

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИ

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ**

**ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

**(ДГТУ)**

Кафедра «Информационные технологии»

**Методические указания**

к выполнению лабораторных работ по дисциплине

«Мировые информационные ресурсы»

Ростов-на-Дону

2024

УДК 004.4

Составители: В.А. Ландышев, Д.Н. Климова

Методические указания к выполнению лабораторных работ по дисциплине «Мировые информационные ресурсы». - Ростов-на-Дону: Донской государственный технический университет, 2024. - 35 с.

Содержат рекомендации и задания для выполнения лабораторных работ по дисциплине «Мировые информационные ресурсы». Рассматриваются исследования применимости законов Зипфа к русскоязычным текстам, информационный поиск в сети Интернет,основы обработки персональных данных в информационных системах, и др.

Предназначены для магистров направления подготовки 09.04.02 «Информационные системы и технологии» всех форм обучения.

Ответственный за выпуск зав. кафедрой «Информационные технологии»,

д-р техн. наук, профессор Б.В. Соболь

В печать \_\_\_\_\_\_

Формат 60×84/16. Объем\_\_\_ усл.п.л.

Тираж \_\_\_экз. Заказ №\_\_\_.

Издательский центр ДГТУ

Адрес университета и полиграфического предприятия:

344000, г. Ростов-на-Дону, пл. Гагарина,1

©Донской государственный

технический университет, 2024

#### Лабораторная работа №1

#### Исследование применимости законов Зипфа к русскоязычным текстам

Цель работы: В ходе лабораторной работы студент получает практические навыки морфологического анализа текста, применимости законов Зипфа к русскоязычным документам и оптимизации поиска соответствующей информации в Интернет.

Теоретическая часть:

Джордж Кингсли Зипф (George Kingsley Zipf (1902-1950)) – известный профессор-лингвист и филолог Гарвардского университета, эмпирически на основании анализа произвольных англоязычных текстов заметил такую закономерность - «слова с большим количеством букв встречаются в тексте реже коротких слов». Продолжая работы в этом направлении, Зипф пришел к выводу, что все тексты построены по единым правилам. Какой бы язык не использовался, кто бы ни писал текст – внутренняя его структура текста останется неизменной. Понять смысл законов Зипфа проще, если посмотреть на расположение букв на клавиатуре. В каждом языке есть буквы, которые встречаются чаще, они располагаются в середине клавиатуры.

Аналогично в каждом языке есть слова, которые встречаются чаще, чем остальные. Проведя анализ любого текста, можно выделить группу наиболее значимых слов, т.е. слов, при помощи которых можно определить, к какой области относится текст и его основные термины. Суть открытия Зипфа состоит в том, что он эмпирически установил статистические закономерности распределения частоты слов. В то время, когда Зипф сформулировал подмеченные им закономерности распределения частоты слов, законами они не считались. В то время еще не было компьютеров и нельзя было провести точные расчеты, подтверждающие выявленные закономерности.

В последующие годы рядом ученых были проведены исследования, которые подтвердили и уточнили подмеченные зависимости. Их стали называть законами Зипфа.

В настоящее время установлено, что законам Зипфа подчиняются и другие явления. Так Лауреат Нобелевской премии физик Гелл-Манн приводит такие факты:

Если взять статистический справочник со списком крупных городов США, расположенных по порядку убывания населения, и каждому городу сопоставить свой ранг: первый – для самого большого, второй – для следующего и так далее, то анализ этих данных показывает, что существует общий закон, показывающий как убывает население города по мере увеличение ранга, - с достаточной точностью население города обратно пропорционально его рангу: если население первого принять за единицу, то население второго – 1/2, третьего – 1/3 и так далее.

Аналогичным образом можно проанализировать список фирм, расположенных в порядке убывания их оборота. Существует ли здесь какой-либо приблизительный закон, связывающий объём оборота с рангом фирмы? Да, и это тот же закон, что для населения: оборот фирмы обратно пропорционален его рангу. Можно привести и другие примеры из различных областей деятельности человека, для которых применимы законы Зипфа.

3.1 Первый закон Зипфа – «ранг - частота»

Выберем любое слово и посчитаем, сколько раз оно встречается в тексте. Эту величину определим, как частоту вхождения слова и измерим её. Некоторые слова будут иметь одинаковую частоту, то есть входить в текст равное количество раз. Сгруппируем их, взяв только одно значение из каждой группы. Расположим частоты по мере их убывания и пронумеруем. Порядковый номер частоты называется её рангом. Так, наиболее часто встречающиеся слова будут иметь ранг 1, следующие за ними - 2 и т.д.

Определим вероятность вхождения произвольно выбранного слова в текст. Очевидно, она будет равна отношению частоты вхождения этого слова к общему числу слов в тексте. Таким образом, справедливо следующее выражение:

*Вероятность = Частота вхождения слова / Число слов*  (1)

Зипф обнаружил закономерность - если умножить вероятность обнаружения слова в тексте на ранг частоты, то получившаяся величина (С) – константа Зипфа приблизительно постоянна:

*С = (Частота вхождения слова х Ранг частоты) / Число слов*  (2)

После незначительного преобразования зависимости (2) (число слов для анализируемого текста есть некоторая постоянная величина), ее можно представить в виде равносторонней гиперболы y=k/x. Следовательно, по этому закону Зипфа, если самое распространенное слово встречается в тексте, например, 100 раз, то следующее по частоте слово вряд ли встретится 99 раз. Частота вхождения второго по популярности слова, с высокой долей вероятности, окажется равной примерно 50.

Значение константы C в разных языках отличается друг от друга, но внутри одной языковой группы остается неизменным, какой бы текст не был использован для анализа. Так, например, для английских текстов константа Зипфа равна примерно 0,1. Интересно, как выглядят с точки зрения законов Зипфа русские тексты? Они не исключение. Для русского языка константа С равна 0,06-0,07.

3.2 Второй закон Зипфа - "количество - частота "

При рассмотрении первого закона, не учитывался факт, что разные слова могут входить в текст с одинаковой частотой. Зипф установил, что частота и количество слов, входящих в текст с этой частотой, тоже связаны между собой.

Если построить график, отложив по одной оси (оси Х) частоту вхождения слова, а по другой (оси Y) - количество слов в данной частоте, то получившаяся кривая будет сохранять свои параметры для всех без исключения созданных человеком текстов! Как и в предыдущем случае, это утверждение верно в пределах одного языка. Однако и межъязыковые различия невелики. На каком бы языке текст ни был написан, форма кривой Зипфа останется неизменной. Могут немного отличаться лишь коэффициенты, отвечающие за наклон кривой Рисунок 3. Следует заметить, что в логарифмическом масштабе, за исключением нескольких начальных точек, график зависимости количества слов от частоты представляет собой прямую линию.

Законы Зипфа универсальны. В принципе, они применимы не только к текстам. Аналогичный вид имеет, например, зависимость количества городов от числа проживающих в них жителей. Характеристики популярности сайтов в сети Интернет - тоже отвечают законам Зипфа. Не исключено, что в них отражается "человеческое" происхождение объекта. Рассмотрим другой пример. Хорошо известно, что ученые давно бьются над расшифровкой манускриптов Войнича.

Манускрипты Войнича (Voynich Manuscript) — таинственная книга, написанная около 500 лет назад неизвестным автором, на неизвестном языке, с использованием неизвестного алфавита. За всё прошедшее время, рукопись Войнича пытались расшифровать, но без всякого успеха. Она стала Святым Граалем криптографии, но совсем не исключено, что рукопись есть лишь мистификация, бессвязный набор знаков. Книга названа в честь американского книготорговца русского происхождения Вилфрида Войнича (мужа известной писательницы Этель Лилиан Войнич), который приобрёл её в 1912 году. Сейчас она хранится в Библиотеке редких книг Байнеке (Beinecke Rare Book And Manuscript Library) Йельского университета. Однако исследование этих манускриптов на соответствие законам Зипфа доказало: это созданные человеком тексты. Графики, построенные для манускриптов Войнича точно повторили графики текстов на известных языках.

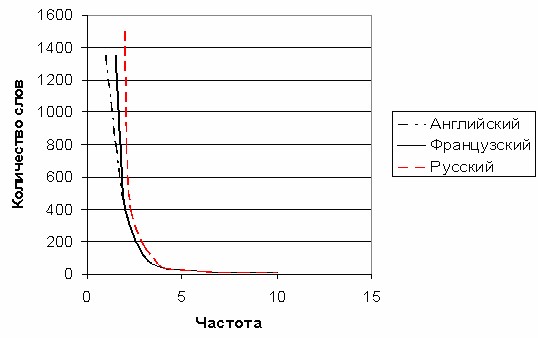


Рисунок 1– Интерпретация второго закона Зипфа для разных языков

Возникает справедливый вопрос: что дают законы Зипфа пользователю Интернета при поиске необходимой информации? Как с их помощью извлечь слова, отражающие смысл текста?

Воспользуемся первым законом Зипфа и построим график зависимости ранга от частоты. Как уже упоминалось, его форма одинакова для всех текстов одного языка Рисунок 2

Исследования показывают, что наиболее значимые слова лежат в средней части представленного рисунка. Это и понятно. Слова, которые попадаются слишком часто, в основном оказываются предлогами, местоимениями, в английском - артиклями и т.п. Редко встречающиеся слова в большинстве случаев, не имеют решающего смыслового значения.

От того, как будет выставлен диапазон значимых слов, зависит многое. При увеличении диапазона - нужные термины потеряются в массе вспомогательных слов; в случае уменьшения диапазона - существует высокая вероятность потери смысловых терминов. Каждая поисковая система решает проблему по-своему, руководствуясь общим объемом текста, специальными словарями и т.п. Проведем эксперимент. Проанализируем абзац любого текста и попытаемся определить список значимых слов.

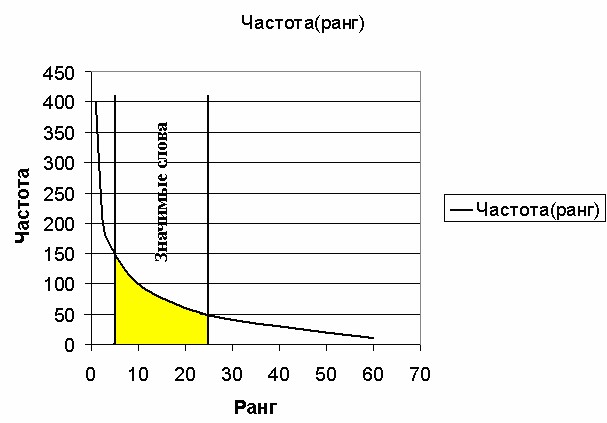


Рисунок 2- Первый закон Зипфа

В качестве примера возьмем один из абзацев текста этих методических указаний. (Например, абзац, начинающийся словами "Законы Зипфа универсальны"). Посмотрим, какие слова попали в область значимых слов, а какие нет. Слово с наибольшей частотой вхождения оказалось предлогом, а слова с меньшей - общими словами.

На Рисунке 3 приведен график частота-ранг этого абзаца. Выделим зону значимых слов. Пусть это будут слова с рангом 2, 3 и частотой 3, 2 соответственно. (Обратите внимание, как смещение или расширение зоны значимых слов влияет на их состав.)

Проанализируем выделенную нами область значимых слов. Не все слова, которые попали в эту область, отражают смысл текста. Смысл очень точно выражают слова: «зипфа», «манускриптов», «войнича», «законам». Запрос типа: "закон\* зипфа" + "манускрипт\* войнича" непременно найдет нам этот документ. Однако в область попали и слова: «на», «не», «для», «, например,», «это». Эти слова являются "шумом", помехой, которая затрудняет правильный выбор. "Шум" можно уменьшить путем предварительного исключения из исследуемого текста некоторых слов. Для этого создается словарь ненужных слов -- стоп-слов (словарь называется стоп-листом). Например, для английского текста стоп-словами станут термины: the, a, an, in, to, of, and, that... и так далее. Для русского текста в стоп-лист могли бы быть включены все предлоги, частицы, личные местоимения и т. п. Наверняка сюда попали бы и слова из нашего "шума": «на», «не», «для», «это». Есть и другие способы повысить точность оценки значимости терминов.

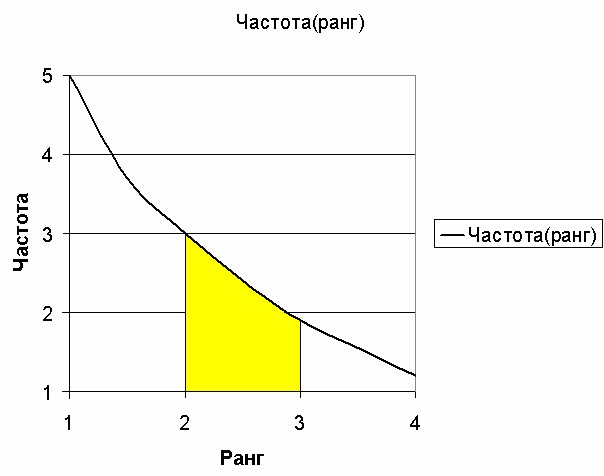


Рисунок 3- Зависимость частоты от ранга для фрагмента выбранного текста

Таким образом, проведенный анализ фрагмента текста показал возможность выделение ключевых слов для последующего его использования при поиске информации в Интернете.

До сих пор рассматривался отдельно взятый документ, не принимался во внимание тот факт, что он входит в базу данных наряду с множеством других документов. Если представить всю базу данных как единый документ, к ней можно будет применить те же законы, что и к единичному документу. Чтобы избавиться от лишних слов и в тоже время поднять рейтинг значимых слов, вводят инверсную частоту термина. Значение этого параметра тем меньше, чем чаще слово встречается в документах базы данных. Вычисляют его по следующей формуле:

*Инверсная частота термина i = log (Количество документов в базе данных / Количество документов с термином i).* (3)

Теперь каждому термину можно присвоить весовой коэффициент, отражающий его значимость:

*“Вес” термина i в документе j =(Частота термина i в документе j)\*(Инверсная частота термина i).*  (4)

Современные способы индексирования не ограничиваются анализом перечисленных параметров текста. Поисковая машина может строить весовые коэффициенты с учетом местоположения термина внутри документа, взаимного расположения терминов, частей речи, морфологических особенностей и т.п.

Изложенный ранее материал позволяет представить, как машина "понимает" суть текста. Теперь необходимо организовать всю совокупность документов так, чтобы можно было легко отыскать в ней нужную информацию. База данных должна взаимодействовать с пользовательским запросом. Запросы могут быть простыми, состоящими из одного слова, и сложными - из нескольких слов, связанных логическими операторами. Простой запрос оправдывает свое название. Пользователь вводит слово, машина ищет его в списке терминов и выдает все связанные с термином ссылки. Структура такой базы данных проста. Взаимодействие со сложными запросами требует более изощренной организации.

Наиболее простой способ представить элементы базы данных в форме, удобной для многовариантного поиска - создать матрицу «документтермин».

Предположим, база данных имеет 8 документов (Д1, Д2, … Д8), в которых содержатся 12 терминов Таблица 1. Если термин входит в документ, в соответствующей клетке таблицы проставляется единица, в противном случае - ноль (в реальной базе поисковой машины все сложнее: помимо прочего, учитываются еще и весовые коэффициенты терминов).

Составим, например, такой запрос: «трубопроводы к сепараторам». Поисковая система обработает запрос: удалит стоп - слова и, возможно, проведет морфологический анализ. Останется два термина: трубопровод и сепаратор. Система будет искать все документы, где встречается хотя бы один из терминов. Посмотрим на матрицу. Пусть указанные в запросе термины есть в документах: Д1, Д2, Д4, Д7, Д8. Они и будут выданы в ответ на запрос. Однако нетрудно заметить, что документы Д4 и Д7 не удовлетворяют нашим запросам - они из области выпечки хлеба и никакого отношения к химико-технологическому оборудованию не имеют. Впрочем, поисковая машина все сделала правильно, ведь, с ее точки зрения, термины трубопровод и сепаратор равноценны.

Таблица 3 Матричное представление «документ-термин»

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Д1 | Д2 | Д3 | Д4 | Д5 | Д6 | Д7 | Д8 |
| Реактор | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| Сепаратор | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| Брожение | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| Трубопровод | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Модель | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| Завод | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| Эстакада | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| Тесто | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| Проектирование | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Мука | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| Инвестиция | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |

Пространственно-векторная модель позволяет получить результат, хорошо согласующийся с запросом. Причем документ может оказаться полезным, даже не имея 100% соответствия. В найденном документе может вовсе не оказаться одного или нескольких слов запроса, но при этом его смысл будет запросу соответствовать. Как достигается такой результат?

Все документы базы данных размещаются в некотором воображаемом многомерном пространстве. Координаты каждого документа зависят от структуры содержащихся в нем терминов (от весовых коэффициентов, положения внутри документа, от расстояния между терминами и т.п.). В результате окажется, что документы с похожим набором терминов разместятся в пространстве ближе друг к другу Рисунок 6.

Предположим, что необходимо найти документы, относящиеся к проектированию трубопроводов, соединенных с сепараторами. Составим тот же запрос: «трубопроводы к сепараторам». Получив его, поисковая машина удалит лишние слова, выделит термины и вычислит вектор запроса в пространстве документов (стрелочка на рисунке). Установив некоторый диапазон соответствия, система выдаст документы, попавшие в заштрихованную область на Рисунке 4. В них может вовсе не оказаться некоторых слов запроса, однако документы останутся в результирующей выборке, и они будут достаточно релевантными. Релева́нтность (англ. *relevant*) — применительно к результатам работы поисковой машины — степень соответствия запроса и найденного, уместность результата. Это субъективное понятие, поскольку результаты поиска, уместные для одного пользователя, могут быть совершенно неприемлемыми для другого.

Термины, относящиеся к хлебопекарне, будут группироваться в другой точке пространства, и запрос их не затронет. Как видите, "уравниловку" терминов удалось преодолеть. В пространственно-векторной модели термины взаимодействуют друг с другом, что повышает релевантность документов. Понятно, что пространственно-векторная модель лучше воспринимает запросы, составленные на естественном языке, чем матричная.

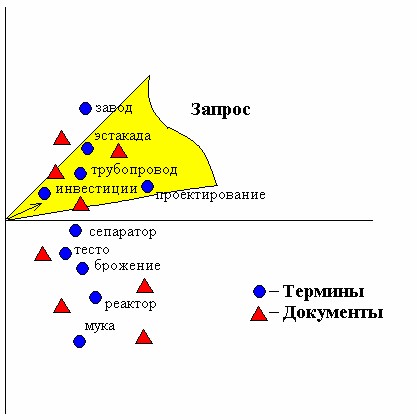


Рисунок 4- Пространственно-векторное представление документов

К сожалению, догадаться, по какой схеме работают поисковые системы Интернета, трудно. Как правило, создатели каждой системы держат ее в секрете. В методическом пособии изложены лишь основы работы поисковой системы. В реальности механизм индексации и структура базы данных значительно сложнее. Однако полученных знаний достаточно, чтобы попытаться выработать стратегию поиска информации в Интернете.

Известно, как поисковая машина выделяет ключевые слова. Воспользуемся этим знанием, чтобы сформировать эффективный запрос.

Изначально сформулируем некоторые исходные предпосылки. Для организации поиска требуется текст-источник, описывающий изучаемую проблему.

Таким образом, необходимо найти документы максимально близкие по смысловому содержанию к выбранному эталонному документу.

Возникает вопрос - как получить документ, на основании которого будет построена стратегия поиска? Предполагается, что сама задача поиска не могла возникнуть из ничего, где-то непременно должна существовать информация, побудившая интерес к проблеме. Может быть, это журнальная статья, книга, веб - страница и т.п. Именно эту информацию следует проанализировать, упорядочить ее и привести в форму, удобную для поиска. Если описание задачи не существует в виде какого-либо источника информации, а она присутствует только в голове конкретного человека, то следует изложить ее в виде сочинения, тогда его текст и станет текстом источником. Если бы нам удалось препарировать текст-источник так же, как это делает поисковая машина, то полученный результат имел бы наивысшее значение релевантности.

В общем случае стратегия эффективного поиска информации в Интернет может быть представлена так:

1. Подбираем текст-источник. Чем четче описание проблемы в тексте-источнике, тем качественнее и точнее окажется результат. Удаляем из текста стоп - слова (их можно просто вычеркивать).
2. Вычисляем частоту вхождения каждого термина. Причем делаем это без учета морфологии слов. Так, слова завод и заводах будут разными терминами. Все буквы считаем строчными.
3. Выписываем на отдельный лист термины в порядке убывания их частоты вхождения (первыми должны идти те, которые встречаются чаще).
4. Строим графическую зависимость между рангом и частотой, как это показано на рисунке 3. Выбираем диапазон частот. Выбор диапазона субъективен. Следует ориентироваться на конкретный смысл текста. Необходимость выбора диапазона вручную не должна смущать, ведь выбираются термины не из текста, а из построенного по определенному правилу упорядоченного списка.
5. Из выбранного диапазона выписываем термины. Здесь может оказаться довольно много слов. Все их применить вряд ли удастся. Достаточно взять 10-20 терминов. Их следует выбирать, руководствуясь, в первую очередь, здравым смыслом. Причем не стоит ограничиваться только характерными терминами, даже если они кажутся наиболее удачными. В список должны попасть и общие слова (их лучше выбирать из средней части диапазона).
6. Составляем запрос, располагая отобранные слова в порядке их следования в списке терминов. Запрос должен пониматься машиной как слова, связанные логическим оператором «ИЛИ». Это очень важное требование. Чтобы результат не исказился, следует изучить особенности синтаксиса запросов конкретной поисковой системы.
7. Отправляем запрос поисковой системе.

В ответ мы можем получить просто громадное число ссылок. Если поисковая машина ранжирует результаты (а это еще одно условие), на первых страницах ответа окажутся практически стопроцентно релевантные документы. Самое любопытное, что документ - источник запроса (если его аналог существует в Интернете) вовсе не обязательно будет возглавлять список, но обязательно должен присутствовать среди найденных документов.

Описание лабораторной работы:

Лабораторную работу необходимо выполнять в следующем порядке.

Первоначально требуется ознакомиться с методическими указаниями. После изучения теоретического материала необходимо ответить на контрольные вопросы, и получить у преподавателя допуск на выполнение работы. Разрешается объединение нескольких студентов во временную группу, которая получает общее задание.

После получения допуска на закрепленном за студентом рабочем месте (компьютере), необходимо выполнить индивидуальное задание, используя указанные в нем поисковые системы.

Непосредственно лабораторная работа проводится в два этапа. Первоначально проводится анализ русскоязычного текста, и оценивается применимость к нему законов Зипфа. После этого студент должен выполнить поиск документов в Интернет, обращая внимание на их релевантность и пертине́нтность. Пертине́нтность (англ. pertinent) - соотношение объема полезной информации к общему объему полученной информации.

Для исследования применимости законов Зипфа к русскоязычным текстам следует использовать текстовый редактор MS Word.

Запустите приложение MS Word и найдите в Интернете сайт научной конференции со сборником статей по направлению вашего обучения. Выберете одну из статей конференции. Последняя цифра страницы соответствует последней цифре зачетки.

Например, номер зачетки [172472**1**](https://edu.donstu.ru/Dek/?mode=stud&f=stud&id=119381), следовательно, подойдут статьи на страницах: 91, 111 и т.д

Перед Вами появится фрагменты статьей. Нужно проанализировать. При этом требуется выявить слова с рангом 1,2,3,4,5 и построить график применимости законов Зипфа к фрагменту статьи. Для сравнения полученных результатов следует использовать значение константы C. (см. теоретическую часть). Анализ проводится без учета морфологических особенностей текста. Каждое слово является уникальным элементом статьи. При анализе текста необходимо исключить «шум» (стопслова). Далее нужно составить запрос в различных поисковых системах.

Зафиксируйте результаты и проведите сравнительный анализ работы поисковых систем (релевантность и пертинен́тность).

Проанализируйте влияние морфологического анализа текста на применимость законов Зипфа и на результаты поиска. Для этого нужно рассмотреть список терминов, содержащийся в задании, и провести его морфологический разбор. Постройте соответствующие графические зависимости, сравните их с предыдущими, сделайте выводы.

При выполнении лабораторной работы для автоматизации процесса анализа текста-первоисточника рекомендуется использовать программу электронных таблиц MS Excel, а также сервис, расположенный по адресу <http://www.shipbottle.ru/ir/>.

5. Оформление результатов работы

Результаты работы представляются студентом индивидуально в виде отчета подготовленного с использованием пакета офисных программ. Отчет должен содержать таблицу с частотными характеристиками текстапервоисточника, графические зависимости применимости законов Зипфа без проведения морфологического анализа так и с ним. Обязательным элементом отчета является анализ результатов, полученных с помощью различных поисковых систем. Подготовленный отчет направляется на проверку преподавателю с помощью средств электронного документооборота

В случае возврата отчета преподавателем студент должен внести необходимые коррективы, новый вариант представляется как версия первоначального документа.

В случае успешной защиты студентом лабораторной работы преподаватель переносит файл отчета в электронный архив кафедры.

Содержание отчета и правила его оформления представлены на информационных стендах кафедры и соответствуют требованиям стандарта предприятия.

Контрольные вопросы

1. Сформулируйте первый закон Зипфа.
2. Сформулируйте второй закон Зипфа
3. Определите области применимости законов Зипфа, приведите примеры.
4. Приведите последовательность операций необходимых для построения графической зависимости, соответствующей первому закону Зипфа для произвольного текста
5. Приведите последовательность операций необходимых для построения графической зависимости, соответствующей второму закону Зипфа для произвольного текста
6. Что такое весовые коэффициенты термина?
7. Что такое матричное представление документов в базе данных?
8. В чем преимущество пространственно-векторного представления базы данных?
9. Что такое оптимальная стратегия поиска и необходимые предпосылки для ее реализации?
10. Что такое релевантность?
11. Что такое пертине́нтность?
12. В чем сходство и отличие релевантности и пертинентности?

#### Лабораторная работа № 2

#### Информационный поиск в интернете

Цель работы: Знакомство с особенностями языков запросов поисковых систем yandex.ru, google.com и rambler.ru, использование режима расширенного поиска на yandex.ru.

Информационный поиск − выделение из общего объема информации части, удовлетворяющей условиям поиска.

Информационный поиск может заключаться как в поиске информации во всем объеме документов, так и поиск самих документов. Объектом поиска могут быть текст, изображение, видео и т.д., областью поиска – реляционные базы данных, интернет, локальные интранет-системы и т.д.

В связи с экспоненциальным ростом объемов неструктурированной информации, размещаемой в интернете в настоящее время и связанное с этим увеличение сложности выделения из общего потока желаемых данных нельзя недооценивать важность информационного поиска.

Наиболее распространенные системы поиска в России – yandex.ru (61,3%), rambler.ru (1,9 %) и др., во всем мире – google.com (83,87%), yahoo.com (6,20%) и др. Наверняка каждый уже хотя бы раз пользовался одной из поисковых систем, но мало кто знает, что помимо обычного поиска существует возможность использования языка запросов и режима расширенного поиска для повышения качества находимой информации Таблица 2.

Таблица 1- Язык запросов поисковой системы yandex.ru

|  |  |
| --- | --- |
| **Пример** | **Описание** |
| «К нам на утренний рассол» | Слова идут подряд в точной форме |
| «Прибыл \* посол» | Пропущено слов в цитате |
| Полгорбушки & мосол | Слова в пределах одного предложения |
| Снаряжайся && добудь | Слова в пределах одного документа |
| Технический прогресс + антирес | Поиск документов, в которых обязательно встречаются определённое слово |
| Глухаря | куропатку | кого-нибудь | Поиск любого из слов |
| Я должон /2 казнить | Расстояние в пределах двух слов в любую сторону (то есть между заданными словами может встречаться одно слово) |
| Государственное дело && /3 улавливаешь нить | Расстояние в 3 предложения в любую сторону |
| Нешто я ~~ пойму | Исключение слова пойму из поиска |
| При моем /+2 уму | Расстояние в пределах двух слов в прямом порядке |
| Чай ~ лаптем | Поиск предложения, где слово чай встречается без слова лаптем |
| Щи /(-1+2) хлебаю | Расстояние от одного слова в обратном порядке до двух слов в прямом |
| !тихий !Дон | Поиск документа, в котором есть слово тихий, начинающееся со строчной или заглавной буквы, или слово Дон, начинающееся с заглавной |
| Получается && (+на | !мне) | Скобки формируют группы в сложных запросах |
| !!политика | Словарная форма слова |
| Title:(в стране) | Поиск по заголовкам документов |
| url:ptici.narod.ru/ptici/kuropatka.htm | Поиск по URL |
| Беспременно inurl:vojne | Поиск с учетом фрагмента URL |
| host:lib.ru | Поиск по хосту |
| rhost:ru.lib.\* | Поиск по хосту в обратной записи |
| site:http://www.lib.ru/PXESY/FILATOW | Поиск по всем поддоменам и страницам заданного сайта |
| mine:pdf | Поиск по одному типу файлов |
| Lang:en | Поиск с ограничением по языку |
| domain:ru | Поиск с ограничением по домену |
| data:200712\* | Поиск с ограничением по дате |

Таблица 2 - Язык запросов google

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Пример** | | | **Описание** | | | | |
| "К нам на утренний рассол" | | | Слова форме | идут | подряд | в | точной |
| глухаря **OR**  нибудь | куропатку **OR** | кого- | Поиск любого из слов | | | | |
| нешто я **-** пойму | | | Исключает страницы, содержащие  слово или фразу. | | | | |
| технический прогресс + антирес | | | Включение общих слов или для  Отключения морфологического поиска | | | | |
| полгорбушки ~ мосол | | | Не только сами слова из запроса,  но и их синонимы. | | | | |
| linux site:com  linux site:(com OR org ) linux site:[www.linux.org/dist/](http://www.linux.org/dist/) | | | Страницы, принадлежащие определенному домену и всем его  поддоменам. | | | | |
| статистика filetype:pdf. | | | Поиск в файлах определенного  типа. | | | | |
| link:[www.shtogrin.com](http://www.shtogrin.com/) | | | Поиск страниц, содержащих  ссылки на указанную страницу. | | | | |
| linux lang:en.  linux lang:ru. | | | Поиск страниц на определенном  языке. | | | | |

Таблица 3 – Язык запросов поисковой системы rambler.ru

|  |  |
| --- | --- |
| **Пример** | **Описание** |
| **1** | **2** |
| технический прогресс + антирес | Используется для включения общих слов |
| "К нам на утренний рассол" | Слова идут подряд в точной форме |
| **1** | **2** |
| чай NOT лаптем | Запросу отвечают документы, удовлетворяющие левой части запроса и не удовлетворяющие правой. |
| &<, &<< &>, &>> | Увеличение и уменьшение в  тексте расстояния между найденными словами. |
| снаряжайся && добудь | Оба слова в пределах одного  документа |
| я должон &&& казнить | Оба слова в пределах одного сайта |
| щи | хлебаю | Поиск любого из слов |
| щи || хлебаю | Поиск любого из слов или обоих  вместе |
| получается && (на мне) | Скобки формируют группы в  сложных запросах |

Ход выполнения работы:

1. Ознакомьтесь с описанием особенностей языков запросов поисковых систем yandex.ru, rambler.ru и google.com (Таблица 2,Таблица 3,Таблица 4).
2. Составьте и заполните следующую таблицу с использованием языков запросов для своего варианта задания.

|  |  |  |  |
| --- | --- | --- | --- |
| **Функция** | **yandex.ru** | **rambler.ru** | **google.com** |
| Поиск слов, которые идут подряд в точной  форме |  |  |  |
| Поиск любого из слов |  |  |  |
| Исключение из результатов страницы, содержащие слово или  фразу. |  |  |  |

1. Для всех функций таблицы, составленной в пункте 2, произведите поиск на поисковых системах yandex.ru, rambler.ru и google.com с использованием языков запросов.
2. На основе результатов поиска сформулируйте пять предложений, объясняющих словосочетание, по которому производился поиск с указанием сайтов, на которых была найдена информация.
3. Сравните качество поиска с использованием систем yandex.ru, rambler.ru

и google.com по следующим критериям:

* + общее количество совпадений,
  + для первых трех страниц с результатами поиска произвести расчет отношения количества релевантных совпадений к общему количеству совпадений.

1. Повторите поиск только в файлах c расширением .pdf на yandex.ru с использованием языка запросов. Сравните результаты поиска с результатами, полученными при выполнении пункта 5. Для первых трех страниц с результатами поиска вычислите отношение количества совпадений для файлов pdf к общему количеству совпадений из пункта 5.
2. Перейдите в режим расширенного поиска на yandex.ru ([http://yandex.ru/search/advanced?&lr=10407)](http://yandex.ru/search/advanced?&amp;lr=10407))
3. Повторите поиск, произведенный в задании 3, на сайте<http://www.wikipedia.ru/> с использованием поисковой системы yandex.ru. Воспользуйтесь для этого режимом расширенного поиска.

Структура отчета:

1. Цель работы
2. Задание
3. Пять предложений из пункта 4.
4. Ход выполнения (заполненная таблица из пункта 2, общее количество найденных документов (пункт 5), отношение количества релевантных совпадений к общему количеству найденных документов для первых трех страниц с результатами поиска (пункт 5), отношение количества совпадений для файлов pdf к общему количеству совпадений для первых трех страниз с результатами (пункт 6).
5. Выводы, содержащие анализ:
   * результатов поиска с помощью различных поисковых систем,
   * количества данных представляемых в формате pdf,
   * режима расширенного поиска на yandex.ru.

Индивидуальные задания выбираем в соответсвии с списочным номером от 1-15 для 16 нумерация с начала.

|  |  |
| --- | --- |
| Номер варианта | Словосочетание для поиска |
| 1 | ЦИФРОВИЗАЦИЯ В АГРОПРОМЫШЛЕННОМ КОМПЛЕКСЕ |
| 2 | МОДУЛЬНЫЕ ТЕХНОЛОГИЧЕСКИЕ ЛИНИИ |
| 3 | НЕРАЗРУШАЮЩИЙ КОНТРОЛЬ МЕТАЛЛИЧЕСКИХ ДЕТАЛЕЙ |
| 4 | КОМПЕТЕНТНОСТНО-ОРИЕНТИРОВАННЫЙ ПОДХОД |
| 5 | РЫНОЧНАЯ СТОИМОСТЬ ИМУЩЕСТВА ПРЕДПРИЯТИЙ |
| 6 | ОЦЕНКА ЭФФЕКТИВНОСТИ СИСТЕМ ОХЛАЖДЕНИЯ |
| 7 | ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ТЕХНИЧЕСКОЙ ЭКСПЛУАТАЦИИ |
| 8 | РЕГЛАМЕНТ ПРОЦЕССА ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ |
| 9 | УСОВЕРШЕНСТВОВАНИЕ СИСТЕМЫ КОНДИЦИОНИРОВАНИЯ ВОЗДУХА |
| 10 | ПОВЫШЕНИЕ ЭКСПЛУАТАЦИОННОЙ НАДЕЖНОСТИ ТОПЛИВНОЙ СИСТЕМЫ |
| 11 | АНАЛИЗ ПРОБЛЕМЫ РАЗРАБОТКИ ПРОГРАММНОГО |
| 12 | АНАЛИЗ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ИНФОРМАЦИИ |
| 13 | ПЕРСПЕКТИВЫ РАЗВИТИЯ АНТЕННЫХ СИСТЕМ  РАДИОЛОКАТОРОВ, |
| 14 | СИСТЕМА РЕЗЕРВИРОВАНИЯ ИСТОЧНИКОВ ПИТАНИЯ |
| 15 | КОНСТРУКТИВНЫЕ  НЕДОСТАТКИ, ВЫЯВЛЕННЫЕ В ПРОЦЕССЕ ЭКСПЛУАТАЦИИ |

Контрольные Вопросы:

1. Понятие информационного поиска.

2. Библиотека как информационно-поисковая система. Поисковые ресурсы библиотек.

3. Методика информационного поиска

4. Информационный поиск в Интернете.

5. Организация эффективного поиска в Интернет

Лабораторная работа №3

Основы обработки персональных данных в информационных системах

Цель работы: Изучить законодательные основы обработки персональных данных в информационных системах в соответствии с законодательством РФ.

Краткие теоретические сведения

Совет Европы опубликовал конвенцию о защите личности при обращении с персональными данными. Цель Конвенции - «гарантировать на территории каждой страны каждому частному лицу, независимо от его национальности и места проживания, соблюдение его прав и основных свобод, и особенно его права на личную жизнь в аспекте автоматизированной обработки данных личного характера (статья 1).

Согласно части 1 статьи 3 «Стороны обязуются применять настоящую Конвенцию в отношении автоматизированных картотек и для автоматизированной обработки данных личного характера в общественном и частном секторах».

Как необходимый шаг для вступления в ВТО, Россия подписала Конвенцию в ноябре 2001 года. Таким образом, только спустя 20 лет со дня публикации Конвенции в России началось движение в сторону создания правовых основ обработки персональных данных. Конвенция была принята формально, но по факту не действовала из-за отсутствия российских нормативных актов.

Вновь законодатели вернулись к этому вопросу в 2005 году, когда был принят Федеральный закон от 19 декабря 2005 года №160-ФЗ «О ратификации Конвенции Совета Европы о защите физических лиц при автоматизированной обработке персональных данных».

Знаковым стал 2006 год, когда в целях наведения порядка в сфере интеллектуальной собственности и защиты персональных данных, в том числе как выполнение одного из условий вступления России в ВТО, были приняты два Федеральных закона.

Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных»

Федеральный закон от 27 июля 2006 г. N 149-ФЗ «Об информации, информационных технологиях и о защите информации»

Основополагающие документы, регламентирующие сохранность персональных данных

Федеральный закон "О персональных данных" от 27.07.2006 N 152-ФЗ- Основной закон в области защиты персональных данных, содержит основные сведения о персональных данных так в законе дано определение принципов обработки персональных данных, определение различных категорий персональных данных определение обязанности оператора и прав субъекта персональных данных.

Постановление Правительства РФ от 01.11.2012 № 1119 "Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных" В постановление правительства определены 4 ровня защищенности информационных систем персональных данных, определен состав организационных и технических мер применимых для защиты информации в информационных системах различных уровней защищенности.

Основополагающие документы в области информационной безопасности приведены на рисунке 5.

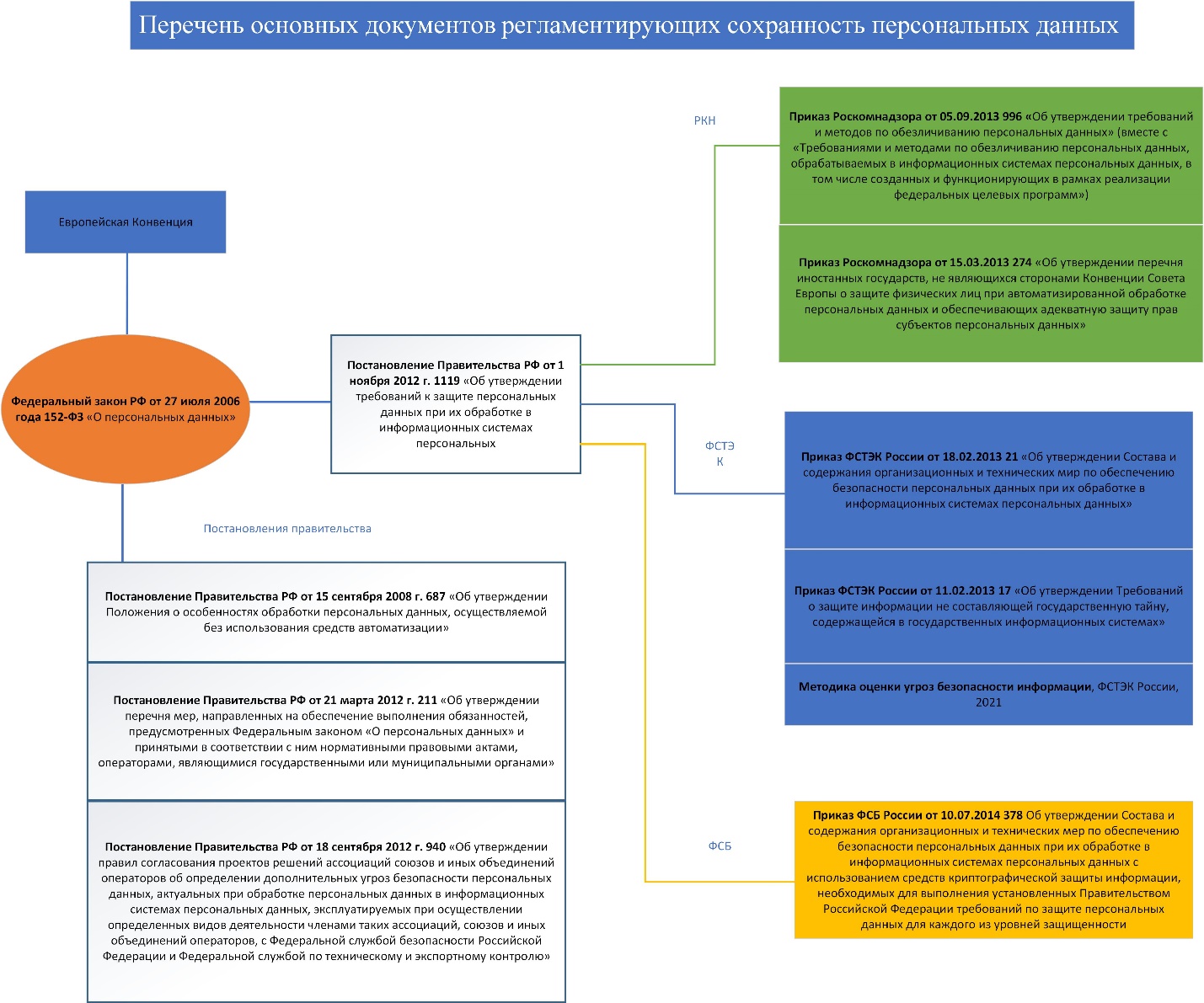


Рисунок 5 - Структура основополагающих документов в области персональных данных

Приказ №21 от 18 февраля 2013 г. «Об утверждении состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных».

Приказ №17 от 1.02.2013г. «Об утверждении требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах».

Приказы определяют состав и содержание организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных. Документы являются ключевыми для обеспечения безопасности персональных данных, что является одной из основных обязанностей оператора. Приказ №21 применяется юридическими лицами с частной формой собственности, а 17 приказ разработан для государственных органов.

Состав и содержание требований организационных и технических мер по обеспечению безопасности ПДн при их обработке в ИСПДн устанавливает уполномоченный в области технической защиты информации ФСТЭК в Приказе №21.

Методический документ, Меры защиты информации в государственных информационных системах ФСТЭК, 2014г. Данный «Методический документ», как следует из названия, был разработан в целях детализации организационных и технических мер защиты информации, применяемые в государственных информационных системах в соответствии с Требованиями, утвержденными в 17-ом Приказе ФСТЭК

Приказ ФСБ России от 10 июля 2014 г. N 378 "Об утверждении Состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных с использованием средств криптографической защиты информации, необходимых для выполнения установленных Правительством Российской Федерации требований к защите персональных данных для каждого из уровней защищенности"

Вопрос применения криптографических средств защиты регулирует ФСБ. По аналогии с Приказом ФСТЭК №21, состав и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных с использованием средств криптографической защиты информации, необходимых для каждого из уровней защищенности, определен в Приказе ФСБ России от 10.07.2014 N 378.

Целесообразность применения СКЗИ можно найти в документе "Методические рекомендации по разработке нормативных правовых актов, определяющих угрозы безопасности персональных данных, актуальные при обработке персональных данных в информационных системах персональных данных, эксплуатируемых при осуществлении соответствующих видов деятельности" (утв. ФСБ России 31.03.2015 N 149/7/2/6-432).

Задание

1. Изучить теоретический материал
2. Зайти на сайт роскомнадзора в раздел посвященный защите персональных данных <https://pd.rkn.gov.ru/>, изучить содержание портала.
3. Зайти в раздел «Реестр операторов» <https://pd.rkn.gov.ru/operators-registry/operators-list/> найти сведенья о следующих организациях:

* Открытое акционерное общество "Ростсельмаш" 6166010558
* Общество с ограниченной ответственностью "Чистый город - Юг"
* Государственное бюджетное учреждение здравоохранения Республики Карелия "Больница скорой медицинской помощи" ИНН 1001117532

1. Определить и проанализировать:

* Цель обработки персональных данных
* ФИО физического лица или наименование юридического лица, ответственных за обработку персональных данных
* Срок или условие прекращения обработки персональных данных
* Список информационных систем и их параметры
* Список информационных систем и их параметры

1. Зайти в раздел «Документы» https://pd.rkn.gov.ru/operators-registry/operators-registry-documents/ изучить пример заполнения основных типов документов по регистрации оператора в системе.
2. Зайти в раздел согласие на обработку ПД разрешенных к распространению. https://pd.rkn.gov.ru/soglasiya/ На основе шаблонов создать типовые документы, «согласие на обработку» персональных данных.
3. Оформить отчет сделать выводы по работе.

Контрольные вопросы

1. Что такое персональные данные?
2. Назовите основные категории персональных данных ?
3. Что такое «Информационная система персональных данных”?
4. Кто является «субъектом персональных данных»?
5. Кто является «оператором персональных данных»?
6. Какие сферы ответственности регуляторов в области защиты персональных данных ( ФСБ, Роскомнадзор, ФСТЭК )?

Лабораторная работа №4

Основы обработки персональных данных в информационных системах

Цель работы: Изучить методы обезличивания персональных данных в информационных системах.

Краткие теоретические сведения

В соответствии со статьей 3 Федерального закона от 27 июля 2006 г. N 152-ФЗ "О персональных данных" (Собрание законодательства Российской Федерации 2006, N 31 (ч. I), ст. 3451; 2009, N 48, ст. 5716; N 52 (ч. I), ст. 6439; 2010, N 27, ст. 3407; N 31, ст. 4173, ст. 4196; N 49, ст. 6409; N 52 (ч. I), ст. 6974; 2011, N 23, ст. 3263; N 31, ст. 4701; 2013, N 14, ст. 1651) под обезличиванием персональных данных понимаются действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту персональных данных.

Обезличивание персональных данных должно обеспечивать не только защиту от несанкционированного использования, но и возможность их обработки. Для этого обезличенные данные должны обладать свойствами, сохраняющими основные характеристики обезличиваемых персональных данных.

К свойствам обезличенных данных относятся:

полнота (сохранение всей информации о конкретных субъектах или группах субъектов, которая имелась до обезличивания);

структурированность (сохранение структурных связей между обезличенными данными конкретного субъекта или группы субъектов, соответствующих связям, имеющимся до обезличивания);

релевантность (возможность обработки запросов по обработке персональных данных и получения ответов в одинаковой семантической форме);

семантическая целостность (сохранение семантики персональных данных при их обезличивании);

применимость (возможность решения задач обработки персональных данных, стоящих перед оператором, осуществляющим обезличивание персональных данных, обрабатываемых в информационных системах персональных данных, в том числе созданных и функционирующих в рамках реализации федеральных целевых программ (далее - оператор, операторы), без предварительного деобезличивания всего объема записей о субъектах);

анонимность (невозможность однозначной идентификации субъектов данных, полученных в результате обезличивания, без применения дополнительной информации).

К характеристикам (свойствам) методов обезличивания персональных данных (далее - методы обезличивания), определяющим возможность обеспечения заданных свойств обезличенных данных, относятся:

обратимость (возможность преобразования, обратного обезличиванию (деобезличивание), которое позволит привести обезличенные данные к исходному виду, позволяющему определить принадлежность персональных данных конкретному субъекту, устранить анонимность);

вариативность (возможность внесения изменений в параметры метода и его дальнейшего применения без предварительного деобезличивания массива данных);

изменяемость (возможность внесения изменений (дополнений) в массив обезличенных данных без предварительного деобезличивания);

стойкость (стойкость метода к атакам на идентификацию субъекта персональных данных);

возможность косвенного деобезличивания (возможность проведения деобезличивания с использованием информации других операторов);

совместимость (возможность интеграции персональных данных, обезличенных различными методами);

параметрический объем (объем дополнительной (служебной) информации, необходимой для реализации метода обезличивания и деобезличивания);

возможность оценки качества данных (возможность проведения контроля качества обезличенных данных и соответствия применяемых процедур обезличивания установленным для них требованиям).

Задание

1. Изучить теоретический материал
2. Зайти на сайт роскомнадзора в раздел, посвященный обезличиванию персональных данных https://rkn.gov.ru/docs/Xerox\_Phaser\_3200MFP\_20131216122746.pdf.
3. Применить методики обезличивания используя учебный файл xls, оценить их эффективность и применимость в практической работе.
4. Автоматизируйте любой из способов обезличивания используя те программные инструменты, которыми вы владеете т.е. в результате продукта вы должны импортировать не обезличенные персональные и выгрузить обезличенные.
5. Оформить отчет сделать выводы по работе.

Контрольные вопросы

1. Что такое обезличивание персональных данных?
2. Чем обезличивание персональных данных отличается от шифрования?
3. Объясните суть обезличивания методом перемешивания?
4. В чем заключаются использование метода декомпозиции?
5. В чем заключаются использование метода идентификаторов?

## Лабораторная работа № 5

## Характеристика объектов интеллектуальной собственности и анализ объема патентных прав

Цель работы: Изучить интеллектуальную собственность, подразделяющуюся на две категории: промышленная собственность и авторское право.

**Промышленная собственность** - товарные знаки, изобретения, промышленные образцы, полезные модели.

**Авторское право** - литературные и художественные произведения. К объектам авторского права относятся также компьютерные программы и базы данных.

**Товарный знак** (торговая марка, логотип, слоган) охраняется в силу его официальной регистрации. Регистрация товарного знака является наиболее эффективным способом защиты товаров/услуг от недобросовестной конкуренции. Срок действия регистрации товарного знака составляет 10 лет и предполагает возможность ее неоднократного продления. Законный владелец товарного знака обладает исключительным правом его использования и без его разрешения пользование этим товарным знаком запрещено.

Наименование **домена** выполняет ту же функцию, что и товарный знак, и подобно ему может быть отнесено к объектам интеллектуальной собственности. Наиболее эффективной охраной доменного имени является регистрация тождественного товарного знака.

**Изобретения, полезные модели и промышленные образцы** защищаются патентом. Патент действителен в течение определенного срока и только на территории той страны, где он выдан, либо на территории ряда стран, между которыми существуют соответствующие договоренности.

Объекты авторского права, в отличие от патентуемых объектов, охраняются в течении всей жизни автора, а имущественное право на объект действует еще 70 лет после смерти и принадлежит наследникам. Авторское право на произведение науки, литературы и искусства возникает в силу факта его создания. Для возникновения и осуществления авторского права не требуется регистрации произведения.

**Компьютерные программы и базы данных** относятся к категории официально регистрируемых объектов авторского права. Несмотря на то, что авторское право на эти объекты возникает в силу их создания, для обеспечения защиты исключительные права подтверждаются официальной регистрацией. Часто практикуют дополнительную защиту программ и баз данных, регистрируя название и/или логотип программы в качестве товарного знака.

Подробнее см.: <http://www.sciteclibrary.ru/express-patent/voprosi.html>, <http://www.lintes.ru/>

### Задание

1. Реализация совместного телекоммуникационного проекта – Презентации «Объем патентных прав \_\_\_\_\_\_\_\_ (объекта интеллектуальной собственности)»

### Порядок выполнения (лабораторная работа обязательно выполняется группой студентов):

1. Запустить любой браузер Интернет
2. Найти полное определение любого объекта интеллектуальной собственности (на усмотрение студента):
   1. изобретение устройства;
   2. полезная модель;
   3. промышленный образец;
   4. фирменное наименование;
   5. товарный знак, знак обслуживания, место происхождения товара;
3. Выбранный объект описать согласно плану:
   1. понятие и признаки описываемого объекта
   2. общие правила оформления заявки на объект
   3. права собственника описываемого объекта
   4. защита прав собственников описываемого объекта (признаки нарушения прав, ветви права: уголовное, административное, способы защиты)
4. Создать телекоммуникационный проект (например? с помощью Google) «Презентация» по теме «Объем патентных прав \_\_\_\_\_\_\_\_ (такого-то объекта интеллектуальной собственности)»
   1. Создать аккаунт Google (<https://www.google.ru>)
   2. Изучить возможности инструмента «Документы Google» (<http://www.google.com/google-d-s/intl/ru/tour1.html>)
   3. Создать совместную презентацию
   4. Этапы работы над презентацией (от различных пользователей: автора и соавторов) в виде скриншотов поместить в отчет о выполнении лабораторной работы
   5. Добавить в презентацию образцы различного контента, защищенного авторским правом добавить описание, и выдержку из лицензии.

Инструкция по защите авторства и интеллектуальной собственности материалов, выставленных в информационном пространстве Интернет.

<https://sites.google.com/site/avtbfa/instrukcia-po-zasite-avtorstva-i-intellektualnoj-sobstvennosti-materialov-vystavlennyh-v-informacionnom-prostranstve-internet>

Лабораторная работа № 6

Государственная регистрация программы для электронных вычислительных машин или базы данных в Роспатент Федеральной службе по интеллектуальной собственности

Цель работы: Изучить процедуры регистрации программы для электронных вычислительных машин. Приобрести навыки создания заявки на регистрацию

Краткие теоретические сведения.

Программой для ЭВМ является представленная в объективной форме совокупность данных и команд, предназначенных для функционирования ЭВМ и других компьютерных устройств в целях получения определенного результата, включая подготовительные материалы, полученные в ходе разработки программы для ЭВМ, и порождаемые ею аудиовизуальные отображения.

Базой данных является представленная в объективной форме совокупность самостоятельных материалов (статей, расчетов, нормативных актов, судебных решений и иных подобных материалов), систематизированных таким образом, чтобы эти материалы могли быть найдены и обработаны с помощью электронной вычислительной машины (ЭВМ)

Правообладатель в течение срока действия исключительного права на программу для ЭВМ или на базу данных может по своему желанию зарегистрировать такую программу или такую базу данных в Роспатент См. Рисунок 6

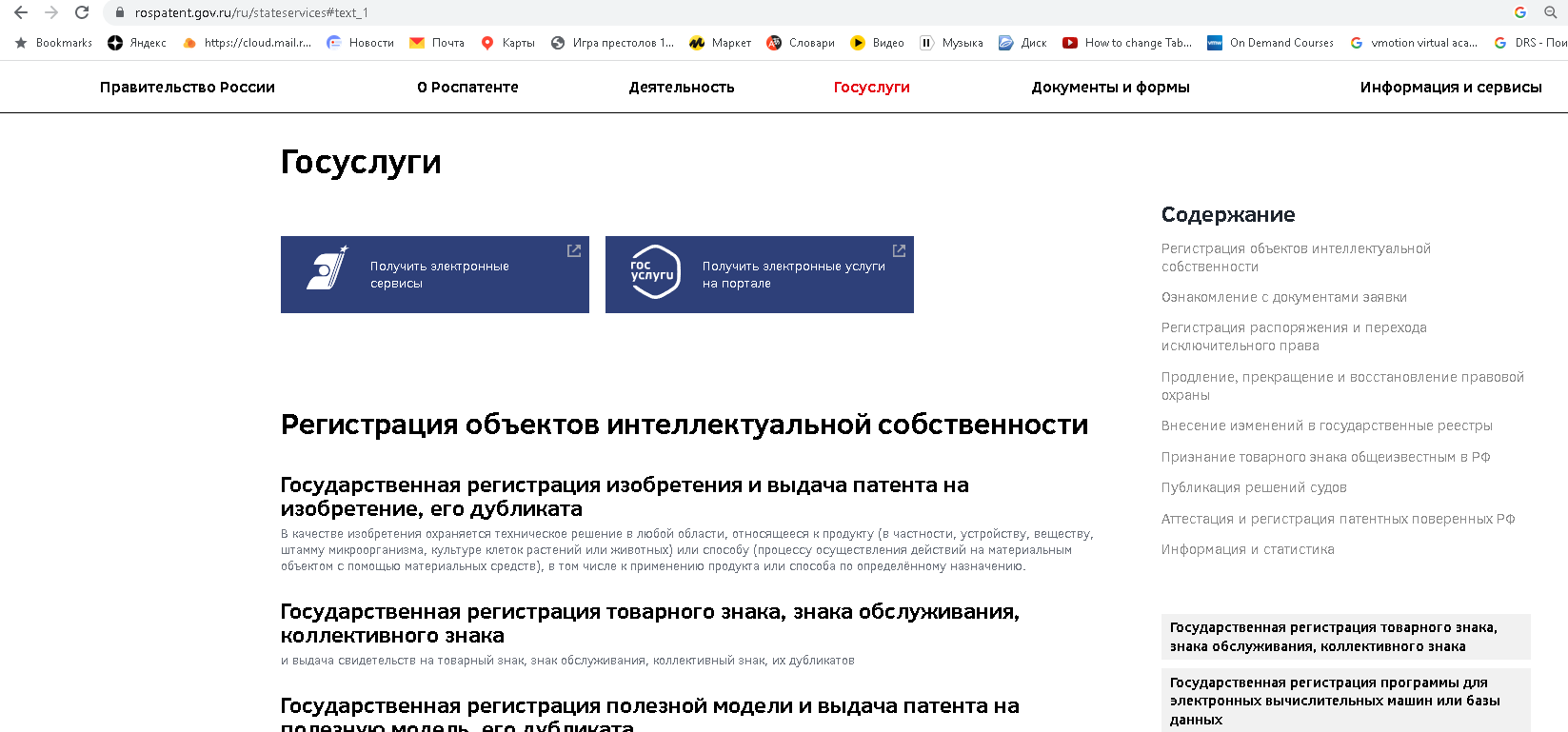


Рисунок 6- Сайт Роспатента

Авторские права на все виды программ для ЭВМ (в том числе на операционные системы и программные комплексы), которые могут быть выражены на любом языке и в любой форме, включая исходный текст и объектный код, охраняются так же, как авторские права на произведения литературы. Программой для ЭВМ является представленная в объективной форме совокупность данных и команд, предназначенных для функционирования ЭВМ и других компьютерных устройств в целях получения определенного результата, включая подготовительные материалы, полученные в ходе разработки программы для ЭВМ, и порождаемые ею аудиовизуальные отображения. При этом авторские права не распространяются на идеи, концепции, принципы, методы, процессы, системы, способы, решения технических, организационных или иных задач, открытия, факты, языки программирования. Базой данных является представленная в объективной форме совокупность самостоятельных материалов (статей, расчетов, нормативных актов, судебных решений и иных подобных материалов), систематизированных таким образом, чтобы эти материалы могли быть найдены и обработаны с помощью ЭВМ. Для возникновения, осуществления и защиты авторских прав не требуется регистрация произведения или соблюдение каких-либо иных формальностей. В отношении программ для ЭВМ и баз данных возможна регистрация, осуществляемая по желанию правообладателя в соответствии с правилами статьи 1262 Кодекса (пункт 4 статьи 1259 Кодекса). При этом ст. 1261 Кодекса п. 5 ст. 1259 Кодекса п. 2 ст. 1260 Кодекса п. 4 ст. 1259 Кодекса 5 регистрируется конкретная созданная программа для ЭВМ или база данных. Для государственной регистрации программы для ЭВМ или базы данных заявителем подается заявка на регистрацию, которая должна содержать: заявление с указанием правообладателя (заявителя), а также автора, если он не отказался быть упомянутым в качестве такового, и места жительства или места нахождения каждого из них (кроме базы данных, государственная регистрация которой осуществляется в соответствии с пунктом 3 статьи 1334 Кодекса); депонируемые материалы, идентифицирующие программу для ЭВМ или базу данных, включая реферат. На основании заявки на регистрацию Роспатент непосредственно или с привлечением подведомственного ему Федерального государственного бюджетного учреждения «Федеральный институт промышленной собственности» (ФИПС) проверяет наличие необходимых документов и материалов, их соответствие предусмотренным требованиям. При положительном результате проверки Роспатент вносит программу для ЭВМ или базу данных соответственно в Реестр программ для ЭВМ и в Реестр баз данных, выдает заявителю свидетельство о государственной регистрации и публикует сведения о зарегистрированных программе для ЭВМ или базе данных в официальном бюллетене Роспатента.

**Выполнения работы**

В качестве программы можно использовать учебную программу, полученную у преподавателя или использовать свою бакалаврскую дипломную работу. Используя материалы сайта <https://rospatent.gov.ru/ru/about> выполнить слелующее.

1. Изучить «Руководство по осуществлению административных процедур и действий в рамках предоставления государственной услуги по государственной регистрации программы для электронных вычислительных машин или базы данных и выдаче свидетельств о государственной регистрации программы для электронных вычислительных машин или базы данных, их дубликатов (PDF)»;
2. Заполнить заявление заявителя. По образцу, представленному на сайте.
3. Согласие на обработку персональных данных
4. Согласие на указание сведений об авторе.
5. Депонируемые материалы, идентифицирующие программу для ЭВМ или базу данных, включая реферат;

Контрольный вопросы:

1. Какие объекты интеллектуальной собственности можно зарегистрировать в Роспатенте?
2. Как регистрируется программ для ЭВМ ?
3. Что является доказательством авторского права на программу для ЭВМ?
4. Как можно подать документы в Роспатент.
5. Какой размер госпошлины определен для физического лица?
6. Каков максимальный срок публикации сведений о зарегистрированных программе для ЭВМ или базе данных в официальном бюллетене Роспатента?
7. Каков максимальный срок проверки соответствия заявки установленным законодательством Российской Федерации требованиям?
8. Что из себя представляют депонируемые материалы?

Лабораторная работа № 7

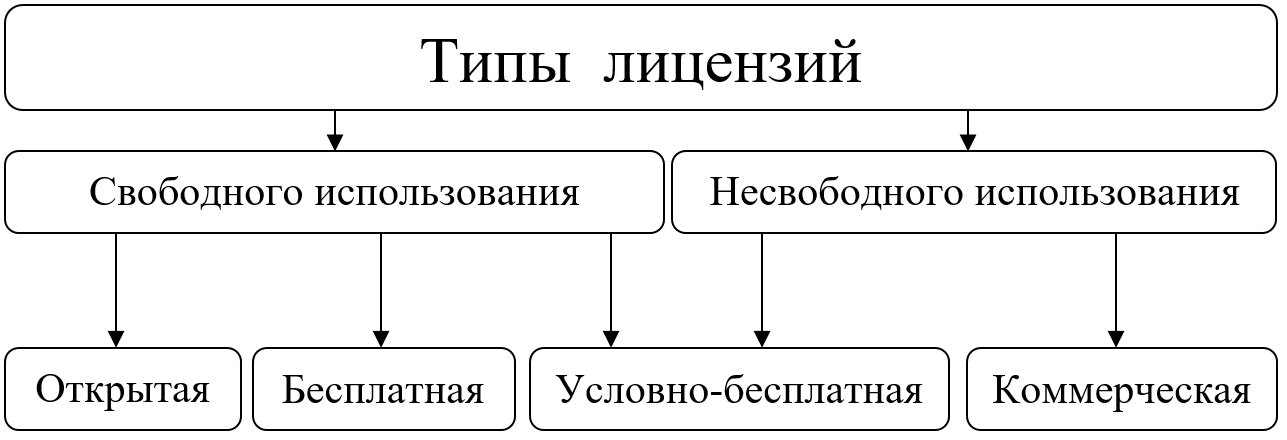
Лицензирование программного обеспечения

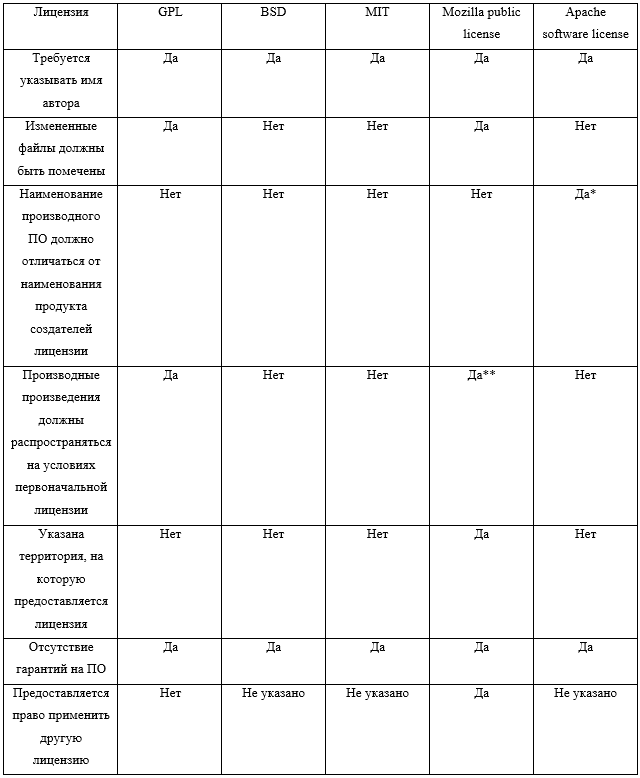
Цель работы: Изучить принципы лицензирования программного обеспечения

Краткие теоретические сведения.

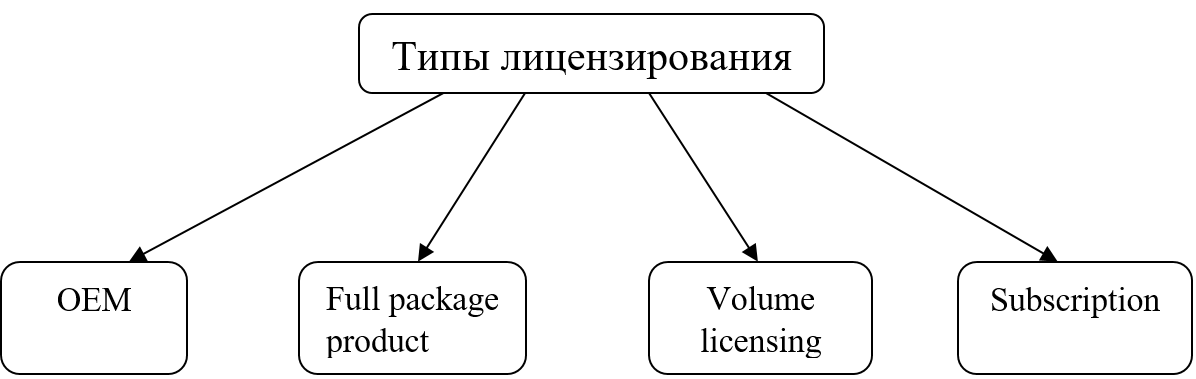
Основным документом, который определяет права и обязанности пользователя программного обеспечения, является лицензионное соглашение (licence agreement), которое прилагается к приобретенному продукту либо в виде бумажного документа, либо в электронном виде. Именно это соглашение определяет правила использования данного экземпляра продукта. По сути, лицензия выступает гарантией того, что издатель ПО, которому принадлежат исключительные права на программу, не подаст в суд на того, кто ею пользуется. Иными словами, издатель программного обеспечения ставит определенные защитные рамки по использованию его программного обеспечения.

Классификация лицензий и типы лицензирования ПО. В основном программы делятся на две большие группы — свободного использования (бесплатная и открытая лицензия) и несвободного (коммерческая лицензия), а также между ними существуют условно-бесплатные программы, которые можно отнести к двум группам пополам, такие программы можно скачать и использовать, но пока ее не оплатить у вас могут возникнуть некоторые проблемы или ограничения.



К открытым относятся: Open Source программы с открытым кодом которые можно модифицировать.К бесплатным относятся: Freeware, GPL, Adware, Postcardware, Donationware, Nagware/Begware.К условно-бесплатным относятся: ShareWare, TrialWare, Demoware.К коммерческим относятся: Commercial главная цель таких программ получение прибыли, код программ закрыт.Для наглядности рассмотрим сравнительную характеристику условий самых распространенных лицензий в виде таблицы, где будет указано о наличии или отсутствии в лицензии тех или иных требований. Все лицензии, которые будут рассматриваться являются лицензиями, одобренными Open Source Initiative для распространения ПО с открытым исходным текстом.

\* Если нет письменного разрешения об использовании наименования продукта создателей лицензии.\*\* В данном случае речь идет об исходном тексте.Защита своих авторских прав для разработчика — это и есть различные схемы лицензирования программного обеспечения. По каждому отдельному виду программного продукта применяются разные типы лицензирования.



Кратко разберем какой тип лицензирования что из себя представляет.  
OEM. Предустановленное ПО является одним из самых дешевых вариантов. Он заключается в том, что пользователь приобретает ПО вместе с самим компьютером или сервером и использовать его можно только на купленном ПК.

Full Package Product. «Коробочный» продукт применяется в основном для розничной торговли и удобен для частных лиц или малого бизнеса. Разрешение на использование программного продукта на одном компьютере дает покупка одной «коробки» и не важно, сколько людей будет пользоваться этим ПК. Так же можно сменить ПК, но определенное количество раз.

Volume Licensing. Корпоративная лицензия удобна для компаний, у которых много сотрудников, компьютеров и поэтому нужно приобретать много лицензий. При этом компания получает одну именную лицензию на программное обеспечение, которая содержит информацию о заказчике (название, адрес и т.д.), перечень ПО и ключи для его установки. В основном при такой схеме лицензирования компаниям, заказывающим именную лицензию, разработчики или распространители ПО предоставляют значительные скидки, техническую поддержку, решения нестандартных ситуаций и т. п. На сегодня она является лучшей для покупки нового ПО или его обновления для компаний.

Subscription. Подписка на лицензирование программного обеспечения предусматривает внесение ежемесячных или ежегодных платежей. Эта схема удобна компаниям, которые покупают более 10 лицензий. Она позволяет пользователям за минимальные начальные затраты получить практически все основные преимущества использования данного продукта.



Итак, теперь посмотрим в чем же разница типов лицензирования, а что бы это было нагляднее представим в виде таблицы.Теперь по этой таблице можно сделать вывод, кому, что больше подойдет.OEM версия подойдет для тех кто закупает новое оборудование. Если ПО будет уже предустановлено сборщиком, то оборудование обойдется намного дешевле чем покупать самому и устанавливать на каждое устройство. Выгода как по времени, так и по цене.FPP версия подойдет для тех у кого уже куплено оборудование, но отсутствует на нем нужное ПО, особенно если компания маленькая и сотрудники будут пользоваться одним ПК по несколько человек.VL версия подойдет для больших компаний которым нужна быстрая тех. поддержка и возможность решения нестандартных ситуаций. А также при покупке лицензии для всей компании всегда существуют очень хорошие скидки.а SUB версия подойдет для тех кто хочет использовать ПО кратковременно, или не знает на сколько данное ПО ему пригодится. Если же продукт нужен на долгое использование, то лучше посмотреть версию из “коробки”.

**Выполнения работы**

В качестве программы можно использовать учебную программу, полученную у преподавателя или использовать свою бакалаврскую дипломную работу. Или учебную программу на одном из известных Web движков.

Используя данные из сети Интернет https://www.gnu.org/licenses/quick-guide-gplv3.ru.html составьте лицензионное соглашение на разработанную вами программу программу по схеме GNU GPL.

**Схема GNU GPL**

Текст GNU GPL состоит из нескольких пронумерованных разделов. Ниже приведена схема версии 2.0 лицензии. Эта схема не имеет никакой юридической силы и служит только для краткого ознакомления.

1. Определения
   * (первый абзац) Определение термина «программа»
   * (второй абзац) Область действия лицензии
2. Право на копирование и распространение
3. Изменение программы
   * (первый абзац) Право на изменения при соблюдении следующих условий:
     + a) добавление информации об изменении в модифицированных файлах;
     + b) лицензирование модифицированных версий на условиях GNU GPL;
     + c) условное требование интерактивного вывода информации об авторских правах и отсутствии гарантии.
   * (абзацы 2—4) Уточнение термина «производная работа»
4. Требование предоставления исходного кода
   * (первый абзац) Возможные варианты распространения исполнимого кода:
     + a) распространение вместе с исходным кодом, или
     + b) распространение с гарантией предоставления исходного кода, или
     + c) (для некоммерческого использования) распространение вместе с такой гарантией, полученной от третьего лица.
   * (второй абзац) Определение термина «исходный код»
   * (третий абзац) Достаточность одинакового доступа для копирования исполнимого и исходного кодов
5. Прекращение действия лицензии при нарушении её условий
6. Акты, означающие принятие лицензии
7. Запрещение дополнительных ограничений при дальнейшем распространении
8. Внешние ограничения не снимают обязательства выполнять условия лицензии
9. Возможность географических ограничений
10. Будущие версии GNU GPL
11. Запросы на исключения из правил
12. Отказ от предоставления гарантий
13. Отказ от ответственности

Контрольный вопросы:

1. Для чего необходимо лицензировать программный продукт?
2. Какие особенности распространения OEM лицензий?
3. В чем заключаться особенности лицензии BSD?
4. В чем заключаются особенности лицензирования GPL?
5. В чем заключается особенности лицензирования по подписке?
6. В чем особенности облачной модели SaaS?

Лабораторная работа № 8

Контроль целостности информации

Цель работы: Использовать программу хэширования для проверки целостности данных.

Краткие теоретические сведенья

Важно понимать, были ли данные повреждены или была совершена попытка их фальсификации. Для определения того, были ли данные изменены или остались такими же, можно использовать программу хэширования. Программа хэширования выполняет преобразование данных или файла используя хэшфункцию, которая выдает некое значение (обычно значительно короче, чем сами исходные данные). Существует множество разных хэш-функций, одни очень простые, другие, напротив, очень сложные. Если одна и та же хэш-функция выполняется для преобразования одних и тех же данных, то значение, которое будет получено, будет всегда одинаково. Если данные были каким-то образом изменены, то полученное значение хэш-функцией будет отличаться.

Как работает хэш-функция? Сравнение хеш функций

Разные виды хэш-функций производят вывод разной величины, но возможный размер данных на выходе для каждого из алгоритмов хеширования всегда является постоянным. Например, алгоритм SHA-256 может производить вывод исключительно в формате 256-бит, в то время как SHA-1 всегда генерирует 160-битный дайджест.

На примере слов “Binance“ и “binance”, которые пропущены через алгоритм хеширования SHA-256 (используется в сети Биткоин), проиллюстрируем это.

|  |  |
| --- | --- |
| SHA-256 | |
| Входные данные | Результат (256 бит) |
| Binance | f1624fcc63b615ac0e95daf9ab78434ec2e8ffe402144dc631b055f711225191 |
| binance | 59bba357145ca539dcd1ac957abc1ec5833319ddcae7f5e8b5da0c36624784b2 |

Обратите внимание, что незначительное изменение (регистр первой буквы) привело к другому значению хэша. Но поскольку использовался SHA-256, данные на выходе имеют фиксированный размер в 256 бит (или 64 символа).

Не имеет значения какое количество раз эти два слова будут пропущены через алгоритм, два выхода не будут видоизменяться, поскольку они являются постоянными.

При обработке тех же входных данных через алгоритм хеширования SHA-1, получаются следующие результаты:

|  |  |
| --- | --- |
| SHA-1 | |
| Входные данные | Результат (160 бит) |
| Binance | 7f0dc9146570c608ac9d6e0d11f8d409a1ee6ed1 |
| binance | e58605c14a76ff98679322cca0eae7b3c4e08936 |

Акроним SHA расшифровывается как Secure Hash Algorithms (безопасный алгоритм хеширования). Он относится к набору криптографических хэш-функций, который включает такие алгоритмы как SHA-0 и SHA-1 вместе с группами SHA-2 и SHA-3. SHA-256 является частью группы SHA-2, наряду с SHA-512 и другими аналогами. В настоящее время, только группы SHA-2 и SHA-3 считаются безопасными.

Коллизии

Коллизии хэш-функций - это появление одного хэш-кода на два различных массива информации.

Коллизии возникают из-за того, что в хеш используется мало символов. Чем меньше знаков использует конечная формула, тем больше вероятность итерации (повтора) одного и того же хэш-кода на разные наборы данных. Чтобы снизить риск появления коллизии, применяют двойное хеширование строк, образующее открытый и закрытый ключ – то есть, используется 2 протокола, как, например, в сети Bitcoin.

Технические параметры

Рассмотрим основополагающие характеристики протоколов хеширования:

Наличие внутрисистемных уравнений, позволяющих модифицировать нефиксированный объем информации в лаконичный набор знаков и цифр заданной длины.

Прозрачность для криптографического аудита.

Наличие функций, которые дают возможность надежно кодировать первоначальную информацию.

Способность к расшифровке хэш-суммы с использованием вычислительного оборудования средней мощности.

Есть еще следующие важные свойства алгоритмов: способность «свертывать» любой массив данных, производить хэш конкретной длины, распределять равномерно на выходе значения функции.

Любые изменения во входящем сообщении (другая буква, цифра, знак препинания, даже лишний пробел) внесут коррективы в итоговый хэш-код. Он будет такой же длины, но с иными символами.

Требования

К эффективной во всех отношениях хэш-функции выдвигаются следующие требования:

протокол должен обладать чувствительностью к изменениям, происходящим во входящих документах – то есть, обязан распознавать перегруппировку абзацев, переносы, другие элементы текстовых данных;

технология обязана таким образом преобразовывать поток информации, чтобы на практике было невозможно осуществить обратную процедуру – восстановить из значения хэш первоначальные данные;

протокол должен использовать такие математические уравнения, которые исключают или значительно снижают факт появления коллизии.

Виды хешей

MD5 - хеш-функция была описана в 1991 году Роном Ривестом. В настоящее время она считается небезопасной и не используется в криптографии.

SHA-1 - алгоритм хеширования 1 используется с 1995 года, но не считается безопасным с 2005 года, после того, как его несколько раз успешно атаковали. Надежнее использовать SHA-2 или SHA-3.

SHA-2 - это семейство хеш-функций, которые являются преемниками SHA-1. Функции SHA-2 содержат многочисленные улучшения, которые делают их безопасными. Однако SHA-256 и SHA-512 бывают уязвимы для атак с увеличением длины, поэтому в определенных ситуациях лучше всего реализовывать SHA-3.

SHA-3 - самый новый член семейства Secure Hash Algorithm, он построен совсем не так, как его предшественники.

Он еще не заменил SHA-2, а просто предоставляет криптографам еще один вариант, который может обеспечить повышенную безопасность в определенных ситуациях.

RIPEMD - еще одно семейство функций, разработанное академическим сообществом. RIPEMD основан на многих идеях MD4 (предшественника MD5) и не ограничен никакими патентами. RIPEMD-160 считается безопасным, но он не получил широкого распространения.

Whirlpool - это хеш-функция, основанная на модификации AES, она также не подпадает под действие каких-либо патентов. Считается безопасной, но работает медленнее, чем альтернативные варианты

**Выполнение работы:**

Шаг 1: Создайте текстовый файл

a. Найдите на своем компьютере программу «Notepad» и откройте ее.

b. Введите любой текст в программу. См. Рисунок 7

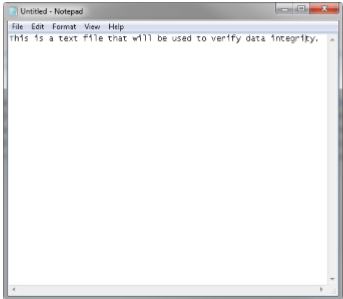


Рисунок 7- Notepad

c. Выберите «Файл» > «Сохранить».

d. Перейдите на «Рабочий стол».

e. Введите «Hash» в поле Имя файла и нажмите Сохранить.

Шаг 2: Установите «HashCalc»

a. Откройте веб-браузер и перейдите по ссылке: [http://www.slavasoft.com/download.htm. См](http://www.slavasoft.com/download.htm.%20См). Рисунок 8

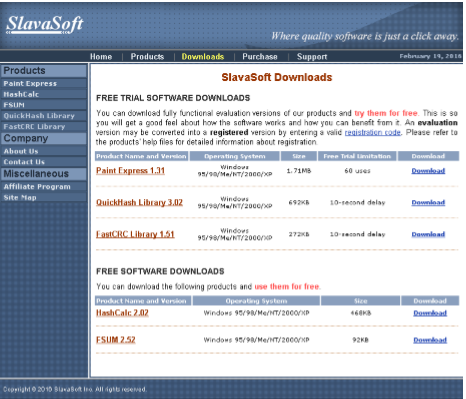


Рисунок 8- www.slavasoft.com

b. Щелкните «Загрузить» в строке HashCalc 2.02.

c. Откройте архив hashcalc.zip и запустите находящийся в нем файл setup.exe.

d. Следуйте инструкциям мастера установки, чтобы установить HashCalc. Если в процессе установки возникнут вопросы, обратитесь за помощью к инструктору.

e. Нажмите «Завершить» на последнем экране и закройте файл README, если он открыт. При желании можно прочитать этот файл.

f. Теперь программа «HashCalc» установлена и запущена. См. Рисунок 9

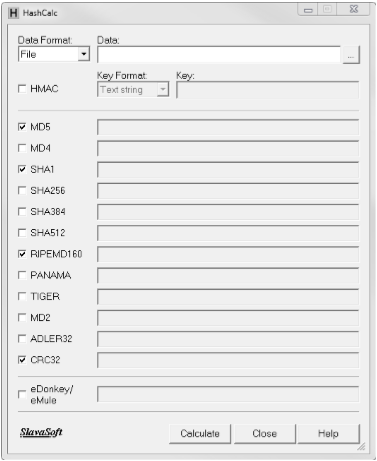


Рисунок 9- HashCalc

Шаг 3: Вычислите хэш файла Hash.txt

a. Задайте следующие параметры в HashCalc.

1) Формат данных: File.

2) Данные: щелкните кнопку … рядом с полем «Данные», перейдите на «Desktop» и выберите файл «Hash.txt».

3) Снимите флажок HMAC.

4) Снимите флажки со всех типов хэшей, кроме MD5.

b. Нажмите кнопку Calculate (Рассчитать).

Какое значение указано рядом с MD5?

Шаг 4: Внесите изменения в файл Hash.txt

a. Перейдите на Desktop и откройте файл Hash.txt.

b. Внесите любое небольшое изменение в текст, например, удалите букву или добавьте пробел или точку.

c. Щелкните File > Save (Файл > Сохранить) и закройте Notepad.

Шаг 5: Вычислите новый хэш файла Hash.txt

a. Снова щелкните кнопку Calculate (Рассчитать) в программе HashCalc.

Какое значение указано рядом с MD5?

Отличается ли это значение от значения, записанного в шаге 3?

b. Поставьте флажки рядом со всеми остальными типами хэша.

c. Щелкните Calculate (Рассчитать).

d. Обратите внимание, что для многих хэш-функций созданы хэши разной длины. Почему?

Контрольные вопросы:

1. Что характеризует целостность информации?
2. Что такое Хэш функция?
3. Для чего используется ХЭШ функция?
4. Что такое коллизия ХЭШ?
5. Какие алгоритмы ХЭШ используются в настоящее время ?

Лабораторная работа № 9

Организация длительного хранения информации (Архивное дело)

Цель работы:Изучить организацию архивного дела структуру информации в архиве.

Краткие теоретические сведенья

Организация хранения документов в архиве:

2.1. Архивные документы в архиве систематизируются по исторически и (или) логически сложившимся комплексам документов – архивным фондам, объединенным архивным фондам и архивным коллекциям.

2.2. Архивный фонд формируется из архивных документов, образовавшихся в процессе деятельности (жизнедеятельности) соответствующего государственного органа, органа местного самоуправления, организации, гражданина, семьи, рода.

2.3. В целях рациональности хранения, комплектования, учета и использования архивных документов допускается формирование:

объединенного архивного фонда – из архивных документов, образовавшихся в процессе деятельности двух и более организаций, действовавших на определенной территории и (или) объединенных одним объектом деятельности, органа государственной власти (органа местного самоуправления) и подведомственных ему организаций, правопредшественников и правопреемников, а также двух и более человек, связанных между собой родственными, профессиональными или творческими связями;

архивной коллекции – из разрозненных архивных документов, объединенных по одному или нескольким признакам (тематическому, объектному, авторскому, хронологическому или другим).

2.4. Использование нефондовой организации архивных документов допускается только в отношении аудиовизуальных документов.

2.5. Фондовая принадлежность архивных документов архивного фонда, объединенного архивного фонда или архивной коллекции может быть изменена (произведено перефондирование) в случае обнаружения ошибок при их первоначальном фондировании.

Исторически сложившаяся архивная коллекция перефондированию не подлежит.

2.6. Документы, являющиеся входящими, относятся к архивному фонду того фондообразователя, который их получил.

Документы, являющиеся исходящими копиями (отпусками), относятся к архивному фонду фондообразователя, который является их создателем.

Документы внутреннего обращения относятся к архивному фонду фондообразователя, в котором они созданы.

2.7. В случае реорганизации или ликвидации организации законченные делопроизводством дела включаются в ее архивный фонд, незаконченные дела, переданные для окончания делопроизводством в организацию-правопреемник, включаются в архивный фонд последними.

2.8. В архивный фонд организации, кроме документов, образовавшихся в процессе деятельности (жизнедеятельности) организации, включаются архивные документы:

представительств, филиалов, иных обособленных подразделений организации;

коллегиальных органов организации;

образовавшиеся в деятельности собраний (комитетов) кредиторов, временного, административного, внешнего, конкурсного, финансового управляющего, предоставивших обеспечение лиц во время прохождения процедур, применяемых в деле о банкротстве к организации – должнику;

созданные при организации профсоюзных организаций, осуществляющих деятельность до вступления в силу Федерального закона от 19 мая 1995 г. № 82-ФЗ «Об общественных объединениях»[[4]](https://archives.gov.ru/documents/rules/pravila-2020.shtml" \l "04);

назначенной в организации ликвидационной комиссии (ликвидатора).

2.9. При наличии на хранении в архиве архивного фонда организации новый архивный фонд создается в случаях:

изменения профиля ее деятельности;

изменения формы собственности ее имущества.

Из архивных документов государственной или муниципальной организации, подвергшейся преобразованию с изменением формы собственности ее имущества, и архивных документов вновь созданной организации-правопреемника допускается формирование объединенного архивного фонда в случае передачи архивных документов организации-правопреемника в государственную или муниципальную собственность.  
 III. Требования к зданию и помещениям архива

3.1. Архив размещается в здании(ях) и (или) отдельных помещениях здания. Основные помещения, предназначенные для хранения, учета, комплектования и использования архивных документов, и вспомогательные (административно-хозяйственного, технического и бытового назначения), должны соответствовать требованиям, установленным главой V Правил, а также общими требованиями безопасности к общественным зданиям и сооружениям[[5]](https://archives.gov.ru/documents/rules/pravila-2020.shtml" \l "05).

Состав, расположение, оборудование помещений основного назначения здания архива должны обеспечивать сохранность архивных документов на всех участках работы с ними, соблюдение технологий архивных работ, а также оптимальное взаимодействие подразделений архива. К помещениям основного назначения относятся:

архивохранилища;

рабочие помещения работников архива;

помещения для приема и временного хранения, для адаптации архивных документов к новым условиям хранения после их перемещения (акклиматизации);

помещения для изоляции, дезинфекции и дезинсекции пораженных архивных документов;

помещения для обеспыливания, переплета и реставрации архивных документов;

помещения для копирования и фотореставрации архивных документов, обработки пленки и проведения технического контроля аудиовизуальных документов;

помещения для хранения учетных документов;

читальный зал с участками выдачи дел и справочно-поисковых средств, временным хранилищем (просмотровый зал, комнаты прослушивания фонодокументов) (далее – читальный зал);

помещения каталогов или при отсутствии специально выделенного помещения – рабочая комната архива;

научно-справочная библиотека;

методический кабинет;

помещения для экспонирования архивных документов (выставочный зал).

3.2. Размещение архива в приспособленном здании (помещениях) допускается только после его обследования (в том числе подвальных и чердачных помещений, работоспособности электросетей, отопительных и вентиляционных систем, прочности межэтажных перекрытий с учетом потенциальных нагрузок) представителями архива, соответствующего уполномоченного органа исполнительной власти в сфере архивного дела или органа местного самоуправления.

Предназначенные для хранения архивных документов помещения в приспособленном под архив здании должны быть изолированы от остальных его помещений.

Не допускается размещение архива:

в зданиях, в которых расположены предприятия общественного питания, готовящие горячую пищу, продуктовые магазины и (или) склады, организации, хранящие и (или) применяющие пожароопасные и агрессивные химические вещества;

в ветхих строениях, деревянных постройках, помещениях с относительной влажностью воздуха более 55% и с печным отоплением;

в подвалах и помещениях, расположенных ниже уровня поверхности земли, за исключением хранилищ (помещений), обеспечивающих защиту от проникновения в них грунтовых, талых и паводковых вод.

3.3. Архивохранилище должно:

быть отделено от других помещений архива огнеупорными стенами и перекрытиями;

быть удалено от лабораторных, производственных, бытовых помещений архива и не иметь общих с ними вентиляционных каналов;

быть гидроизолировано;

иметь высоту не менее 2,2 м от пола до нижней части выступающих конструкций перекрытия потолка;

иметь покрытие стен, полов, потолков, оборудования и средств хранения архивных документов, не собирающих пыль и не выделяющих агрессивные химические вещества;

иметь беспрепятственные выходы к лифтам и лестничным клеткам;

не иметь внутри распределительных щитов, предохранителей и отключающих рубильников, труб водоснабжения и канализации, технологических или бытовых выводов воды.

3.4. Не допускается хранение фотодокументов, видеодокументов и электронных документов на физически обособленных носителях с магнитным рабочим слоем в архивохранилище с постоянными или переменными магнитными полями напряженностью более 400 а/м.

**Выполнение работы:**

<https://ya.ru/archive>

Войдите на сайт проекта «Поиск по архивам» См Рисунок 10

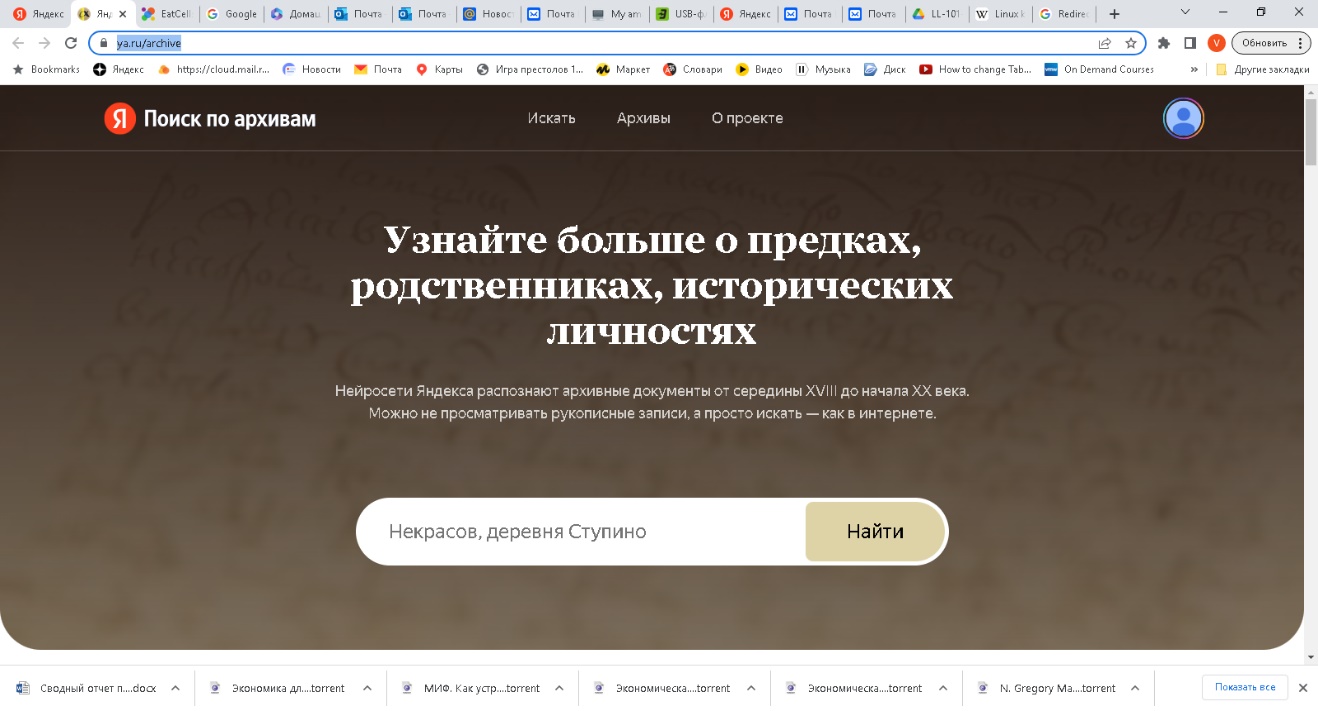


Рисунок 10- Сайт проекта «Поиск по архивам»

1. Определите какие архивы участвуют в проекте.
2. Определите какие документы оцифрованы, что это за документы.
3. Осуществите поиск по отсканированным документам по своей фамилии, фамилии своего деда, девичьей фамилии матери преподавателя. См. Рисунок 11

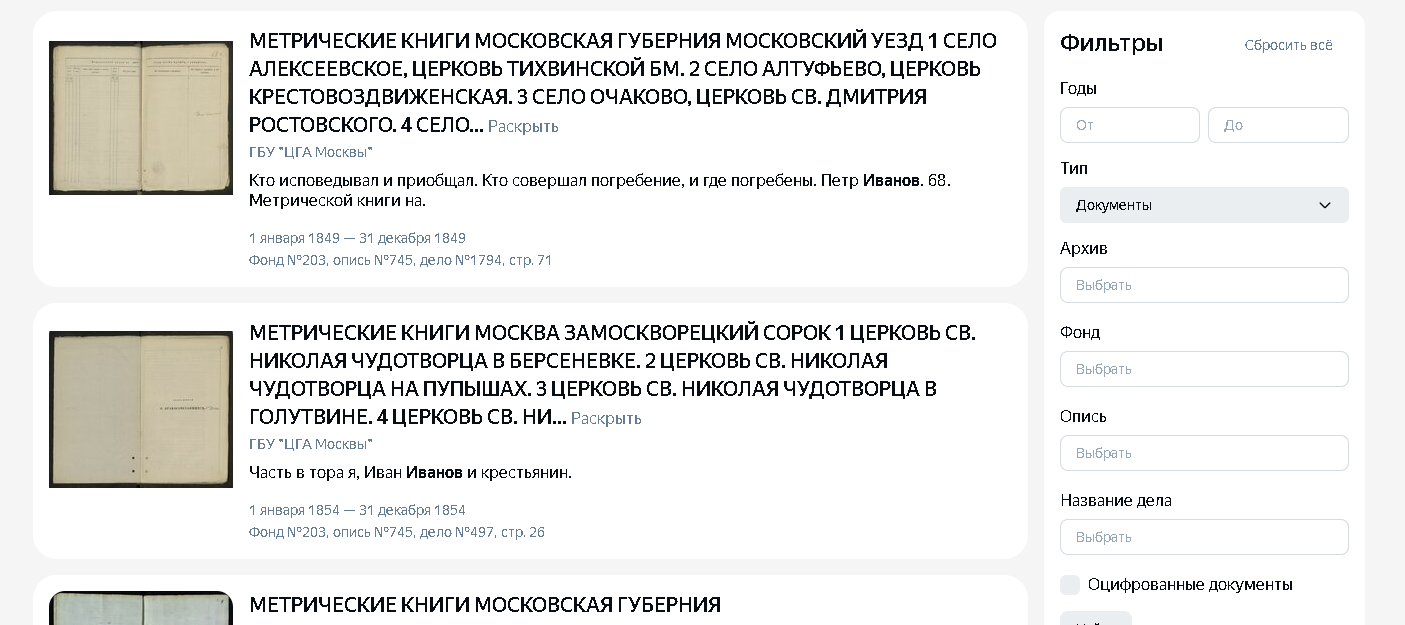


Рисунок 11- поиск по отсканированным документам по своей фамилии

1. Обратите внимание где хранятся и как организовано хранение документов.
2. Определите срок хранения документа его реквизиты хранения.

Контрольные вопросы:

1. Что такое архивное хранение документов?
2. Что такое архив?
3. Как организуется архив на предприятии?
4. Что такое Фонд, что такое опись?
5. Какие проблемы оцифровки архивов ?

ЛИТЕРАТУРА

1. Блюмин, А.М. Мировые информационные ресурсы / А.М. Блюмин, Н.А. Феоктистов. – Москва : «Дашков и К°», 2022. – 296 с.
2. Инструмент для формирования запросов к поисковым системам [Электронный ресурс]. – Режим доступа: <http://www.shipbottle.ru/ir/>.
3. studme.org [Электронный ресурс]. – Режим доступа: https://studme.org/54444/informatika/informatsionnyy\_poisk\_internete.
4. studopedia.su [Электронный ресурс]. – Режим доступа: https://studopedia.su/17\_23196\_metodika-informatsionnogo-poiska.html.
5. yandex.ru [Электронный ресурс]. – Режим доступа: https://yandex.ru/dev/.
6. Хабр [Электронный ресурс]. – Режим доступа: https://habr.com/company/yandex/blog/221139/.
7. reactiv.top [Электронный ресурс]. – Режим доступа: https://www.reactiv.top/kak-sozdat-blog-v-blogger/.
8. РОСКОМНАДЗОР[Электронный ресурс]. – Режим доступа: https://pd.rkn.gov.ru
9. moysklad.ru [Электронный ресурс]. – Режим доступа: <https://www.moysklad.ru/poleznoe/statyi/chto-takoe-oblachnye-servisy/>.
10. Журавлев, А.Е. Информатика. Практикум в среде Microsoft Office 2016: учебное пособие/ А.Е.Журавлев.-Москва:Изд-во Лань,2020.-96 с.
11. rospatent.gov.ru [Электронный ресурс]. – Режим доступа: <https://rospatent.gov.ru/ru/about>
12. Поиск по архивам [Электронный ресурс]. – Режим доступа: <https://ya.ru/archive>
13. Реестр операторов, осуществляющих обработку персональных данных[Электронный ресурс]. – Режим доступа: <https://pd.rkn.gov.ru/operators-registry/operators-list/>
14. Серогодский, В. В. Microsoft Office 2016 / Office 365. Полное руководство / В. В. Серогодский, А. П. Тихомиров, Д. П Сурин. – Москва : Наука и Техника, 2020. – 448 с.
15. Мicrosoft [Электронный ресурс]. – Режим доступа: <https://devblogs.microsoft.com/dotnet>
16. Мicrosoft [Электронный ресурс]. – Режим доступа: <https://www.microsoft.com/ru-ru/microsoft-365/sharepoint/collaboration>
17. Инструкция по работе с платформой MICROSOFT TEAMS [Электронный ресурс]. – Режим доступа: <https://docviewer.yandex.ru/view/983121343>
18. MS-TEAMS.RU [Электронный ресурс]. – Режим доступа: https://ms-teams.ru/.
19. PowerShell Documentation // Мicrosoft URL: <https://docs.microsoft.com/en-us/powershell/>
20. Weekend Scripter: The Best Ways to Learn PowerShell // Мicrosoft URL: <https://devblogs.microsoft.com/scripting/weekend-scripter-the-best-ways-to-learn-powershell/>