символов без HTML разметки

Проверяем:

по аннотации можно предсказать релевантность документа

Эталон:

результаты оценки дорожек поиска 2004

Аннотирование (оценка)

- Асессор видит фрагмент + заголовок документа
- Мало дубликатов
- Оценка аннотаций индивидуально (а не списком)
- Асессор смотрит на одинаковое число ответов разных систем
- Асессор не видит альтернативные ответы разных систем
- Смесь: 50% сильно-, 25% слабо и 25% нерелевантных
- Всего: 960 (39) + 1970 (56)
- Асессоры «НЕ-юристы»

Извлечение фактов

Основа: новостная коллекция

1. Извлечение имен собственных

- Ответ: все вхождения имен собственных во всех документах
- Классы: персона/организация/геогр. объект/прочее

2. Извлечение фактов

- факт = 2 фигуранта + отношение + фрагмент текста
- Отношения: «Владеет»/«Работает на»

Сложность:

как получить переиспользуемый результат?

Извлечение фактов (оценка)

- Ассессор проверяет факт, указывает тип ошибки и правильное выделение
- Ошибки: не тот фигурант, неверное отношение
- Отдельно
 - Проверка фигурантов
(неточно выделенный фигурант, не имя собственное)
 - Выявление одинаковых фактов
- Выбор фактов для оценки
 - Все факты в наборе документов (полнота выделения)
 - Все факты для заданных фигурантов (полнота досье)
 - Требуется согласования синонимов
- Всего оценили 3000 фактов

Организация новостей

Задача: структурировать поток новостей в набор событийных сюжетов, связанных между собой.

Сюжет: отражение события из реального мира в потоке новостных сообщений
(«смерть папы римского»)

Надсюжеты – сюжеты, связанные между собой нетематическими связями
(«перенос футбольного тура в связи с кончиной папы»)

- Только 1 участник дошел до финиша
- Оценка еще не произведена

Проблемы 2005

- Опоздания со сдачей результатов участниками
- Большая пиковая нагрузка на оргкомитет
- Опоздание с оценкой QA и новостной дорожкой
- Юридический статус
 - Права собственности на коллекции
 - Оплата участия

Эволюция

	2003	2004	2005
Коллекции	1	3	4
Дорожки	2	5	10
Участники (заявки)	7 (9)	9 (11)	14 (17)
Число прогонов	14	34	98
Объем оценки	550	1300	2000