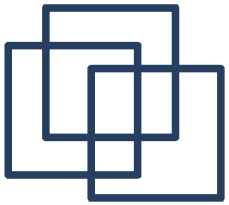
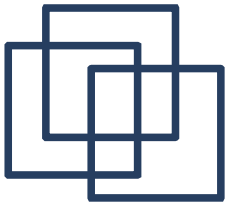


Опыт построения системы защиты электронных библиотек от несанкционированного копирования документов

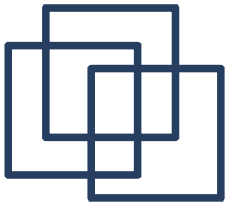
Евгений Ивашко
Наталья Никитина



Информационная безопасность —
обеспечение

1. целостности
2. доступности
3. конфиденциальности
4. защиты от **несанкционированного полного копирования**

Несанкционированное полное копирование (НПК) —
получение копий большой части документов библиотеки
без согласия владельцев



Угрозы несанкционированного полного копирования:

1. Нарушение авторских прав
2. Фишинг (подложный web-сайт)

Как защищаться от НПК?

Методы:

1. ограничение круга доступа
2. платный доступ
3. юридические ограничения
4. технические ограничения (ограничение интенсивности и количества обращений)

Технические ограничения

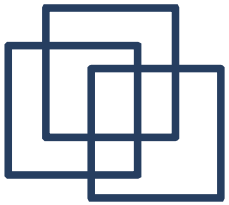
Ограничение интенсивности обращений

Ограничение количества обращений

Какой порог устанавливать?

Будет ли эффективное ограничение зависеть от
* ЭБ?
* количества документов в ЭБ?
* тематики контента?

Цель: создание интеллектуальной системы защиты от НПК



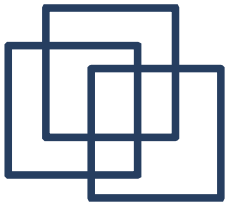
Каковы цели пользователей ЭБ?

1. обычные пользователи: обращение к документам с целью найти необходимую информацию
2. злоумышленники: обращение к документам с целью скачать как можно больше уникальных документов

Семантическая связь между документами

* скачиваемых обычными пользователями: есть
* скачиваемых злоумышленниками: нет

Вывод: если научиться определять имеют ли документы смысловую связь, то можно обнаруживать НПК



Этапы построения системы защиты от НПК:

1. построение модели «нормального» поведения
   * на основе «нормальных» данных
   * на основе онтологий

2. создание классификатора поведения пользователей:
   оцениваем отклонение поведения пользователя от «эталонного» «нормального» поведения

3. настройка параметров
   эффективное допустимое отклонение
   * по соотношению числа ложных срабатываний и пропущенных атак
   * по скорости обнаружения

Модель «нормального» поведения – взвешенная сеть

Для обнаружения НПК высчитываются веса переходов м/у документами, определяя отклонение поведения пользователя от «эталонного»

Эксперименты:

* объект: Электронная библиотека Республики Карелия

* период: июнь 2007 г. - февраль 2009 г.

* объем: > 1000 документов, обращения зафиксированы к > 700

Исходные данные:

* лог-файл с IP-адресами пользователей и названиями документов

* отброшены потенциальные proxy-сервера

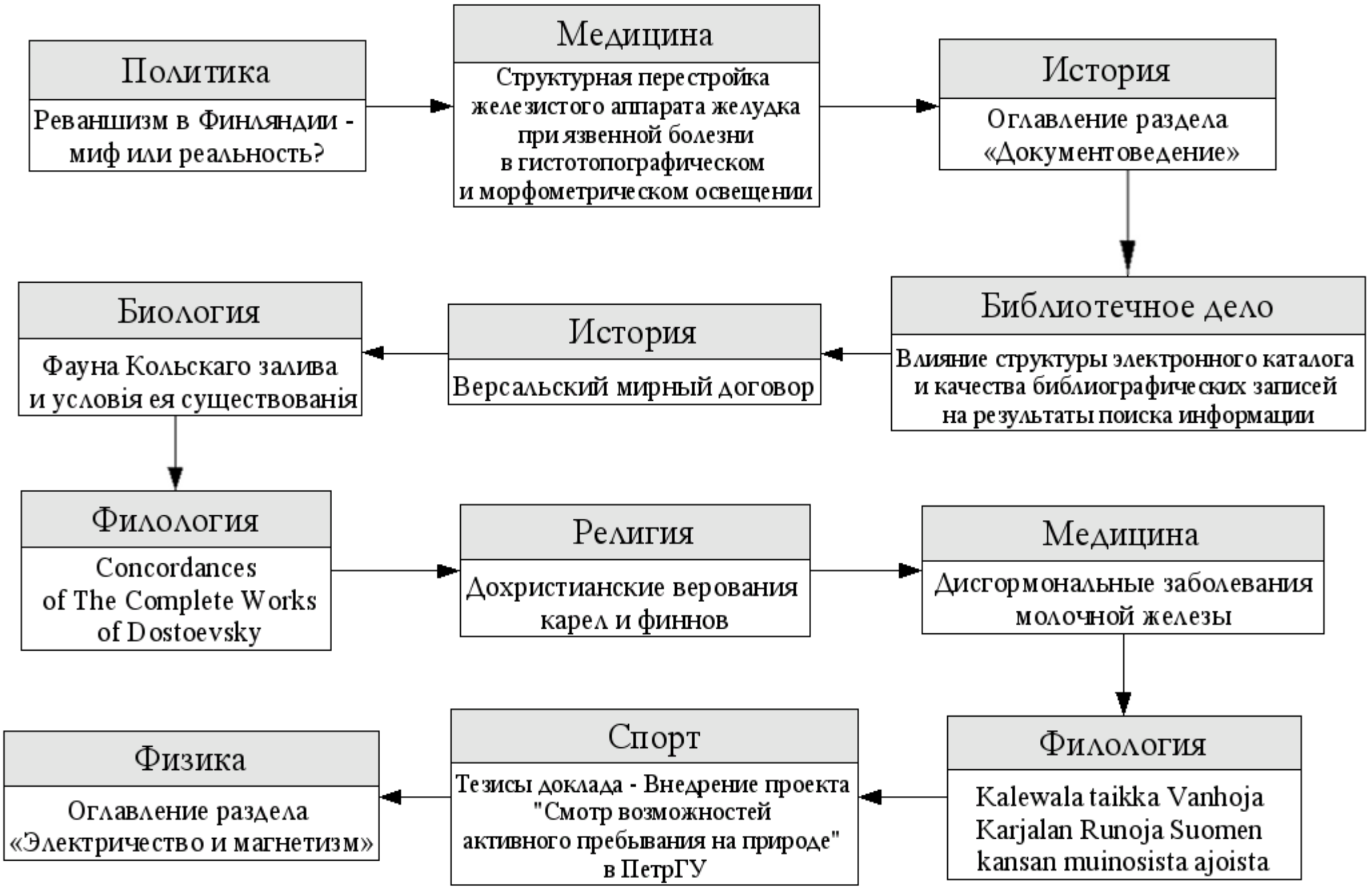
* отброшены неинформативные сессии (<10 обращений к документам)

* сессия работы — последовательность всех обращений с одного IP-адреса

* лог-файл содержит 4718 сессий

Пример «нормального» профиля

Пример аномальной сессии

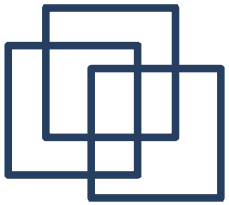


Результаты

1. Разработана программная система для проведения экспериментов.

2. На основе анализа поведения пользователей возможно автоматически выявлять семантические связи между электронными документами.

3. Возможно автоматическое выявление последовательностей обращений, противоречащих семантическим связям между документами.



Спасибо за внимание!

«Копирование целых томов или выпусков журналов, а также использование для этих целей автоматических поисковых систем (роботов) категорически запрещено. Организации, нарушившие это правило, лишаются доступа в библиотеку на год (такие случаи уже были), а при повторном нарушении — навсегда» - «Условия пользования научной библиотекой РФФИ»

Лицензиат не может сам или давать разрешение другим лицам

(viii) извлекать идеи, алгоритмы, процедуры, трудовые процессы или иерархическую последовательность из Программного Обеспечения или Документации с целью создания какой-либо работы, которая может использоваться Лицензиатом или третьей стороной в качестве замены для Программного Обеспечения или Документации;

(ix) создавать какой-либо продукт, предназначенный для конкурирования со Службами Данных, Программным Обеспечением или Документацией в течение периода действия Лицензии и в течение одного (1) года после него; или помимо разрешенных копий, копировать для любой цели Информационные Услуги или какую-либо их часть.

ДОГОВОР О МЕЖДУНАРОДНОЙ ЛИЦЕНЗИИ КОНЕЧНОГО ПОЛЬЗОВАТЕЛЯ

[Только для непосредственных клиентов Value Line Publishing, Inc.]

Онтология - спецификация концептуального представления предметной области, соответствующая поставленным задачам (T. Gruber, M. Ushold).