

Автономная некоммерческая образовательная организация высшего образования «Сибирский институт бизнеса и информационных технологий»

## РЕФЕРАТ

ПО ДИСЦИПЛИНЕ: ИнфБиог Культура НА ТЕМУ:

**Понятие «информация». Основные пути поиска информации**

Выполнил(а): \_\_\_\_\_ (Ф.И.О. студента)

\_\_\_\_\_ (направление, группа) Проверил(а):

\_\_\_\_\_ (ученая степень, звание, Ф.И.О. преподавателя)

\_\_\_\_\_ (дата) Омск 20\_\_ г.

## **СОДЕРЖАНИЕ**

<b>ВВЕДЕНИЕ</b>	<b>1</b>
1.Понятие информации	3
2.Восприятие информации	5
3.Свойства информации	8
<b>ЗАКЛЮЧЕНИЕ</b>	<b>11</b>
<b>СПИСОК ЛИТЕРАТУРЫ</b>	<b>12</b>



## ВВЕДЕНИЕ

Человек получает информацию о внешнем мире с помощью своих органов чувств. Практически около 90% информации человек получает при помощи органов зрения, примерно 9% - при помощи органов слуха, и только 1% при помощи остальных органов чувств (обоняния, вкуса, осязания). Следует отметить, что органы чувств человека получили название анализаторов, поскольку именно через эти органы информация попадает в головной мозг. А вот, например, для лисы, собаки и многих других животных основная информация та, которая поступает через нос. У них хорошо развито обоняние. Для летучих мышей главная информация - звуковая, они воспринимают ее своими большими, чуткими ушами.

Компьютер, помогающий человеку хранить и обрабатывать информацию, приспособлен в первую очередь для обработки текстовой, числовой, графической информации.

Понятие «информация» является одним из фундаментальных в современной науке вообще и базовым для информатики. Информацию наряду с веществом и энергией рассматривают в качестве важнейшей сущности мира, в котором мы живем. Однако, если задаться целью, формально определить понятие «информация», то сделать это будет чрезвычайно сложно.

В простейшем бытовом понимании с термином «информация» обычно ассоциируются некоторые сведения, данные, знания и т.п. Информация передается в виде сообщений, определяющих форму и представление передаваемой информации. Примерами сообщений являются музыкальное произведение; телепередача; команды регулировщика на перекрестке; текст, распечатанный на принтере и т.д. При этом предполагается, что имеются «источник информации» и «получатель информации».

В связи с этим цель работы - изучить понятие, восприятие информации и ее свойства и основные пути поиска информации. Для достижения этой цели нужно достичь следующих задач:

- 1) рассмотреть и изучить литературу, соответствующую теме работы;
- 2) дать понятие информации;
- 3) изучить восприятие информации и пути поиска информации;
- 4) рассмотреть свойства информации.

## **ГЛАВА 1. ПОНЯТИЕ ИНФОРМАЦИИ**

Термин «информация» происходит от латинского слова «informatio», что означает сведения, разъяснения, изложение.

Информация - это настолько общее и глубокое понятие, что его нельзя объяснить одной фразой. В это слово вкладывается различный смысл в технике, науке и в житейских ситуациях.

Информация - сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые воспринимают информационные системы (живые организмы, управляющие машины и др.) в процессе жизнедеятельности и работы.

Одно и то же информационное сообщение (статья в газете, объявление, письмо, телеграмма, справка, рассказ, чертёж, радиопередача и т.п.) может содержать разное количество информации для разных людей - в зависимости от их предшествующих знаний, от уровня понимания этого сообщения и интереса к нему.

Информация есть характеристика не сообщения, а соотношения между сообщением и его потребителем. Без наличия потребителя, хотя бы потенциального, говорить об информации бессмысленно.

В случаях, когда говорят об автоматизированной работе с информацией посредством каких-либо технических устройств, обычно в первую очередь интересуются не содержанием сообщения, а тем, сколько символов это сообщение содержит.

Применительно к компьютерной обработке данных под информацией понимают некоторую последовательность символических обозначений (букв, цифр, закодированных графических образов и звуков и т.п.), несущую смысловую нагрузку и представленную в понятном компьютеру виде. Каждый новый символ в такой последовательности символов увеличивает информационный объём сообщения.

## **ГЛАВА 2. ВИДЫ ИНФОРМАЦИИ**

Особую роль для общества играет документированная информация. Документы - это информация, зафиксированная на материальном носителе - бумаге или машинном носителе, имеющем реквизиты, позволяющие его идентифицировать.

Возможность записи информации в письменном виде - в форме последовательности знаков - привела к образованию государств, возникновению бюрократии и появлению почтовых служб. Параллельно это привело - к появлению грамотных людей - людей, умеющих читать, писать и искать информацию для решения различных проблем.

Возникновение письменности позволило людям не только передавать информацию, но и накапливать ее в форме записок, писем и рукописей в архивах, а также в личных и публичных библиотеках.

Развитие промышленного производства в XVIII-XIX веках потребовало большого числа специалистов, для подготовки которых было открыто большое число университетов. Это дало мощный толчок для развития естественных наук - химии, физики, механики, математики и подготовки инженерных кадров.

Развитие печатных станков привело к появлению и распространению газет как средств массовой информации и информатизации общества, а также появлению и распространению журналов для распространения литературных произведений. В это же время появились первые законы, регулирующие авторские права.

Изобретение в XIX - начале XX века телеграфа, радио и телефона открыло новые возможности в передаче информации и информатизации общества. Эти технические средства дали возможность практически мгновенно передавать информацию на любые расстояния.

Следующим шагом технического прогресса стало появление и развитие электроники, телевидения и радиовещания к середине XX века. Изобретение телевидения позволило людям видеть на экранах телевизоров события, происходящие в самых различных точках планеты, а изобретение магнитофона - накапливать звуковую и видеoinформацию на магнитных носителях.

Информация может существовать в самых разнообразных формах [2]:

- в виде текстов, рисунков, чертежей, фотографий;
- в виде световых или звуковых сигналов;
- в виде радиоволн;
- в виде электрических и нервных импульсов;
- в виде магнитных записей;
- в виде жестов и мимики;
- в виде запахов и вкусовых ощущений;
- в виде хромосом, посредством которых передаются по наследству

признаки и свойства организмов и т.д.

Предметы, процессы, явления материального или нематериального свойства, рассматриваемые с точки зрения их информационных свойств, называются информационными объектами.

### **ГЛАВА 3. ВОСПРИЯТИЕ ИНФОРМАЦИИ**

На сегодняшний день нет недостатка информации в какой – либо области; напротив, наблюдается ее переизбыток. Люди часто путаются в одних и тех же понятиях, поскольку информация об одном и том же предмете может быть противоречивой.

Восприятие – это процесс отражения в сознании личности явлений и предметов в сумме их свойств, состояний, компонентов.

Процесс восприятия информации представляет собой высокоорганизованную внутреннюю работу, в которой участвуют все психические процессы: внимание, воображение, память, мышление. Для того, чтобы поступающая в мозг информация лучше усваивалась, ее необходимо осознавать или осмысливать. Восприятие как раз и выполняет функцию своеобразного проводника между новой информацией и ее осознанием.

Под каналами восприятия понимают преобладающую направленность в сторону одного органа чувств, которая обеспечивает лучшее усвоение поступающей информации.

- Визуальный канал. Направлен на усвоение информации путем большего сосредоточения на зрительных образах. В данном случае личности достаточно прочесть материал, и информация прочно «закрепится» в мозге.

- Аудиальный канал. Направлен на усвоение информации путем концентрации преимущественно на слуховых образах.

- Кинестетический канал. Направлен на усвоение информации путем сосредоточения преимущественно на физических ощущениях. Запах, вкус тоже имеют для этой личности первостепенное значение – она более всего внимательна к деталям и собственным чувствам.

- Дигитический канал. Направлен на усвоение информации путем концентрации на абстрактно – логических образах.

Перечисленные каналы восприятия являются ведущими, но кроме них имеются и другие: вкусовой, обонятельный, семантический и т.д. В соответствии с представленными особенностями каждого из каналов в психологии выделяют следующие виды восприятия информации: зрительный, слуховой, осязательный, речемыслительный. Каждый из перечисленных видов полностью соотносится с выше обозначенными каналами восприятия информации.

## ГЛАВА 4. СВОЙСТВА ИНФОРМАЦИИ

Как и всякий объект, информация обладает свойствами. Характерной отличительной особенностью информации от других объектов природы и общества, является дуализм: на свойства информации влияют как свойства исходных данных, составляющих ее содержательную часть, так и свойства методов, фиксирующих эту информацию.

С точки зрения информатики, наиболее важными представляются следующие общие качественные свойства: объективность, достоверность, полнота, точность, актуальность, полезность, ценность, своевременность, понятность, доступность, краткость и пр.

**Объективность информации.** Объективный – существующий вне и независимо от человеческого сознания. Информация объективна, если она не зависит от методов ее фиксации, чье-либо мнения, суждения. Отражаясь в сознании конкретного человека, информация перестает быть объективной, так как, преобразовывается (в большей или меньшей степени) в зависимости от мнения, суждения, опыта, знаний конкретного субъекта.

Информация достоверна, если она отражает истинное положение дел. Объективная информация всегда достоверна, но достоверная информация может быть как объективной, так и субъективной. Недостоверной информация может быть по следующим причинам:

- преднамеренное искажение (дезинформация) или непреднамеренное искажение субъективного свойства;
- искажение в результате воздействия помех («испорченный телефон») и недостаточно точных средств ее фиксации.

Информацию можно назвать полной, если ее достаточно для понимания и принятия решений. Неполная информация может привести к ошибочному выводу или решению.

Точность информации определяется степенью ее близости к реальному состоянию объекта, процесса, явления и т. п.

Актуальность информации – важность для настоящего времени, злободневность, насущность. Только вовремя полученная информация может быть полезна.

Самая ценная информация – объективная, достоверная, полная, и актуальная. При этом следует учитывать, что и необъективная, недостоверная информация (например, художественная литература), имеет большую значимость для человека. Социальная (общественная) информация обладает еще и дополнительными свойствами:

- имеет семантический (смысловой) характер, так как именно в понятиях обобщаются наиболее существенные признаки предметов, процессов и явлений окружающего мира.

- имеет языковую природу (кроме некоторых видов эстетической информации, например изобразительного искусства).

С течением времени количество информации растет, информация накапливается, происходит ее систематизация, оценка и обобщение. Это свойство назвали ростом и кумулированием информации. (Кумуляция – от лат; *simulatio* – увеличение, скопление).

Старение информации заключается в уменьшении ее ценности с течением времени. Старит информацию не само время, а появление новой информации, которая уточняет, дополняет или отвергает полностью или частично более раннюю. Научно-техническая информация стареет быстрее, эстетическая (произведения искусства) – медленнее.

Логичность, компактность, удобная форма представления облегчает понимание и усвоение информации.

## **ГЛАВА 5. ИНФОРМАЦИОННЫЙ ПОИСК: ВИДЫ, ЭТАПЫ ПОИСКА**

Побудительная причина информационного поиска - информационная потребность, выраженная в форме информационного запроса. Объектами

информационного поиска могут быть документы, сведения о их наличии и (или) местонахождении, фактографическая информация.

Условно информационный поиск делится на четыре основных вида: библиографический, документальный, фактографический и аналитический. Например, чтобы найти информационный источник (документ или издание), надо знать определенную совокупность библиографических данных (фактов), характеризующих его, отличающих от многих других: хотя бы от написанных тем же автором, на одну и ту же тему и т. д.

Виды информационного поиска в зависимости [7]:

- от цели - адресный (формально-механический) и семантический (тематический);

- от объекта поиска - документный и фактографический;

- от степени использования технических средств - ручной или автоматизированный.

- от функциональной роли – доминирующие / второстепенные, центральные / периферические, устойчивые / ситуативные потребности.

Все виды информационного поиска пересекаются, так как их цели и объекты часто взаимосвязаны. Например, документный и фактографический виды поиска могут быть как адресными, так и семантическими.

Информационный поиск производится при помощи информационно-поисковых систем (ИПС). ИПС - это комплекс связанных друг с другом отдельных частей, предназначенный для выявления в каком-либо множестве элементов информации, отвечающих на предъявленный информационный запрос. Информационный поиск осуществляется по определенным правилам, определяющим стратегию поиска. Стратегия информационного поиска зависит от типа поисковой задачи, критериев выдачи и характера диалога между потребителями информации и ИПС.

В общем случае процедура информационного поиска состоит из четырех этапов [7]:

- уточнения информационной потребности и формулировки запроса

- определения совокупности держателей информационных массивов
- извлечения информации из информационных массивов;
- ознакомления пользователя с полученной информацией и оценки результатов поиска.

Эффективность информационного поиска определяют показатели, характеризующие нахождение релевантных документов. Они делятся на семантические (точность и полнота поиска, коэффициент информационного шума и коэффициент потерь) и технико-экономические (оперативность поиска, стоимость и трудоемкость поиска).

С проблемой информационного поиска раньше всего столкнулись библиотекари. Для того, чтобы читатели могли находить в фондах библиотеки, интересующие их документы, в ней создаются различные каталоги и указатели.

## **5.1. ВИДЫ ИСТОЧНИКОВ ИНФОРМАЦИИ**

### **ДОКУМЕНТАЛЬНЫЙ ТИП ИСТОЧНИКОВ ИНФОРМАЦИИ**

Документ – это материальный носитель записи с зафиксированной на нём информацией для передачи её во времени и пространстве.

Под производственными документами имеется в виду совокупность текстов которые обеспечивают информационное обслуживание производственной жизни трудовых коллективов, нужды управления в государственной и производственной сферах

Бытовыми документами – совокупностью официальных и личных материалов, которая обеспечивает информационное обслуживание людей в быту они, как правило, представляют собой личную собственность человека

Существует следующее деление документов [1]:

1. По способу фиксирования информации (рукописные, печатные документы, кино и фотоплёнки, магнитные ленты).

2. По типу авторства (личные и общественные, например, расписка в получении денег и протокол собрания коллектива).

3. По статусу документа (официальные и неофициальные, например, постановление правительства и пояснительная записка).

4. По степени близости к эмпирическому материалу (первичные, например, заполненные анкеты, и вторичные – отчёт, написанный по результатам анкетирования на основе обобщения данных анкет).

5. По способу получения документа (естественно функционирующие в обществе, например, статистические отчёты по установленному образцу и «целевые», т.е. созданные по заказу журналиста – допустим, справка о деятельности учреждений).

По целевому назначению информационные издания делят на библиографические, реферативные и обзорные

Библиографическое издание – представляет собой библиографическое пособие содержащее упорядоченное множество библиографических записей, объединенных по какому-либо признаку. К библиографическим изданиям относят библиографические указатели, обзоры, списки, информационные издания.

Реферативные издания – содержат аннотации и рефераты. К реферативным изданиям относят – экспресс-информацию, реферативные журналы, реферативные сборники, информационные листки.

Экспресс-информация – самый оперативный вид реферативных изданий, обработка первичных документов при их подготовке занимает не более 1–1,5 месяца, выпускаются с частой периодичностью..

Реферативные журналы – охватывают с возможной полнотой как отечественные, так и зарубежные публикации. Срок подготовки реферативного журнала – 3-4 месяца с момента получения первичного документа.

Источники информации «ИНТЕРНЕТ» подразделяются на группы:

По способу представления информации:

Web-страницы – сегодня это основной и наиболее распространенный тип информационных ресурсов в Сети. Этот ресурс представляет собой

страницы так называемого гипертекста, то есть текста, который может содержать в себе ссылки. Взаимосвязанная логически и посредством ссылок совокупность гипертекстовых страниц, расположенная в одном месте, представляет собой единицу, называемую сайтом. Кроме текста web-страница может также содержать информацию в произвольной форме: графической, звуковой, видео и т.д.

Базы данных – также могут иметь интерфейс в Интернете, т.е. могут быть доступны через Сеть. Базы данных могут содержать произвольную информацию: публикации, табулированные данные и т.д.

Файловые серверы – традиционный способ хранения данных в Интернете. Представляет собой компьютеры, часть дискового пространства которых доступна по Сети.

Телеконференции – способ общения людей, имеющих доступ в Сеть. Предназначены для обсуждения каких-либо вопросов или распространения информации. Все телеконференции разбиты по тематическому признаку на рубрики, или группы новостей.

Chart – также, в определенной степени, является источником социологического порядка. Здесь можно выяснить общественное мнение по насущным вопросам или интересующей исследователя тематике.

По языковому признаку: в силу историко-географических причин основным языком в сети Интернет является английский, но практически все языки мира представлены в Сети. Многие сайты поддерживают несколько языков – на выбор пользователя. Российский сегмент Интернет называется «Рунет».

По территориальному признаку: произвольный информационный ресурс в подавляющем большинстве случаев принадлежит какой-либо организации, осуществляющей свою деятельность на определенной территории, и может быть предназначен для аудитории, находящейся преимущественно в пределах другого региона.

По виду и характеру представляемой информации: данный вид разделения информации с практической точки зрения является наиболее важным, т. к. именно информационное наполнение, в конечном итоге, оказывается решающим при отборе и использовании источников.

Тематическая информация – наиболее ценная информация, непосредственно относящаяся к конкретной предметной области.

Научные публикации – статьи, рефераты, обзоры и прочие публикации научного характера, хранящиеся в Интернете.

Рекламная информация – большинство web-сайтов коммерческих компаний, представленных в Сети, носят в той или иной степени рекламный характер. Но данный вид информации может помочь в поисках нужных вам данных.

Справочная информация – разнообразные справочные материалы, ссылки на web-сайты компаний, часто представленные аналогично традиционным «желтым страницам», нормативные базы и т.д.

Новости – эта информация легко доступна по Сети, представляет собой вид «сырой», необработанной информации, которая ценна не столько сама по себе, сколько в контексте прочих событий или в динамике развития, и часто нуждается в последующей обработке и анализе.

## **5.2. МЕТОДЫ ИНФОРМАЦИОННОГО ПОИСКА**

Сплошной метод. Это обследование сплошь и без пропусков все наличие имеющихся пособий и источников. Очевидно, что в современных условиях этот прием трудно реализуем даже для узкой темы. Поэтому часто он рассматривается только в качестве чисто теоретического.

Выборочный метод – более рациональный и реальный путь поиска информации, который сводится к подбору и изучению определенной совокупности источников.

Интуитивный метод. Под ним понимается умение найти нужный материал посредством библиографического чутья, продолжительного опыта.

Типологический, или рецептурный метод, когда использование задач информационного поиска позволяет для каждого конкретного информационного поиска намечать определенный путь и документальные средства (пособия, источники) решения.

Индуктивный метод информационного поиска. Процесс познания с использованием индукции совершается от единичных суждений и фактов к общим правилам и обобщениям, в которых выражается общая закономерность.

Дедуктивный метод информационного поиска. Обратная картина сложится, если мы будем осуществлять поиск дедуктивным методом. Такая направленность мысли может иметь место при исследовании определенной области явлений, создании развернутой научной теории и т.п.

Метод информационного поиска по библиографическим ссылкам. Возможность его использования заключается и в самом характере литературы, состоящей из определенной взаимосвязи фактов, произведений, документов и изданий. Каждое новое литературное произведение – это в известной степени анализ, оценка и обобщение социальной информации, уже имеющейся в литературе.

Изложенные методы позволяют найти необходимую информацию в традиционной информационной среде. [1]

Однако в настоящее время гораздо более важны методы и методика поиска информации в новой информационной среде – Интернете.

По используемым поисковым технологиям информационные системы можно разбить на 4 категории [2]:

1. Тематические каталоги;
2. Специализированные каталоги (онлайновые справочники);
3. Поисковые машины (полнотекстовый поиск);
4. Средства метапоиска.

В Интернете информационно-поисковые системы размещаются на одном или нескольких серверах. В информационно-поисковой системе

собирается, индексируется и регистрируется информация о документах, имеющихся в обслуживаемой системой группе веб-серверов. В документах индексируются все значащие слова или только слова из заголовков.

Тематические каталоги предусматривают обработку документов и отнесение их к одной из нескольких категорий, перечень которых заранее задан. Фактически это индексирование на основе классификации. Индексирование может проводиться автоматически или вручную с помощью специалистов, просматривающих популярные веб-узлы и составляющих краткое описание документов-резюме (ключевые слова, аннотация, реферат).

Поисковые машины (самое развитое средство поиска в Интернете) реализуют технологию полнотекстового поиска. Индексируются тексты, расположенные на опрашиваемых серверах. Индекс может содержать информацию о нескольких миллионах документов.

При использовании средств метапоиска запрос осуществляется одновременно несколькими поисковыми системами. Результат поиска объединяется в общий, упорядоченный по степени релевантности список.

Базы информационных данных могут содержать практически любые виды информации, в том числе в любой комбинации. Для формирования запросов используются специальные информационно-поисковые языки.

Информационно-поисковые системы внутри найденной выборки обычно пытаются расположить документы в порядке их "релевантности", то есть близости к введенному пользователем запросу. Критерием такой близости много и выявление близких "по смыслу" к запросу документов не решает проблемы получения информации при отсутствии релевантного документа. Подобная ситуация достаточно тривиальна, в том числе и потому, что пользователь зачастую ищет документ, который сам собирается написать.

Информационно-поисковые системы фактически являются системами информационного обеспечения и представляют собой базы и банки данных. В качестве объекта в них выступает индивид, организация, отрасль, регион и

т.п. Субъектом информационного обеспечения является специалист-информатик, любой потребитель информации.

### 5.3. ПОИСК С ИСПОЛЬЗОВАНИЕМ ПОИСКОВЫХ МАШИН

Наиболее широко используемым, но в то же время наиболее сложным является метод поиска с использованием поисковых систем. Его широкая распространенность обусловлена тем, что поисковые системы содержат в себе индексы громадного количества сайтов и при правильно сформированном запросе можно сразу же получить ссылки на интересующие ресурсы. Сложность метода состоит в том, что для того, чтобы результат был качественным, необходимо уметь выбрать наиболее подходящие поисковые системы, правильно формулировать запросы к ним, учитывать их особенности и функциональные возможности.

Для эффективного использования поисковых серверов, прежде всего необходим список ключевых слов, организованный с учетом семантических отношений между ними, то есть тезаурус.

Одним из подходов к составлению тезауруса может стать использование законов Ципфа.

Число, показывающее сколько раз встречается слово в тексте, называется частотой вхождения слова. Если расположить частоты по мере убывания и пронумеровать, то порядковый номер частоты называется рангом частоты. Вероятность обнаружения слова в тексте равно отношению частоты вхождения слова к числу слов в тексте. Ципф определил, что если умножить вероятность обнаружения слова в тексте на ранг частоты, то получившаяся величина приблизительно постоянна для всех текстов на одном языке [3]:

$$C = \frac{f \times r}{n},$$

где  $f$  - частота вхождения слов,

$r$  - ранг частоты,

$n$  - число слов

Это значит, что график зависимости ранга от частоты представляет из себя равностороннюю гиперболу.

Исследование вышеуказанных зависимостей для различных текстов показали, что наиболее значимые слова текста лежат в средней части диаграммы, так как слова с максимальной частотой, как правило, являются предлогами, частицами, местоимениями, в английском языке — артиклями (так называемые «стоп-слова»), а редко встречающиеся слова в большинстве случаев не имеют решающего значения. Таким образом, данная особенность может помочь правильно выбрать ключевые слова для проведения поиска информации.

Процедура оптимального выбора ключевых слов, основанная на применении законов Ципфа, заключается в следующем: берут любой текст-источник, близкий к искомой теме, то есть «образец», и анализируют его, выделяя значимые слова. В качестве текста-источника может служить книга, статья, web-страница, любой другой документ. Анализ текста производится в следующем порядке [3]:

1. «стоп-слова» удаляются из текста;
2. вычисляется частота вхождения каждого слова и составляется список, в котором слова расположены в порядке убывания их частоты;
3. выбирается диапазон частот, лежащий в середине списка, и из него отбираются слова, наиболее полно соответствующие смыслу текста;
4. составляется запрос к поисковой машине в форме перечисления отобранных таким образом ключевых слов, связанных логическим оператором OR(ИЛИ) Запрос в таком виде позволяет обнаружить тексты, в которых встречается хотя бы одно из перечисленных слов.

Число документов, полученных в результате поиска по этому запросу, может быть огромно. Однако, благодаря ранжированию документов, то есть расположению их в порядке убывания частоты вхождения в документ слов запроса, применяемому в большинстве поисковых машин, на первых

страницах найденных ресурсов практически все документы должны оказаться релевантными.

Отбор поисковых систем требует установить последовательность использования поисковых машин в соответствии с убыванием ожидаемой эффективности поиска с применением каждой машины.

Всего известно около нескольких сотен поисковых систем, различающихся по регионам охвата, принципам проведения поиска, объему индексной базы, скорости обновления информации, способности искать «нестандартную» информацию и т. д. Основными критериями выбора поисковых систем являются объем индексной базы сервера и степень развитости самой поисковой машины, то есть уровень сложности воспринимаемых ею запросов.

Составление и выполнение запросов к поисковым машинам - наиболее сложный и трудоемкий этап, связанный с обработкой значительного количества информации, большая часть которой обычно является шумовой. На основе тезауруса формируются запросы к выбранным поисковым серверам. После получения первоначальных результатов возможно уточнение запросов с целью отсекаания очевидно нерелевантной информации. Затем производится отбор ресурсов, начиная с наиболее интересных, с точки зрения целей поиска, и данные с ресурсов, признанных релевантными, собираются для последующего анализа.

Языки запроса различных машин поиска в основном являются сочетанием следующих функций [3]:

осуществление поиска документов при помощи операторов булевой алгебры AND, OR, NOT. AND (И) - содержащих все термины, соединенные им, OR (ИЛИ) = искомый текст должен содержать хотя бы один из терминов, соединенных данным оператором; NOT (НЕ) - поиск документов, в тексте которых отсутствуют термины, следующие за данным оператором;

осуществление поиска документов при помощи операторов расстояния, ограничения порядка следования и расстояния между словами. NEAR -

второй термин должен находиться на расстоянии от первого, не превышающем определенного числа слов; FOLLOWED BY - термины следуют в заданном порядке; ADJ - термины, соединенные оператором, являются смежными;

возможность усечения терминов - использование символа \* вместо его окончания термина; позволяет включить в искомый список все слова, производные от его начальной части шаблона;

учет морфологии языка - машина автоматически учитывает все формы данного термина, возможные в языке, на котором ведется поиск;

возможность поиска по словосочетанию, фразе;

ограничение поиска элементом документа (слова запроса должны находиться именно в заголовке, первом абзаце, ссылках и т. д.);

ограничение по дате опубликования документа;

ограничение на количество совпадений терминов;

возможность поиска графических изображений;

чувствительность к строчным и прописным буквам.

Результат запроса, то есть выведенный системой список ссылок на найденные ресурсы, обрабатывается в два этапа. На первом этапе производится отсечение очевидно нерелевантных источников, попавших в выборку в силу несовершенства поисковой машины или недостаточной «интеллектуальности» запроса. Параллельно проводится семантический анализ, имеющий целью уточнение тезауруса для модификации последующих запросов. Дальнейшая обработка производится путем последовательного обращения на каждый из найденных ресурсов и анализа находящейся там информации.

Конечной стадией поиска является анализ ресурсов и сбор искомой информации. Первичный анализ ресурсов может основываться на аннотациях, если они есть, а при их отсутствии - на ознакомлении с информационным наполнением ресурса. Далее информация извлекается с отобранных источников и используется в соответствующих поиску целях.

## Заключение

Понятие информации рассматривалось еще античными философами. До начала промышленной революции, определение сути информации оставалось прерогативой преимущественно философов. С развитием электроники, рассматривать вопросы теории информации стала новая на то время наука кибернетика.

Информация — сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые воспринимают информационные системы (живые организмы, управляющие машины и др.) в процессе жизнедеятельности и работы.

Применительно к компьютерной обработке данных под информацией понимают некоторую последовательность символических обозначений (букв, цифр, закодированных графических образов и звуков и т.п.), несущую смысловую нагрузку и представленную в понятном компьютеру виде. Каждый новый символ в такой последовательности символов увеличивает информационный объём сообщения.

Свойства информации:

достоверность;	понятность;
полнота;	доступность;
ценность;	краткость;
своевременность;	и др.

Информация может существовать в самых разнообразных формах:

в виде текстов, рисунков, чертежей, фотографий;

в виде световых или звуковых сигналов;

в виде радиоволн;

в виде электрических и нервных импульсов;

в виде магнитных записей;

в виде жестов и мимики;

в виде запахов и вкусовых ощущений;

в виде хромосом, посредством которых передаются по наследству признаки и свойства организмов и т.д.

В качестве единицы информации условились принять один бит (англ. bit — binary, digit — двоичная цифра).

Бит в теории информации — количество информации, необходимое для различения двух равновероятных сообщений.

А в вычислительной технике битом называют наименьшую "порцию" памяти, необходимую для хранения одного из двух знаков "0" и "1", используемых для внутримашинного представления данных и команд.

Бит — слишком мелкая единица измерения. На практике чаще применяется более крупная единица — байт, равная восьми битам. Именно восемь битов требуется для того, чтобы закодировать любой из 256 символов алфавита клавиатуры компьютера ( $256=2^8$ ).

Широко используются также ещё более крупные производные единицы информации:

1 Килобайт (Кбайт) = 1024 байт =  $2^{10}$  байт,

1 Мегабайт (Мбайт) = 1024 Кбайт =  $2^{20}$  байт,

1 Гигабайт (Гбайт) = 1024 Мбайт =  $2^{30}$  байт.

В последнее время в связи с увеличением объёмов обрабатываемой информации входят в употребление такие производные единицы, как:

1 Терабайт (Тбайт) = 1024 Гбайт =  $2^{40}$  байт,

1 Петабайт (Пбайт) = 1024 Тбайт =  $2^{50}$  байт.

Можно привести немало разнообразных свойств информации. Каждая научная дисциплина рассматривает те свойства, которые ей наиболее важны. С точки зрения информатики наиболее важными представляются следующие свойства: объективность, полнота, достоверность, адекватность, доступность и актуальность информации.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Информатика. Конспект лекций. - М.: ЭКСМО, 2016.
2. Информатика. Учебное пособие. А.Г. Гейн. - М.: Просвещение, 2017.
3. Босова Л.Л., Цветкова М.С. Информатика. - М.: Астрель, 2017.
4. Информатика и информационные технологии. Под ред. Ю.Д.Романовой. - М.: Эксмо, 2018.
5. Восприятие Информации Человеком. Человеческое Восприятие [Электронный ресурс] - Электрон. дан. - сор. 2015-2016.
6. Информатика. Учебник. Под ред. А.Н. Данчула. - М.: РАГС, 2018.
7. Шрайберг Я.Л., Гончаров М.В. Справочное руководство по основам информатики и вычислительной техники. - М.: Финансы и статистика, 2017.
8. Информатика: Учебник. Под ред. Н.В.Макаровой. - М.: Финансы и статистика, 2017.
9. Могилев А.В., Пак Н.И., Хеннер Е.К. практикум по информатике. Учебное пособие. - М.: Академия, 2016.
10. Морозевич А.Н. Основы информатики. - М.: Новое знание, 2016.
11. Рыжиков Ю.В. Информатика: лекции и практикум. - СПб.: КОРОНАпринт, 2018.
12. Савельев А.Я. Основ информатики. Учебник для ВУЗов. - М.: МГТУ им. Н.Э.Баумана, 2019.
13. Семакин И.Г., Хеннер Е.К. Информатика и ВТ. Учебник. - М.: Бином, 2016.
14. Угринович Н.Д. Информатика и информационные технологии. - М.: ЛБЗ, 2016.
15. Каймин В.А. Информатика: Учебник. - М.: ИНФРА-М, 2017.