

1 курс

ПЛАН – КОНСПЕКТ
проведения лекционного занятия по дисциплине «Информатика»

**Раздел 1. «Информация и информационная деятельность
человека»**

**Тема 1.2:
«Подходы к измерению информации»**

Подготовил: преподаватель
В.Н. Борисов

Вопросы занятия:

1. Подходы к измерению информации (содержательный, алфавитный, вероятностный).
2. Единицы измерения информации.
3. Информационные объекты различных видов.
4. Универсальность дискретного (цифрового) представления информации.
5. Передача и хранение информации.
6. Определение объёмов различных носителей информации.
7. Архив информации.

Время проведения лекционного занятия – 2 часа

Первый вопрос: Подходы к измерению информации (содержательный, алфавитный, вероятностный).

Важным вопросом является измерение количества информации. Количество информации в одном и том же сообщении должно определяться отдельно для каждого получателя, то есть иметь субъективный характер. При этом нельзя объективно оценить количество информации, содержащейся даже в простом сообщении. Поэтому, когда информация рассматривается как новизна сообщения для получателя (бытовой подход), не ставится вопрос об измерении количества информации.

Информация – это содержание сообщения, сигнала, памяти, а также сведения, содержащиеся в сообщении, сигнале или памяти.

Информация – сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределённости, неполноты знаний.

Информация – это понимание (смысл, представление, интерпретация), возникающее в аппарате мышления человека после получения им данных, взаимоувязанное с предшествующими знаниями и понятиями.

Информация, первоначально – сведения, передаваемые людьми, устным, письменным или другим способом (с помощью условных сигналов, технических средств и т.д.); с середины 20 века общенаучное понятие, включающее обмен сведениями между людьми, человеком и автоматом, автоматом и автоматом; обмен сигналами в животном и растительном мире; передачу признаков от клетки к клетке, от организма к организму.

Информация – содержание сообщения или сигнала, сведения, рассматриваемые в процессе их передачи или восприятия; одна из исходных общенаучных категорий, отражающая структуру материи и способы её познания, несводимая к другим, более простым понятиям.

Существует три основные интерпретации понятия "информация".

Научная интерпретация. Информация – исходная общенаучная категория, отражающая структуру материи и способы ее познания, несводимая к другим, более простым понятиям.

Абстрактная интерпретация. Информация – некоторая последовательность символов, которые несут как вместе, так в отдельности некоторую смысловую нагрузку для исполнителя.

Конкретная интерпретация. В данной плоскости рассматриваются конкретные исполнители с учетом специфики их систем команд и семантики языка. Так, например, для машины информация – нули и единицы; для человека – звуки, образы, и т.п.

– в житейском аспекте под информацией понимают сведения об окружающем мире и протекающих в нем процессах, воспринимаемые человеком или специальными устройствами;

– в технике под информацией понимают сообщения, передаваемые в форме знаков или сигналов;

– в теории информации (по К.Шеннону) важны не любые сведения, а лишь те, которые снимают полностью или уменьшают существующую неопределенность;

– в кибернетике, по определению Н. Винера, информация – эта та часть знаний, которая используется для ориентирования, активного действия, управления, т.е. в целях сохранения, совершенствования, развития системы;

– в семантической теории (смысл сообщения) – это сведения, обладающие новизной, и так далее...

Такое разнообразие подходов не случайность, а следствие того, что выявилась необходимость осознанной организации процессов движения и обработки того, что имеет общее название – информация.

Можно выделить следующие подходы к определению информации:

Традиционный (обыденный) – используется в информатике:

Информация – это сведения, знания, сообщения о положении дел, которые человек воспринимает из окружающего мира с помощью органов чувств (зрения, слуха, вкуса, обоняния, осязания).

Вероятностный - используется в теории об информации: **Информация** – это сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределённости и неполноты знаний.

Для человека: **Информация** – это знания, которые он получает из различных источников с помощью органов чувств.

Вся информация, которую обрабатывает компьютер, представлена **двоичным кодом** с помощью двух цифр – **0 и 1**. Эти два символа 0 и 1 принято называть **битами** (от англ. **binary digit** – двоичный знак). **Бит** – наименьшая единица измерения объема информации.

В информатике используются различные подходы к измерению информации:

1. Содержательный подход к измерению информации.

Сообщение, уменьшающее неопределенность знаний человека в два раза, несет для него **1 бит** информации.

Количество информации, заключенное в сообщении, определяется по формуле Хартли:

2. Алфавитный (технический) подход к измерению информации – основан на подсчете числа символов в сообщении.

Формула Хартли определяет количество информации, содержащееся в сообщении длины n . Была предложена Ральфом Хартли в 1928 году как один из научных подходов к оценке сообщений.

Имеется алфавит A , из букв которого составляется сообщение:

$$|A| = m$$

Количество возможных вариантов разных сообщений:

$$N = m^n,$$

где N – возможное количество различных сообщений, m – количество букв в алфавите, n – количество букв в сообщении.

Пример: Алфавит состоит из двух букв «В» и «Х», длина сообщения 3 буквы – таким образом, $m = 2$, $n = 3$. При выбранных нами алфавите и длине сообщения можно составить $N = m^n = 2^3 = 8$ разных сообщений: «ВВВ», «ВВХ», «ВХВ», «ВХХ», «ХВВ», «ХВХ», «ХХВ», «ХХХ» – других вариантов нет.

Формула Хартли определяется:

$$I = \log_2 N = n \log_2 m,$$

где I – количество информации в битах.

Для задач такого рода американский учёный **Клод Шеннон** предложил в 1948 г. другую **формулу определения количества информации, учитывающую возможную неодинаковую вероятность сообщений в наборе**.

Формула Шеннона: $I = -(p_1 \log_2 p_1 + p_2 \log_2 p_2 + \dots + p_N \log_2 p_N)$, где p_i – вероятность того, что именно i -е сообщение выделено в наборе из N сообщений.

Легко заметить, что если вероятности p_1, \dots, p_N равны, то каждая из них равна $1/N$, и формула Шеннона превращается в формулу Хартли.

Помимо двух рассмотренных подходов к определению количества информации, существуют и другие. **Важно помнить, что любые теоретические результаты применимы лишь к определённому кругу случаев, очерченному первоначальными допущениями.**

Второй вопрос: Единицы измерения информации.

Бит в теории информации – количество информации, необходимое для различения двух равновероятных сообщений (типа "орел"- "решка", "чет"- "нечет" и т.п.).

В вычислительной технике битом называют наименьшую "порцию" памяти компьютера, 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 **восми 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.

За единицу информации можно было бы выбрать количество информации, необходимое для различения, например, десяти равновероятных сообщений. Это будет не двоичная (бит), а десятичная (дит) единица информации.

Третий вопрос: Информационные объекты различных видов.

Мы живем в реальном мире, окруженные разнообразными материальными объектами. Наличие информации об объектах реального мира порождает другой мир, неотделимый от сознания конкретных людей, где существует только информация. Этому миру мы даем разнообразные названия. Одно из таких названий – *информационная картина мира*.

Познание реального мира происходит через информационную картину мира. Человек формирует собственное представление о реальном мире, получая и осмысливая информацию о каждом реальном объекте, процессе или явлении. При этом у каждого человека существует своя информационная картина мира, которая зависит от множества факторов как субъективного, так и объективного порядка. Конечно,

большую роль здесь играет уровень образованности человека. Информационные картины мира у школьника, студента и преподавателя будут существенно различаться. Чем объемнее и разнообразнее информация, которую может воспринять человек, тем более красочной получается эта картина. Так, например, информационная картина мира у ребенка совсем не такая, как у его родителей.

Один из способов познания реального мира – это моделирование, которое прежде всего связано с отбором необходимой информации и построением информационной модели. Однако любая информационная модель отражает реальный объект только в ограниченном аспекте – в соответствии с поставленной человеком целью. Отсюда и возникает определенная «ущербность» восприятия мира, если человек изучает его только с одной стороны, определяемой одной целью. Всестороннее познание окружающего мира возможно только тогда, когда существуют разные информационные модели, соответствующие разным целям.

Предположим, мы создали несколько информационных моделей для одного объекта реального мира. Их количество определяется количеством заданных целей. Например, информационные модели нашей планеты у школьника, астронома, метеоролога и геодезиста будут существенно различаться, так как у них разные цели, а значит, и информация, отобранная ими и положенная в основу информационной модели, будет разной.

При разработке модель постоянно сопоставляется с объектом-прототипом для оценки ее соответствия оригиналу.

Что же произойдет, если мы будем иметь дело только с информационными моделями, отстранившись от реального мира? В этом случае отпадает необходимость в понятии адекватности, так как, устранив объект, мы тем самым разорвем виртуальную связь, устанавливающую объектно-модельное отношение. А это значит, что мы полностью погрузимся в виртуальный, несуществующий мир, где циркулирует только информация. Сравнить модель будет не с чем, а значит, отпадет необходимость в самом моделировании.

Таким образом, модель превращается в некий самостоятельный объект, который представляет собой совокупность информации.

Вспомнив понятие *объекта*, которое определяется как *некоторая часть окружающего мира, рассматриваемая как единое целое*, можно высказать предположение, что информационную модель, которая не имеет связи с объектом-оригиналом, тоже можно считать объектом, но не материальным, а информационным. Таким образом, информационный объект получается из информационной модели путем «отчуждения» информации от объекта-оригинала.

Информационный объект – это совокупность логически связанной информации. Тогда информационный мир будет представлять собой множество разнообразных информационных объектов.

Информационный объект, «отчужденный» от объекта-оригинала, можно хранить на различных материальных носителях. Простейший материальный носитель информации – это бумага. Есть также магнитные, электронные, лазерные и другие носители информации.

С информационными объектами, зафиксированными на материальном носителе, можно производить те же действия, что и с информацией при работе на компьютере:

вводить их, хранить, обрабатывать, передавать. Однако технология работы с информационными объектами будет несколько иная, нежели с информационными моделями. Создавая информационную модель, мы определяли цель моделирования и в соответствии с ней выделяли существенные признаки, делая акцент на исследовании. В случае с информационным объектом мы имеем дело с более простой технологией, так как никакого исследования проводить не надо. Здесь вполне достаточно традиционных этапов переработки информации: ввода, хранения, обработки, передачи.

При работе с информационными объектами большую роль играет компьютер. Используя возможности, которые предоставляют пользователю офисные технологии, можно создавать разнообразные профессиональные компьютерные документы, которые будут являться разновидностями информационных объектов. Все, что создается в компьютерных средах, будет являться информационным объектом.

Литературное произведение, газетная статья, приказ – примеры информационных объектов в виде текстовых документов.

Рисунки, чертежи, схемы – это информационные объекты в видеографических документах.

Ведомость начисления заработной платы, таблица стоимости произведенных покупок в оптовом магазине, смета на выполнение работ и прочие виды документов в табличной форме, где производятся автоматические вычисления по формулам, связывающим ячейки таблицы, – это примеры информационных объектов в виде электронных таблиц.

Результат выборки из базы данных – это тоже информационный объект.

Довольно часто мы имеем дело с составными документами, в которых информация представлена в разных формах. Такие документы могут содержать и текст, и рисунки, и таблицы, и формулы, и многое другое. Школьные учебники, журналы, газеты – это хорошо знакомые всем примеры составных документов, являющихся информационными объектами сложной структуры. Для создания составных документов используются программные среды, в которых предусмотрена возможность представления информации в разных формах.

Другими примерами сложных информационных объектов могут служить создаваемые на компьютере презентации и гипертекстовые документы. Презентацию составляет совокупность компьютерных слайдов, которые обеспечивают не только представление информации, но и ее показ по заранее созданному сценарию. Гипертекстом может быть назван документ, в котором имеются гиперссылки на другие части этого же документа или другие документы.

Четвёртый вопрос: Универсальность дискретного (цифрового) представления информации.

Чтобы сообщение было передано от источника к получателю, необходима некоторая материальная субстанция – носитель информации. Сообщение передаваемое с помощью носителя назовём сигналом. В общем случае сигнал – это изменяющийся во времени физический процесс. Характеристика которая используется для представления сообщений называется параметром сигнала.

В случае когда параметр сигнала принимает последовательное во времени конечное число значений сигнал называется дискретным, а сообщение передаваемое с помощью такого сигнала – дискретным сообщением. Информация, передаваемая источником в этом случае тоже является дискретной. Если же источник вырабатывает непрерывное сообщение то соответствующая информация называется непрерывной.

Пример дискретного сообщения – процесс чтения книги (информация в письменном виде), пример непрерывного сообщения человеческая речь.

Непрерывное сообщение может быть представлено с помощью непрерывной функцией $f(t)$, заданной на некотором отрезке $[a, b]$.

Непрерывное сообщение преобразовать в дискретное (такая процедура называется дискретизацией). Область определения функции разбивается на равные промежутки dt (равномерная дискретизация), и непрерывная функция заменяется импульсной, причём с одинаковой частотой. Получается дискретное представление непрерывной функции, точность которого можно неограниченно улучшать путем уменьшения длин отрезков разбиения области значений аргумента. Таким образом, любое сообщение может быть представлено как дискретное.

Возможность дискретизации непрерывного сигнала с любой желаемой точностью принципиально важна с точки зрения информатики. Компьютер – цифровая машина, т.е. внутренне представление информации в нем дискретно. Дискретизация входной информации позволяет сделать её пригодной для компьютерной обработки.

Представление информации в двоичной системе счисления.

По своему назначению компьютер — универсальное, программно-управляемое автоматическое устройство для работы с информацией. Из свойства универсальности следует то, что компьютер осуществляет все три основных типа информационных процессов: хранение, передачу и обработку информации. Современные компьютеры работают со всеми видами информации: числовой, символьной, графической, звуковой. Информация, хранимая в памяти компьютера и предназначенная для обработки, называется данными.

Для представления всех видов данных в памяти компьютера используется двоичный алфавит. Однако интерпретация последовательностей двоичных цифр для каждого вида данных своя. Еще раз подчеркнем, что речь идет о внутреннем представлении данных, в то время как внешнее представление на устройствах ввода-вывода имеет привычную для человека форму.

Представление числовой информации. Исторически первым видом данных, с которым стали работать компьютеры, были числа. Первые ЭВМ использовались исключительно для математических расчетов. В соответствии с принципами Джона фон Неймана, ЭВМ выполняет расчеты в двоичной системе счисления. Вопрос о внутреннем (машинном) представлении чисел рассмотрим несколько подробнее, чем это делается в учебниках.

Структурные единицы памяти компьютера — бит, байт и машинное слово. Причем понятия бита и байта универсальны и не зависят от модели компьютера, а размер машинного слова зависит от типа процессора ЭВМ. Если машинное слово для данного компьютера равно одному байту, то такую машину называют 8-разрядной (8

бит); если машинное слово состоит из 2 байтов, то это 16-разрядный компьютер; 4-байтовое слово у 32-разрядных ЭВМ.

Пятый вопрос: Передача и хранение информации.

Передача, обработка и хранение информации происходит в форме сигналов или знаков.

Сигналы: можно разделить на несколько типов:

- по физической природе (электромагнитный, световой, тепловой, звуковой, биохимический);
- по способу восприятия (зрительный, слуховой, осязательный, вкусовой, болевой, физиологический).

Знаками можно считать алфавит любого языка, знаки языка жестов, любые коды и шифры, ноты и т.д.

Рассмотрим по отдельности передачу, обработку и хранение информации.

Обработка информации — вся совокупность операций (сбор, ввод, запись, преобразование, считывание, хранение, уничтожение, регистрация), осуществляемых с помощью технических и программных средств, включая обмен по каналам передачи данных. При современном развитии программного обеспечения существует множество различных программных средств обработки информации, написанных на разных языках программирования на основе выше перечисленных методов.

Обработка информации подразумевает переработку информации определённого типа (текстовой, звуковой, графической и др.) и преобразования её в информацию другого определённого типа. Так, например, принято различать обработку текстовой информации, изображения (графики, фото, видео и мультипликация) и звуковой информации (речь, музыка, другие звуковые сигналы). Использование новейших технологий обеспечивает их комплексное представление. При этом человеческое мышление может рассматриваться как процесс обработки информации.

Передача информации.

Информация представляется и передается в форме последовательности сигналов, символов. От источника к приёмнику сообщение передается через некоторую материальную среду. Если в процессе передачи используются технические средства связи, то их называют каналами передачи информации (информационными каналами).

К ним относятся телефон, радио, ТВ. Органы чувств человека исполняют роль биологических информационных каналов.

Передача информации — физический процесс, посредством которого осуществляется перемещение знаков (сведений, способных предоставлять информацию) в пространстве или осуществляется физический доступ субъектов к знакам. Передача информации — заблаговременно организованное техническое мероприятие, результатом которого становится воспроизведение информации, имеющейся в одном месте (так называемый источник информации) или в другом месте (приёмник информации).

1.2.1 Основные термины и предмет теории информации

В любой системе информация представлена в виде *сообщений* — совокупности знаков, либо непрерывных сигналов, являющихся переносчиком информации.

Дискретные сообщения формируются в результате последовательной выдачи источником сообщений отдельных элементов - *знаков*. При этом все множество возможных различных знаков называют *алфавитом сообщения*, а размер множества - *объемом алфавита*.

Непрерывные сообщения в свою очередь не разделены на элементы, а описываются непрерывными сигналами - функциями времени, принимающими значения из непрерывного континуума.

Среда, по которой передаются сообщения между источником и приемником сообщений называется *каналом связи*, либо *каналом передачи информации*. При этом преобразование сообщения в сигнал, подходящий для передачи по заданному каналу связи, называется *кодированием* (в широком смысле слова); обратную операцию называют *декодированием*.

Во время прохождения сообщения по каналу связи в данном канале могут действовать мешающие воздействия - помехи (как внешние, так и внутренние).

Итак, исходя из введенных терминов, определим предмет теории информации:

Теорией информации исследуются информационные системы (кибернетические системы с ярко выраженными процессами передачи, хранения и преобразования информации), подчиняющиеся следующим постулатам:

- 1 Источник сообщения осуществляет выбор сообщения из некоторого множества (с определенными вероятностями выбора каждого из сообщений).
 - 2 Сообщения могут передаваться по каналу связи в закодированном виде с возможностью однозначного декодирования на приемной стороне.
-

- 3 Сообщения следуют друг за другом, при этом количество сообщений может быть сколь угодно большим.
- 4 Сообщение считается принятым при успешно осуществленной (и однозначной!) операции декодирования. При этом не имеет значения, сколько времени прошло с момента передачи сообщения и какова вычислительная сложность операций кодирования и декодирования.
- 5 Количество информации является математической абстракцией; не зависит от смыслового содержания сообщения, его эмоционального воздействия, полезности и отношения к реальной действительности.

1.2.2 Количественная мера информации

Рассмотрим источник дискретных сообщений (дискретный источник информации). Пусть каждое отдельное i -е сообщение представляет собой информационный символ, выбираемый из ансамбля U размерности m с определенной для каждого элемента ансамбля вероятностью появления:

$$U = \begin{pmatrix} u_1 & u_2 & \dots & u_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix} \quad (1.1)$$

Представим информацию как меру неопределенности источника сообщений. Так, детерминированные (представляющие собой сингулярный случай 1.1 при $m = 1$ сигналы не несут в себе полезной –информационной нагрузки (величина количества информации, обозначим его $I = 0$).

Возможно перечислить следующие естественные условия к I , как к количественной характеристике меры неопределенности:

- 1 Функция $I(m)$ должна быть неотрицательной и монотонно возрастающей (за исключением введения в ансамбль вырожденных элементов с вероятностью появления $p = 0$).
- 2 Функция I_X для любых сообщений X должна обладать свойством аддитивности:

$$I(m_{X1}) + I(m_{X2}) = I(m_{X1} + m_{X2}) \quad (1.2)$$

- 3 Количество информации I должно зависеть от вероятностей появления элементов ансамбля.
-

- 4 Количество информации I должно зависеть от вероятностей появления элементов ансамбля. Действительно, интуитивно ясно, что более редкое событие несет в себе большее количество информации.

Базовым условиям **1** и **2** удовлетворяет функция $I = \log m$; при этом указанная формула достаточно легко расширяется на случай сообщения из n символов, каждый из которых выбирается из ансамбля размерности m . Действительно, в этом случае разнообразие N сообщений дискретного источника определяется как число перестановок с неограниченными повторениями из m по n : $N = m^n$. Таким образом, результирующая формула, полученная Ральфом Хартли в 1928 позволяет определить количество информации в виде следующей функции:

$$I = \log N = \log m^n = n \log m. \quad (1.3)$$

В формуле 1.3 возможно использовать произвольное основание логарифма; от выбранного основания зависит единица измерения количества информации¹. Наиболее распространенные основания - e , 10 и 2; соответствующие единицы измерения - нат, дит и бит.² В современной вычислительной технике, в связи с двоичной природой абсолютного большинства современных ЭВМ, в качестве безусловного стандарта принят *бит*.

Хранение информации.

Под хранением обычно понимают запись данных на некоторые накопители данных (запоминающие устройства) с целью их (данных) дальнейшего использования.

Хранение информации осуществляется записью на носителях и накопителях информации. Носителем данных может являться любой материальный предмет, используемый человеком для записи, хранения, чтения и передачи информации (например, книги, диски, фотографии, flash-карты, облачные хранилища и так далее). Накопителями считаются приспособления позволяющие хранить и дополнять информацию.

Хранение является одной из основных операций, осуществляемых над информацией, с целью обеспечения её доступности в течение некоторого промежутка времени.

Шестой вопрос: Определение объёмов различных носителей информации.

Носители информации характеризуются информационной емкостью, то есть количеством информации, которое они могут хранить. Наиболее информационно

емкими являются молекулы ДНК, которые имеют очень малый размер и плотно упакованы. Это позволяет хранить огромное количество информации (до 10^{21} битов в 1 см^3), что дает возможность организму развиваться из одной единственной клетки, содержащей всю необходимую генетическую информацию.

Современные микросхемы памяти позволяют хранить в 1 см^3 до 10^{10} битов информации, однако это в 100 миллиардов раз меньше, чем в ДНК. Можно сказать, что современные технологии пока существенно проигрывают биологической эволюции.

Однако если сравнивать информационную емкость традиционных носителей информации (книг) и современных компьютерных носителей, то прогресс очевиден:

- Лист формата А4 с текстом (набран на компьютере шрифтом 12-го кегля с одинарным интервалом) - около 3500 символов
- Страница учебника - 2000 символов
- Гибкий магнитный диск – 1,44 Мб
- Оптический диск CD-R(W) – 700 Мб
- Оптический диск DVD – 4,2 Гб
- Флэш-накопитель - несколько Гб
- Жесткий магнитный диск – сотни Гб

Таким образом, на дискете может храниться 2-3 книги, а на жестком магнитном диске или DVD - целая библиотека, включающая десятки тысяч книг.

Созданную или полученную каким-либо образом информацию хранят в течение определённого времени, в течение которого её временно или долговременно содержат на различных носителях электронных данных. Если информация представляет интерес для её создателей или правообладателей, то им приходится создавать электронные архивы.

Седьмой вопрос: Архив информации.

Электронный архив - это файл, содержащий один или несколько файлов в сжатой или несжатой форме и информацию, связанную с этими файлами (имя файла, дата и время последней редакции и т.п.).

Электронные архивы позволяют в любой момент времени извлекать из них необходимые данные для дальнейшего их использования в различных ситуациях (например, для обновления или восстановления утерянных данных). Такие архивы

называют страховочными копиями. Их используют в случае утраты или порчи основной машиночитаемой информации, а также для длительного её хранения в месте, которое защищено от вредных воздействий и несанкционированного доступа. Как правило, компьютерными архивами информации являются электронные каталоги, базы и банки данных, а также коллекции любых видов электронной информации.

Для обеспечения надёжности хранения и защиты данных рекомендуют создавать по 2–3 архивные копии последних редакций файлов. В случае необходимости осуществляется разархивирование данных.

Процесс сжатия файлов называется **архивированием**. **Разархивирование** - это процесс точного восстановления электронной информации, ранее сжатой и хранящейся в файле-архиве. Процесс восстановления сжатых файлов – **разархивированием**.

Архиваторы – это программы (комплекс программ) выполняющие сжатие и восстановление сжатых файлов в первоначальном виде. Другие названия архиваторов: утилиты - упаковщики, программы - упаковщики, служебные программы, позволяющие помещать копии файлов в сжатом виде в архивный файл.

Основные возможности архиваторов:

- просмотр содержания архива и файлов, содержащихся в архиве
- распаковка архива или отдельных файлов архива;
- создание простого архива файлов (файлов и папок) в виде файла с расширением, определяющим используемую программу-архиватор;
- создание самораспаковывающегося архива файлов (файлов и папок) в виде файла с пусковым расширением EXE;
- создание многотомного архива файлов (файлов и папок) в виде группы файлов томов заданного размера (раньше - в размер дискеты).

Самым распространенным архиватором считается **WinRar**. может открывать следующие типы файлов: RAR, ZIP, CAB, ARJ, LZH, ACE, 7-Zip, TAR, GZip, UUE, BZ2, JAR, ISO, Z Базовыми функциями являются:

- извлечение файлов из архивов;
- создание новых архивов;
- добавление файлов в имеющийся архив;
- создание самораспаковывающихся архивов;
- создание распределенных архивов на носителях малой емкости (дискета);
- тестирование целостности архивов;
- полное или частичное восстановление поврежденных файлов;
- защита архивов от просмотра и несанкционированной модификации.