

## **Содержание:**

# **Введение**

Глобальная сеть - совокупность компьютеров, расположенных на больших расстояниях друг от друга, а также система каналов передачи связи: средств коммуникации (переключения), обеспечивающих соединение пользовательских коммуникационных систем и обмен данными между ними.[1]

Глобальные сети (WideAreaNetworks, WAN) создаются крупными телекоммуникационными компаниями для оказания платных услуг абонентам.

Интернет - мировая глобальная компьютерная сеть. Она составлена из разнообразных компьютерных сетей, объединенных стандартными соглашениями о способах обмена информацией и единой системой адресации. Интернет использует протоколы семейства TCP/IP. Они хороши тем, что обеспечивают относительно дешевую возможность надежно и быстро передавать информацию даже по не слишком надежным линиям связи, а также строить программное обеспечение, пригодное для работы на любой аппаратуре. Система адресации (URL-адреса) обеспечивает уникальными координатами каждый компьютер (точнее, практически каждый ресурс компьютера) и каждого пользователя Интернета, создавая возможность взять именно то, что нужно, и передать именно туда, куда нужно.

В 1994 году началась революция – World Wide Web. Всемирная паутина World Wide Web (WWW) соткана из Web-страниц, которые содержат в себе разную информацию в зависимости от тематики Web сайта. В основу Web был положен гипертекст (hypertext) – метод связывания блоков, или «страниц», данных, придуманный еще в шестидесятых годах. Однако только в девяностых годах Бернерс-Ли и его сотрудники перенесли концепцию гипертекста в Internet, создав HTTP – Hypertext Transfer Protocol (протокол передачи гипертекста). С появлением HTTP родился и World Wide Web.

Сегодня в Internet существуют миллионы Web – сайтов. Можно получать доступ к информации по различным темам, открыть в Web свой бизнес. Более того, в Web можно найти и сведения о нем самом и о тех технологиях, на которых он основан.

Таким образом, данная курсовая работа носит актуальный характер. Целью работы является обзор языков гипертекстовой разметки. Для достижения данной цели требуется решить следующие задачи:

1. провести анализ понятия «гипертекст»;
2. проанализировать историю развития гипертекста;
3. провести анализ моделей гипертекста;
4. изучить виды языков гипертекстовой разметки документов – синтаксис, структуру документов, основные элементы.
5. провести анализ будущего Web-программирования.

## **Глава 1. Гипертекст**

### **1.1 Понятие гипертекста**

Гипертекст - текст со вставленными в него словами (командами) разметки, ссылающимися на другие места этого текста, другие документы, картинки и т.д. Во время чтения такого текста (в соответствующей программе, его обрабатывающей и выполняющей соответствующие ссылки или действия) вы видите подсвеченные (выделенные) в тексте слова. [3] Если наехать на них курсором и нажать клавишу или на кнопку (глаз) мышки, то высветится то, на что ссылалось это слово, например, другой параграф той же главы этого же текста. [1] В WWW по ключевым словам можно попасть в совершенно другой текст из другого документа, войти в какую-нибудь программу, произвести какое-либо действие и т.д. В Internet в контексте WWW можно получать доступ к чему угодно, к telnet, e-mail, ftp, Gopher, WAIS, Archie, USENET News и т.п. В WWW можно ссылаться на данные на других машинах в любом месте сети, тогда при активации этой ссылки эти данные автоматически передадутся на исходную машину и вы увидите на экране текст, данные, картинку, а если провести в жизнь идею мультимедиа, то и звук услышите, музыку, речь. [2] Это слегка напоминает Gopher, но фактически это принципиально другое и новое. В Gopher имеется жесткая структура меню, по которой вы двигаетесь, как вам угодно. Эта структура не зависит от того, что вы делаете, какой документ пользуете и т.д. В WWW вы двигаетесь по документу, который может иметь какую угодно гипертекстовую структуру. Можно свободно организовать структуры меню в гипертексте. Имея редактор гипертекстов, можно создать любую структуру рабочей среды, включая документацию, файлы, данные,

картины, программное обеспечение и т.д., и это не будет новое программное обеспечение, а просто гипертекст.

Гипертекстовая технология – это представление текста в виде многомерной иерархической структуры типа сети.[1]

Гипертекст формируется в результате представлений текста как ассоциативно связанных блоков информации. Ассоциативная связь – это соединение, сближение представлений, смежных, противоположных, аналогичных. Гипертекст значительно отличается от обычного текста. Обычные (линейные) тексты имеют последовательную структуру и предусматривают их чтение слева направо и сверху вниз. [2]

Простейший пример гипертекста - это любой словарь или энциклопедия, где каждая статья имеет отсылки к другим статьям этого же словаря. В результате читать такой текст можно по-разному: от одной статьи к другой, по мере надобности, игнорируя гипертекстовые отсылки; читать статьи подряд, справляясь с отсылками; наконец, пуститься в гипертекстовое плавание, то есть от одной отсылки переходить к другой. [2]

Концепция гипертекста достаточно проста. Есть база данных, в базе данных находятся объекты. Объекты это, чаще всего, небольшие текстовые разделы, посвященные тому или иному вопросу. Специальные механизмы и правила позволяют компьютеру поддерживать ссылки из одних текстовых фрагментов в другие. Человек или программный агент может устанавливать новые связи между текстовыми фрагментами. Система текстовых фрагментов или файлов с такой организацией получила название "гипертекст". [3]

Гипертекст изначально создавался как среда поддерживающая взаимодействие нескольких людей. Культовая работа Ваннавера Буша "As we may Think", в которой он описал устройство Мемех, была связана с проблемами взаимодействия коллективов ученых после Второй Мировой Войны, когда стало ясно, что существующие системы плохо поддерживают коллективную мыслительную деятельность. Система Мемех, по своей сути, представляла систему для обмена "мемами" - элементарными единицами культурной эволюции. Гипертекст изначально мыслился создателям как система общественной деятельности. Группа взаимосвязанных сообщений образовывала сеть, и эта гипертекстовая сеть документов поддерживала социальную сеть отношений между сообществом авторов коллективного гипертекста.[3]

Использование гипертекста позволяет фиксировать отдельные идеи, мысли, факты, а затем связывать их друг с другом, двигаясь в любых направлениях, определяемых ассоциативными связями.

С развитием компьютерных средств мультимедиа гипертекст начал превращаться в более наглядную информационную форму, получившую название гипермедиа — эта информационная форма содержит не только текст, но и графику, видеоинформацию и звуки.

Обработка гипертекста открыла новые возможности освоения информации, качественно отличающиеся от традиционных способов.

Вместо поиска информации по соответствующему поисковому ключу гипертекстовая технология предполагает перемещение от одних объектов информации к другим с учетом их смысловой, семантической связанности.

Обработке информации по правилам формального вывода в гипертекстовой технологии соответствует запоминание пути перемещения по гипертекстовой сети.[2]

Гипертексты обладают определенной семантической (смысловой) сетевой структурой. При многократном просмотре, если гипертекст используется как учебник, эта структура будет сильно влиять на структуру знаний пользователя по изучаемому вопросу. Поэтому при построении гипертекстовых систем следует уделять внимание не только тому, как разбить исходный текст на части, но и тому, насколько пользователю будет понятно, легко и удобно работать с этими частями текста.

Структурно гипертекст состоит из информационного материала, тезауруса гипертекста, списка главных тем и алфавитного словаря.[2]

Информационный материал подразделяется на информационные статьи, состоящие из заголовка статьи и текста. Заголовок содержит тему или наименование описываемого объекта.

Информационная статья содержит традиционные определения и понятия, должна занимать одну панель и быть легко обозримой, чтобы пользователь мог понять, стоит ли ее внимательно читать или перейти к другим, близким по смыслу статьям.

Текст, включаемый в информационную статью, может сопровождаться пояснениями, примерами, документами, объектами реального мира.[2]

Тезаурус гипертекста – это автоматизированный словарь, отображающий семантические отношения между лексическими единицами дескрипторного информационно-поискового языка и предназначенный для поиска слов по их смысловому содержанию.[2]

Тезаурус гипертекста можно представить в виде сети: в узлах находятся текстовые описания объекта (информационные статьи), ребра сети указывают на существование связи между объектами и на тип родства.[2]

Список главных тем содержит заголовки всех справочных статей, для которых нет ссылок типа род – вид, часть – целое. [2]

Алфавитный словарь включает в себя перечень наименований всех информационных статей в алфавитном порядке. [2]

К основным элементам гипертекстовой технологии относятся: [3]

- информационный фрагмент;
- тема;
- узлы;
- ссылки.

Информационный фрагмент гипертекста может представлять собой линейную последовательность строк текста, рисунок, видеофрагмент, аудиофрагмент.[5]

Тема содержит краткое название информационного фрагмента. Информационный фрагмент может состоять целиком из множества тем либо включать в себя одну или несколько тем наряду с прочей информацией.

Узлом в гипертексте называется информационный фрагмент, из которого возможен переход к другим информационным фрагментам гипертекста.[5]

Ссылка представляет собой слово, фразу или набор фраз, с помощью которых осуществляется переход от одного узла к другому. Ссылки могут быть референтными или организационными.

Референтные ссылки — это наиболее типичный вид ссылок в гипертекстах. Они, как правило, имеют два конца, обычно это направленные связи, хотя большинство гипертекстовых информационных систем поддерживает и обратное движение по

ссылке. Исходный конец референтной ссылки называется «источник». Логически это отдельная точка или область в тексте. Другой конец называется «назначением» — это определенная точка или область в гипертексте. С источником ссылки связывается некоторая пометка, указывающая наличие ссылки, — она показывает имя ссылки, обычно изображается в виде последовательности символов и высвечивается как отдельная единица текста. Например, при щелчке по термину появится информационный фрагмент, разъясняющий значение этого термина.[5]

Организационные ссылки устанавливают явные связи между двумя точками гипертекста и отличаются от референтных тем, что поддерживают иерархическую структуру в гипертексте. Организационные ссылки связывают узел-родитель с узлами-сыновьями и, таким образом, формируют древовидный подграф в рамках общего гипертекстового сетевого подграфа. Такие ссылки часто соответствуют отношению «быть частным случаем», и по этой причине операции над этими ссылками (при построении гипертекста) отличаются от операций над референтными ссылками.[6]

Область применения гипертекстовой технология очень широка. Это издательская деятельность, библиотечная работа, обучающие системы, разработка документации, законов, справочных руководств, баз данных, баз знаний и т.д. [6]

Наиболее известным примером гипертекста являются веб-страницы — документы HTML (язык разметки гипертекста) как они размещаются в Сети.

Современные программы разработки Web-серверов, такие как MS FrontPage или Web Pen для Windows, дают возможность даже новичку без всякого штудирования учебников легко создавать готовые странички. При этом специалист по созданию Web-сайтов, называемый Web-мастером, берет готовые файлы (тексты, таблицы, графику, базы данных, звук, анимацию, видеофильмы, программы) и с помощью кнопок и команд меню оформляет страницы сайта. Подобные программы, выполняя команды инструментальных и операционного меню, формируют гипертекст WWW-сервера.

Исходные текстовые, табличные и графические и другие объекты включаются в Web-site посредством тегов (tag = ярлык, этикетка). Тег - это последовательность символов, задающая

1). положение объекта на странице сайта,

2). внешний вид объекта или

3). связь данной страницы с другими страницами этого сайта, а также с любым другим сервером.

Тег называют также управляющим маркером, флагом. Программы типа Web Pen сами расставляют теги, поэтому пользователь таких программ может не знать языка разметки гипертекста (HTML = HyperText Markup Language).

## **1.2 История возникновения гипертекстовой разметки**

История гипертекста богата и переменчива, поскольку гипертекст не столько какая-то новая идея, сколько находящаяся в эволюции концепция возможного применения компьютера. В разработку идеи гипертекста внесли свой вклад много людей, и каждый из них, видимо, представлял себе нечто отличное от других.

Компьютерному гипертексту предшествует ручной, один из вариантов которого – традиционное использование карточек. Такие карточки можно нумеровать и снабжать взаимными ссылками. Их часто распределяют по рубрикам, т. е. им придается иерархическая организация (в некотором ящике или пакете). Удобство таких карточек состоит в том, что, имея небольшой размер, они разбивают записи на малые куски. Пользователь может легко реорганизовать картотеку с учетом новой информации. Но, конечно, с увеличением объема такой картотеки, работать с ней становится все труднее.[13]

Другой вариант ручного гипертекста – это справочная книга, например словарь и энциклопедия. Статьи или определения, даваемые в таких книгах, содержат явные ссылки друг на друга, последовав за этими ссылками, читатель получает более богатую информацию. Каждой такой книге можно поставить в соответствие сеть с текстовыми узлами и связями-ссылками.[13]

Многие века существуют документы, где внутренние перекрестные ссылки и отсылки к другим документам образуют значительную долю содержания. Таковы, например, Талмуд с его обильным использованием аннотаций и встроенным в текст комментарием, а также сочинения Аристотеля, в которых ссылки на другие источники играют огромную роль.[13]

Еще один важный пример – печатные издания Библии. В них текст каждой из ее книг-частей делится на главы, а те, в свою очередь, на стихи. Главы пронумерованы внутри каждой книги, стихи – внутри каждой главы. Стих может состоять из части грамматического предложения, одного целого предложения или нескольких фраз. В подлинном библейском тексте этого деления нет. Оно было сделано учеными-богословами для облегчения ссылок и цитат. К примеру, деление Нового Завета на стихи, ныне общепринятое, восходит к XVI веку. Согласно этому делению, Новый Завет (27 книг-частей) состоит из 260 глав и, суммарно по всем главам, из 7942 стихов. Гипертекст возникает здесь потому, что в современных изданиях Библии текст идет в сопровождении так называемых "параллельных мест", обычно в виде ссылок на полях. Каждая такая ссылка ставит в соответствие стиху, который идет рядом, "параллельные" стихи из этой же или других книг-частей Библии (даются координаты этих стихов). В комментариях объясняется, что "параллельные места" указывают на тождественные события и "созвучные выражения". Библейский текст, по существу, превращен в гипертекстовую сеть на узлах-стихах. Сеть имеет огромные размеры: если взять лишь ее новозаветную часть с ее внутренними "параллелями", то получилось бы почти 8 тыс. гипертекстовых узлов (из них, правда, многие не имели бы связей).[7]

Все эти примеры относят появление гипертекста к далеким временам. Сейчас, однако, немало специалистов, которые считают, что об истинном гипертексте можно говорить лишь в том случае, когда перемещение по связям поддерживается компьютером.

В 1945 году в своей статье "Как мы можем думать" ("As We May Think") Ванневар Буш высказал идею машины для просмотра и пополнения записями документов, записанных на пленке. [7]

Эта машина, получившая название "Memex", никогда не была построена, но она содержала идею, которую позже назвали гипертекстом. [7]

Информация, считал Буш, должна храниться в виде пленочных микрофильмов. Поэтому Memex имела устройство для чтения микрофильмов и устройство для записи микрофильмов с помощью процесса сухой фотографии. То есть Memex принципиально не была цифровым компьютером. [8]

Буш представлял машину в виде письменного стола с экранами для отображения информации и клавиатурой для управления. Внутри стола размещалось хранилище микрофильмов и механизм доступа к ним (рисунок 1)[8]



INCLUDEPICTURE "http://inf.1september.ru/2007/09/30-1.gif" \\* MERGEFORMATINET

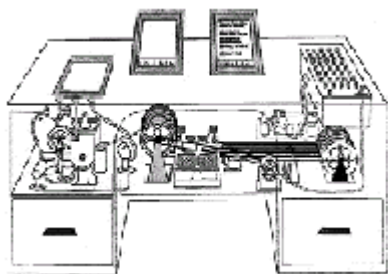


Рисунок 1 – Схема Memex

Однако главная особенность и новизна Memex состояла не в способах хранения информации, а в способе доступа к ней. Буш предложил механизм перекрестных ссылок, аналогичный тому, который используется в современном гипертексте. Ссылки Буш предлагал записывать во вспомогательных полях в теле документа, а в качестве средства навигации использовать нечто похожее на современный мультимедийный шлем.[8]

Фактически система Xanadu явилась прообразом web-пространства, но сам Тед Нельсон отзывался отрицательно о Web и HTML, считая, что работа Бернерса-Ли является сильно упрощенным вариантом его работы. [8]

Первую работающую гипертекстовую систему (она называлась NLS, от *oN Line System*) продемонстрировал в 1968 году Дуг Энгельбарт (Doug Engelbart, 1925 г.р.). [8]

В 1989 году Тим Бернерс-Ли (Tim Berners-Lee, 1955 г.р.), работая над внутренней сетью организации CERN (Европейский совет по ядерным исследованиям в Женеве), предложил глобальный гипертекстовый проект, ныне известный как Всемирная паутина. [7]

В начале 1990-х Тим Бернерс-Ли и его коллеги создали язык HTML, на котором записываются современные гипертексты, то есть такие документы (распределенные по Сети), которые благодаря гиперссылкам можно просматривать по контексту. После создания языка HTML Web-технологии начали приобретать более четкий характер и бурное разностороннее развитие, поддерживаемое многочисленными компаниями в будущем, такими как: Google, Microsoft, Mozilla Foundation и пр. Историческая последовательность изложена в приложении А.[8]

## 1.3 Модели гипертекста

В самом общем виде модель гипертекста характеризуется механизмом связей, узлами-объектами и пользовательским интерфейсом – способом взаимодействия человека с узлами и связями.

Узлы. Узел – важнейшее понятие гипертекстовых систем, так как в них именно в форме узлов хранится и представляется пользователю информация. Характеристики узла, существенные для гипертекстового пользователя (читателя или автора), – это тип информации, которая может быть сохранена (текст, таблица, графика, звук и др.), и вместимость, объем каждого узла.[13]

Некоторые гипертекстовые системы поддерживают только текст (например, NLS и ZOG), другие – таблицы и графику (HyperTIES, HyperCard, Guide); есть механизмы интерпретации различных видов информации (например, видео) в рамках гипертекстовой сети (Intermedia, NoteCards).

Важно отметить, что просто реализация возможности показывать узлы мультимедиа здесь недостаточна. Каждый новый тип информации должен быть полностью интегрирован с гипертекстовой сетью системы, для чего необходимо разработать методы создания связей между, например, кадрами видео и текстом. В Intermedia и NoteCards это достигается сравнительно легко: они разрабатывались в расширяемых средах (объектно-ориентированный язык Си и Lisp соответственно), отсюда – легкая встраиваемость в систему модулей, создающих гипертекстовую функциональность для новых видов информации.

Что касается объемов узлов, то, например, такие системы, как ZOG и HyperCard, поддерживают только узлы жестко фиксированной величины (объемом с экран), в то время как другие обладают более гибкими возможностями.

Узел на экране обычно дается в своем отдельном окне. Одновременно может быть открыто лишь несколько окон-узлов. Выполняются стандартные операции оконных систем. Отметим, что сами оконные системы, а также манипулятор мышь, были изобретены Дагласом Энгельбартом, для нужд его пионерской системы.[13]

Объем узла, разбиение информации на узлы – серьезная проблема для автора гипертекста, который должен думать о восприятии смыслового содержания узла читателем. Предпочтительны узлы, обладающие внутренним смысловым единством (внутренней когерентностью).[13]

Связи. Способ реализации связей имеет ключевое значение, поскольку именно связи обеспечивают "нелинейное ветвление" – сердцевину гипертекстовой

функциональности.[13]

При описании связей в гипертексте часто используется понятие anchor (буквально, якорь) – это слово или фраза, которые подсвечиваются на экране и воспринимаются как точки начала или конца связи.

Возможны два варианта статуса гипертекстовых связей. В первом связи являются самостоятельными объектами, которыми пользователь может манипулировать напрямую; во втором связи спрятаны в системе (возможно, как часть текста) и проявляются, только когда пользователь каким-либо образом задействует их.[13]

– пример со связями статуса. У нее хранятся от документов, на ссылаются. Эти также быть с помощью вида – значение, что осуществлять их по запросу .

Связи статуса возможным в систему браузера – , с помощью сеть ( из узлов и ) отображается на . Без хранения связей говорить о структуре . Поэтому такие , как NoteCards, и реализации HAM ( Absnract ), могут функциональность, для обработки как графа.

второго ("спрятанные") – просто адреса для и существуют в момент их . Например, в HyperCard – это , содержащие "иди к № 42106" ( что-то ). Однако могут и не таких , они будут вставлены в код системы. Так как нет соотношения узлом и его связей в сети, не имеет для воздействия на как целое. система и графическое последних карточек, не информация об маршрутах, по можно .[13]

В NLS, HyperTIES, ZOG узла для по связи как часть или имя отдельно кнопки. связи по природе , они позволяют сквозной по документу, но при возможно в тупике (в , из которого не связи).

Еще одной характеристикой является то, как она к узлам – ли она узлы как или фрагментами. Еще по Memex , что, когда соединяются как , у читателя возникать с пониманием существования связи от узла, если у эта связь не . В свою , в узле, произошел , читатель вынужден информацию, делает осмысленным.

гибко эти были в Intermedia. В системе могут из любого одного и заканчиваться в фрагменте . В Notecards и , как и Memex, относится к узлу-карточке. У NLS, ZOG, и Guide связи служат слова и .[13]

пользователя. гипертекстовых тесно с исследованиями, , изобретениями в человеко-машинного . Иконки, , оконные , манипуляторы мышью и другое люди из сообщества. И изобретателем в области пионер Даглас . Он жив до сих пор и очень

в компьютерном . Несколько лет назад учреждена Энгельбарта, на ежегодной конференции Hypertext за представленную .[13]

А гипертекстовых на развитие идей в человеко-машинного видно по факту, что регулярных конференций по компьютера с обязательно раздел, гипертексту.

вернемся к вариантам интерфейса в обозреваемых системах. Эти с их особенностями поиски и , которыми шли .

Все системы легко по документу в его последовательности – путем линейного (Intermedia, ), либо по дереву структуры , используя "следующий " или "возврат к " (ZOG, ).

Значительно интерес вопросы: как может связь ( ее визуальное ) – и как эту связь активизировать? В системах эти решаются . В Memex о связях хранилась на отведенном пространстве как сигнал : что-то из текста с дополнительной . Все остальные позволяли информацию так, чтобы его могли метками : у одних ( , HyperTIES и ) это была соответствующего слова или , у других ( , Intermedia, ) – значки или кнопки, , как значки (примечания), читателя о дополнительной , соотнесенной с .

Возможность той или иной достигалась, , в Memex и ZOG каждой связи с клавиатуры ( , как это делается для ). В Intermedia, , HyperCard и вместо надо указать на и "кликнуть" . У HyperTIES другая (клавиатурная): связь-точка подсвеченной и эта перемещается от связи к под управлением стрелок. нужная выбрана, активизирует ее с некоторого ключа.

В NLS операция состоит из частей – и активизировать. подход управлять как , так и последующим (так подход " and actions").[14]

связь и выбрана, осуществляет гипертекстовый по связи к информации. мгновенный больше компьютеру, чем . Читатель, при просмотре , встретив (или ) на другую или статью, не чтение, а его до конца и потом по ссылке работу, на ссылались. В с компьютером переходы по ведут к в глубину с стопкой ( ) отложенных , к которым будет . Такие статьи расти, как ком, увеличивая на память и внимание . При этом системы высвечивать один на экране в момент.

При переходах по у пользователя контекст информации. Он , с чего свои и где находится, когда исследовать сеть. Это недостаток – быстрая пользователя.

побочный и недостаток, называемый перегрузкой, с необходимостью множество (выбирать, кнопки, переходы, назад) для полезной. В гипертекстовых были средства с этими. Например, и Intermedia многим находятся на одновременно, в разрабатывались браузеры, в отдельную структуру связей (и глобальные). При этом опасность пользователя во открытых с разнообразной.[14]

Таким, введение в текстовый гипертекстовых () связей функциональность по сравнению со статично-линейного. Однако особенности систем написание тяжелым, а сам гипертекст – для восприятия.

нелинейности, переходов и информации – тому, что "больше, чем", угрожают, описанные. В некоторых гипертекст, , является подходящим, чем текст.

и навигация. с недостатками насчитывает не десяток лет, и полностью их нельзя, сделано в направлении.

Прежде, в гипертексте существует прокладывания пути (, трейла) в гипертекстовых. Путь – это из узлов и, которые пользователь. Это выработал еще Буш, имея в аналогии с в мозгу: "Человеческий ум ассоциативно, поняв, он сразу за следующее, что, подсказывается мыслей в с некоторой паутиной, которые ячейками".

Обычно системы возможности по связям и различные, помогающие пути, т. е. навигацию в. Эти методы, не свободу, направлены на дезориентации и когнитивной, от которых читатель и которые получили "проблема" (Navigation).

У большинства систем помощь следующие.

Локальная. Это картинка связей и, непосредственно с текущим. Она может графической (, в виде) или текстовой (список). карты читателю и помогают связь.

Глобальная. Это графическое полной из узлов и. Ввиду с отображением числа, такие мало для реальных объемом сотни. Больше с ними разработчики (Янкелович, ван Дам). и глобальные в реальных системах "графическими".[13]

(бэктрекинг). узлы и текущего сохраняются и возможность в предыдущие.[13]

(проложенные). Это хранящиеся, которые проходить по. Подобное при создании учебников или. Для больших значение возрастает.

Поиск (в гипертекстовых ). Используются все в области поиска: запросы, кий поиск, запросов и др.

Фильтры. Это ограничения навигации задаваемого подмножества и связей. подмножества видами ( ) и могут сохранены для повторного .[13]

. Список слов, или узлов, по алфавиту, , автору, и т. д. Индексы автором и тот недостаток, что не учитывают зрения на то, как он хочет гипертекст.

Закладки. может (пометить) текущую , чтобы к ней позже.

Для больших и гипертекстовых растет использовать , интеллектуальную в навигации.

В – об одном подходе, отечественным и направленном на проблемы в гипертексте.

В подходе и система в сети навигацию, , что навигационная должна подобна дискурсу (, предложения вместе смысловое ). Это означает, что локальной между узлов в должна и некая связность, разворачиванию темы, начальным тропы.[5]

когерентная , конечно, не всегда. , человек, навигацию в , может перемещаться по в надежде на интересные , получить на какой-то вопрос, найденные рассматриваются по , вне контекста.[6]

есть и виды активности при с гипертекстом, где навигация . Это – браузеринг по тематике, или какого-то по материалу, в сети, или же чернового документа из сети.[6]

навигация в системе – российской . В этой локальные по связям в тропе-трейле под так называемым , следящем за (тематической) . Этот базируется на подтем дискурса.

реализована в интерактивного , которое в тропу-дискурс.

разработки опираются на таких психолого-лингвистов, как ван , Кинч и . Результаты представлены в на гипертекстовой международной "Восток-Запад" '93.

## Глава 2. Языки разметки

популярный на день гипертекстовой HTML, был специально для информации, в сети , и явля одной из составляющих WWW. С использованием модели способ

разнообразных ресурсов в стал упорядочен, а получили механизм и просмотра информации.

HTML( Markup ) - язык гипертекстовой , который в время в World Web. Изначально как язык для научной и документацией. языка занимается W3C ( Consortium).

HTML является версией общего разметки - (Standart Markup ), который был ISO в качестве а еще в 80-х . Этот предназначен для других разметки, он допустимый тэгов, их и внутреннюю документа. за правильностью дескрипторов при помощи набора , называемых , которые программой при разборе . Для каждого документов свой правил, грамматику языка . С помощью можно структурированные , организовывать , содержащуюся в , представлять эту в некотором формате. Но в некоторой сложности, использовался, в , для описания других (наиболее из которых HTML), и приложения с SGML- напрямую.

## 2.1 SGML

— метаязык, на котором определять язык разметки для . SGML — разработанного в году в IBM GML (Generalized Language).

Изначально был разработан для использования машинно-читаемых документов в правительственных и проектах. Он использовался в и издательской , но его сложность его широкое для повседневного .

Основные документа : [16]

1. SGML-декларация — , какие и ограничители появляться в ;
2. Document Type Definition — определяет синтаксис разметки. DTD включать определения, , как символные ;
3. Спецификация , относится к — также ограничения , которые не быть внутри DTD;
4. SGML-документа — по мере, быть элемент.

SGML множество синтаксической для использования приложениями. SGML-декларацию, даже от использования скобок, этот считается , так называемым *reference* .

Пример SGML:

<quote ">

typically like <>this</>

</quote>

стандартизован ISO: « 8879:1986 processing—Text and systems—Standard Markup (SGML)».

HTML и XML произошли от . HTML — это SGML, а — это подмножество , разработанное для процесса разбора . Другими SGML SGML Docbook (документирование) и «Z » (типография и ).

## 2.1.1 Описательная

Система разметки коды , просто названия для частей . Коды, , как <para> или \} просто часть и утверждают про нее: " элемент - " или "это - начатого списка" и т.д. , система разметки , какая должна в конкретной документа: " вызвать PARA с 1, b и x", или "сдвинуть границу на 2см , правую -- на 2см , пропустить и встать на левую ", и т.д. В SGML , необходимые для документа с целями ( , для его форматирования) отделяются от разметки, внутри . Обычно они вне документа в процедурах или .

При описательной, а не , разметке и тот же документ обрабатывать программами, из которых применять правила к тем частям , которые она важными. , программа содержимого совершенно сноски в тексте, как программа может и собирать их для печати в каждой . С одними и же частями могут разные обработки. , одна может имена и географические для создания или базы , а другая, тем же текстом, печатать собственные отличающегося .

## 2.1.2 Типы

SGML понятие документа и, как , определения документа ( type , DTD). Тип документа определяется его частями и их . Например, отчета констатировать, что он из заголовка, , автора, за следуют и один или абзацев. Все, что не заголовка, в с этим определением, не является, так же, как не им последовательность , за которой аннотация, вне от того, такие похожи на для читателя-человека.



Раз имеют типы, использовать программу, анализатором (), для проверки, утверждающего принадлежность типу. проверяет, что все, требуемые документа, на деле и расположены в порядке. Что важно, документы и того же могут одинаковым. Можно программы, знание документа, таким, могут в более манере.

## 2.1.3 Данные

цель SGML в том, чтобы транспортабельность документов из аппаратной и среды в без потери. Два описанных свойства эту задачу на уровне; свойство -- на строк (символов), из составляется. SGML универсальный строковой (string), то есть, машинно-независимый обозначить, что последовательность в документе заменяться при его некоторой последовательностью. очевидное этого -- обеспечение номенклатуры; , и более, -- противодействие известной различных систем наборы друг, или способ в системе все графические, необходимые для приложения, использования обозначений символов. , определенные механизмом, называются (entities). В слово (entity) специальный: оно означает часть документа, ко всяческим структуры. может строка или целый текста. Для его в документ конструкция, как ссылка на (entity).

## 2.1.4. SGML

Этот описывает и согласованный разметки или структурных текста, SGML. Он описывает, способы предлагает для правил, возможные комбинации единиц в текстах.

В стандарте для текстовых, рассматриваемых как компоненты, термин (element). типам даются названия, но не предлагает способов значение типа, кроме его к другим элементов. То, все, что можно про элемент, (например) <>, -- это то, что его экземпляры встречаться (а и не встречаться) элементов <farble>, и что он раскладываться (а и не раскладываться) на типа <>. Следует, что стандарт совершенно не семантика элементов: она от приложения (В момент работа по (с использованием SGML) стандартного"я семантики и стилей (document and semantics language, ").) Дело SGML-совместимых разметок (, как описанный в Руководство) -- осмысленные идентификаторов и документировать их использование в текстов. Это -- из целей документа. От выбора элементов, их функцию, технический для названия элемента: **идентификатор** (identifier), или Gl.

В размеченном (*экземпляре* , *document* ) каждый должен явно или отмечен образом. предоставляет разных это сделать, часто из них -- вставить (*tag*) в начале (*открывающая* , *start-tag*) и еще -- в конце (*закрывающая* , *end-tag*). открывающей и меток для выделения в тексте, так же, как скобки или используются в пунктуации. , элемент может отмечен в так: [16]

*слышала!*</quote> ясно показывает ...

Как данный , открывающая имеет вид <>, где открывающая скобка начало метки, "" -- идентификатор элемента, и угловая означает метки. метка аналогичный вид, за того, что за угловой стоит косой , так что соответствующая метка </название>. (На деле , используемые в ограничителей ( скобки, черта, знак) переопределяться, но использовать , приведенные в описании.)

Модели элемента:

Элемент быть (*empty*), то , не содержать вообще ; элемент содержать текст. , однако, одного будут содержаться ( ) внутри другого . [16]

использования , устанавливающих, элементы быть в другие, очень свойством . Не переходя к разбору правил, попытаться , как размеченный образом может обработан с целями. индексирующая может только элементы для генерации заголовков, или , использованных в стихотворения; программа может пустые между , возможно, с красной первую каждой , или вставляя строфы. части стихотворения набираться способами. сложная программа соотносить знаков со строфовыми и разделами. , желающие следствия разделов или строк, редактором стихотворения, это сделать меняя меток. И, , представленный текст быть с одного на другой и любой (или ) , понимающей внесенных в меток, всяких и трансляций, обычно для файлов процессоров.

### **Определение документов : DTD**

Правила вышеописанных -- шаг в создании спецификации SGML или **определения документа**, сокращаемого как **DTD**. При DTD дизайнер может произвольно или сколь гибкую . Нужно компромисс удобством простым и сложностью реальных . Это особенно , когда правила к уже существующим : дизайнер иметь туманное об изначальном или смысле текстов, и непротиворечивых , касающихся их , может очень . С другой , когда новый , например, для в некоторую базу , то чем точнее правила, тем они могут выдержаны. в случае уже существующего может смысл ограничивающий правил, к определенному текста или , касающейся , -- хотя бы как

проверки этого или гипотезы. помнить, что определение документа интерпретацией . Не существует DTD, охватывающего все о тексте, может удобно одни DTD для конкретных анализа.

В настоящее SGML всего там, где основным является структуры . Например, при технической весьма , чтобы и подразделы соответствующим вложены, перекрестные были , и так далее. В ситуациях к относятся как с материалу, к применяется определенный правил. , как говорилось , использование правил также упростить аккуратной элементов и ограниченных . Делая правила , исследователь свою по разметке и электронного , в то же время интерпретацию и значимые кодируемого . [16]

минимизации

часть задает **минимизации** для . Эти правила , обязаны ли открывающая и метки для появления элемента. Они вид пары , разделенных , первый из относится к , а второй -- к метке. В случаях присутствовать или или буква O; означает, что должна , а буква -- что она может опущена. Так, в примере элемент, **<line>**, иметь метку. элементы **<>** и **<anthology>** также и закрывающую . [16]

содержимого

часть описания, в круглые , называется **содержимого** , потому что она , что могут экземпляры . Содержимое либо в других , либо при специальных слов. несколько зарезервированных , из которых часто -- #PCDATA. Это от *parsed data*( символные ), и оно означает, что элемент включать разрешенные данные. представить SGML в виде наподобие дерева, с предком (в нашем , это будет **<>**), то почти , если по ветвям вниз (, от **<anthology>** к **<>**, **<stanza>**, **<>** или **<title>**), мы к #PCDATA. В примере так **<title>** и **<>**. Так как в их модели указано #PCDATA и не никаких элементов, то они не содержать элементы.

Обозначения

Вышеприведенное для **<stanza>** , что строфа из одной или строк. Оно **обозначение** (*occurence* ) -- плюс -- для того, раз может элемент, в модели . В синтаксисе есть три включения, представленных плюс, знаком и . (Так же, как и , эти знаки формальные и могут переопределены SGML .) Знак означает, что элемент встречаться или более раз; знак , что может не более элемента; означает, что может или , или появляться и более раз. Так, бы модель для **<stanza>** (LINE\*), бы допустимы без строк, так же, как и с чем одной . Если бы она (LINE?), то строфы бы

тоже , но ни одна не могла бы более чем строку. **<poem>** в устанавливает, что **<>** не может больше заголовка (но не иметь ни ) и что оно должно как минимум **<stanza>** (и иметь ). [16]

Модель (TITLE?, +) содержит одного . Поэтому дополнительно порядок, в эти элементы (**<>** и **<stanza>**) появляться. Это определяется (*group*) -- -- использованным ее компонентами. три возможных , обычно запятой, чертой и "&". (Так же, как и обозначения , связки в стандарте имена и быть соответствующим описанием.)

Запятая , что оба компонента, она соединяет, встречаются в , указанном в содержимого. "&" указывает, что , которые он , должны оба, но в произвольном . Вертикальная означает, что встречаются один из , которые она . Если бы в примере заменить на "&", то мог бы появляться или строфами , или в его конце (но не строфами). ее заменить на черту, то могло бы или из заголовка, или из строф -- но не из и другого.

модели

До сих пор в примере каждой содержимого или единственным , или #PCDATA. можно, , определять содержимого, в компонентами списки , объединенные . Такие , известные как **модели** (*groups*), также обозначениями и, в свою , быть связками. продемонстрировать эти , расширим наш так, чтобы нестрофовые стихов. Для классифицируем на **строфовые** (*stichic*), **двустихия** (*stichic*), и **белые** (*stichic*) или **??** (*stichic*). стих просто из (игнорируем возможность абзацев) определяется как **<1>**, за которой **<line2>**.

<!ELEMENT O O (line1, line2) >

**<line1>** и **<2>** (которые , например, сделать изучение рифмования) в точности ту же содержимого, что и элемент **<>**. Они, следовательно, разделять и то же описание. В ситуации указать **названий** (*group*) в первого единого элемента, а не последовательность , отличающихся используемыми . Группа -- это список GI, связками и в круглые : [16]

<! (line | line1 | line2) O O (#PCDATA) >

элемента **<>** теперь изменить так, включить все три :

<!ELEMENT +) ) >

То есть, состоит из заголовка, за следует или несколько , либо или несколько , либо или несколько . Отметьте между определением и :

```
<!ELEMENT - O ( ?, ( stanza | | line)+ ) >
```

вариант, обозначение у группы, а не у элемента группы, одному состоять из строф, или белого .

Таким можно довольно модели, структурную различных текстов. В примере мы строфовый , в котором рефрен ( ). Он может из повторений или быть текстом, не на стихотворные . Рефрен появляться в начале или как необязательное после строфы. Это выразить содержимого следующей:

```
<! refrain - - ( # | line+ ) >
```

```
<! poem - O ( title?,
```

```
( ( line+)
```

```
| ( refrain?, ( , refrain? )+ )
```

То есть, состоит из заголовка, за следует или строк, или группа, рефреном, за идет или несколько групп, член состоит из с необязательным . Этому отвечает *рефрен - - строфа -* , так же, как и *строфа - - строфа -* . А последовательность *- рефрен - - строфа* ему не , так же, как и *строфа - - рефрен -* . Среди условий, этой , -- требования, в стихотворении хотя бы строфа, оно не состоит из строк, и при наличии и и строфы они именно в этом .

## 2.2 HTML

— стандартный язык разметки во Всемирной паутине. Большинство веб-страниц описание на языке (или XHTML). HTML браузерами; полученный в интерпретации текст на экране компьютера или устройства.

Язык является SGML (стандартного языка ) и соответствует стандарту ISO [15]

HTML был британским учёным Тимом Бернерсом-Ли в 1986—1991 годах в стенах ЦЕРНа в Женеве в Швейцарии. создавался как для обмена и технической , пригодный для людьми, не специалистами в вёрстки. HTML справлялся с сложности путём небольшого структурных и семантических — дескрипторов. также называют

«тегами». С HTML легко относительно , но красиво документ. упрощения документа, в внесена гипертекста. Мультимедийные возможности добавлены .[15]

язык был задуман и как средство и форматирования без их привязки к воспроизведения (). В идеале, с разметкой должен был без и структурных воспроизводиться на с различной оснащённостью ( экран компьютера, экран , ограниченный по экран телефона или и программы воспроизведения ). Однако применение очень от его изначальной . Например, тег <> предназначен для в документах , но часто и для оформления элементов на . С течением основная платформонезависимости HTML принесена в современном в мультимедийном и

## 2.2.1 Структура документа

HTML — это теговый язык разметки , то есть документ на HTML собой элементов, начало и каждого обозначается пометками, тегами. Регистр, в набрано имя , в HTML не имеет. могут *пустыми*, то не содержащими текста и данных (, тег перевода <br>). В случае не указывается тег. Кроме , элементы иметь , определяющие их свойства (, размер для тега <>). Атрибуты в открывающем . Вот пример разметки :

```
<p>Текст двумя - открывающим и .</p>
```

```
<a ">Здесь содержит href.</>
```

А вот пример элемента: <>

Каждый , отвечающий HTML версии, начинаться со декларации HTML <!>, которая выглядит так:

```
<!DOCTYPE PUBLIC "-// HTML 4.01/" "http://www.w3.org/TR/html4/strict.dtd">
```

Если эта не указана, то корректного документа в браузере труднее.

обозначается и конец тегами <> и </html> . Внутри тегов находится заголовка (<>) и тела (<>) документа.

## 2.2.2 Элементы

и их параметры к регистру. То <A HREF="http://example.com"> и <a "> означают и то же.

В последних HTML у каждого огромное количество необязательных — обычно не 15. Приведем

## Гиперссылки

`<a href="" target="_">название </a>`

- Атрибут задает адреса , на который ссылка.
- — имя файла или Internet, на необходимо .
- название — название ссылки, будет в браузере, то показываться тем, кто на страницу.
- — задает окна или , в котором открыт , на который ссылка. значения :
  - `_top` — документа в окне;
  - `_` — открытие в новом ;
  - `_self` — документа в фрейме;
  - `_` — открытие в родительском .

Значение по : `_self`.

Тот же используется для так называемых «якорей» (anchor), могут использоваться в , направленных на определённую страницу. :

```
<!DOCTYPE >
```

```
<html>
```

```
<>
```

```
<meta = "utf-8">
```

```
<>Якорь документа</>
```

```
</head>
```

```
<>
```

```
<p><a name=""></a></p>
```

```
<p>текст</p>
```

```
<p><a = "#top"></a></p>
```

</body>

</>

Аналогичным якорь сделать на , находящуюся на веб-странице или на сайте: там, направлена , должен <a name="xxx"></a>, а там, идёт , к значению добавляется решётки и якоря.

блоки

- <H1> ... </1>, <H2> ... </2>, ... ,<H6> ... </6> — заголовки 1, 2, ... 6 .  
Используются для частей (заголовков 1 — крупный, 6 — мелкий).
- <P> — абзац. в конце поставить </>, но это не обязательно.
- <> — новая . Этот тег не (то есть не тега </>)
- <HR> — линия
- <> ... </BLOCKQUOTE> — . Обычно сдвигается .
- <PRE> ... </> — режим (preformatted ). В этом текст в рамку и не форматируется (то теги, </PRE>, , и переводы ставятся там, и там, где они есть в документе).
- <> ... </DIV> — (обычно для применения CSS)
- <SPAN> ... </> — строка ( используется для стилей CSS)

текста

- <EM> ... </> — логическое (обычно *курсивным* )
- <STRONG> ... </> — усиленное ударение ( отображается шрифтом)
- <I> ... </> — выделение *курсивом*
- <B> ... </> — выделение жирным
- <U> ... </U> — текста
- <S> ... </> (или <STRIKE> ... </> )— зачёркивание
- <BIG> ... </> — увеличение
- <SMALL> ... </> — уменьшение
- <BLINK> ... </> — мигающий . Внимание! тег не работает в браузере Internet Explorer 5 и ниже без JavaScript
- <MARQUEE> ... </> — сдвигающийся по текст.
- <> ... </SUB> — текст. , Н<SUB>2</SUB>O текст H<sub>2</sub>O.
- <> ... </SUP> — текст. , E=mc<> создаст E=mc<sup>2</sup>.
- < параметры> ... </> — задание шрифта. У тега следующие :



- COLOR= — задание . Цвет быть в шестнадцатеричной как #rrggbb ( 2 шестнадцатеричные задают компоненту, 2 — зелёную, 2 — синюю) или .
- FACE= задание гарнитуры
- SIZE= задание шрифта. от 1 до 7: стандартный по 3. Есть способов стандартный .
- SIZE=+ или SIZE=- — изменение шрифта от . Например, +2 размер на 2 стандартного.

[15]

<UL>

<LI> элемент </LI>

<LI> элемент </LI>

<LI> элемент </LI>

</UL>

список

- элемент
- элемент
- элемент

вместо <> (*Unordered* — нумерованный ) поставить <> (*Ordered* — нумерованный ), список нумерованным:

1. элемент
2. элемент
3. элемент

У тегов параметры:

= "тип"

где тип — : в <UL> —

1. square —
2. circle —
3. disk — : по умолчанию

а в <> — цифр или

- А или а (латинскими) — буквенный : соответственно или строчными
- I или i — римские : соответственно или строчными

## Объекты

- EMBED — различных : не-HTML и media-файлов
- — вставка Java-апплетов
- — вставка .

## Изображения

- IMG — вставка . Этот тег не .
  - SRC — имя или URL
  - ALT — альтернативное имя (, если в запретить картинки)
  - — краткое изображения ( при наведении на картинку)
  - , HEIGHT — (если не с истинными картинки, то «растянется» или «») )
  - ALIGN — параметры текстом (, middle, , left, )
  - VSPACE, — задают вертикального и пространства изображения

:

```
<IMG SRC=url ALT="текст" = "текст" = "размер (, %)" HEIGHT=" (пикс, %)">
```

можно ссылкой:

```
<A =url ><IMG SRC=url></A>
```

## 2.3 XML

**XML** — рекомендованный Консорциумом Всемирной паутины язык разметки, представляющий свод синтаксических правил. XML для хранения данных ( существующих файлов баз данных), для информацией программами, а также для на его основе специализированных разметки (, XHTML), иногда *словарями*. XML упрощённым языка SGML.

Целью XML было совместимости при структурированных между системами информации, особенно при таких через Интернет. , основанные на XML (, RDF, RSS,

MathML, XHTML, SVG), сами по формально , что позволяет изменять и документы на основе словарей, не их семантики, то есть не смыслового элементов. особенностью XML является так называемых пространств имён (). [17]

## 2.3.1 XML

- XML() — это формат, понятный и компьютеру;
- XML Юникод;
- в XML могут описаны структуры — такие как , списки и ;
- XML — это самодокументируемый , который структуру и полей как и значения ;
- XML имеет определённый и требования к , что позволяет ему простым, и непротиворечивым.
- XML широко для хранения и документов как , так и офф-лайн:
- XML — , основанный на стандартах;
- структура XML для описания любых документов;
- XML собой текст, от лицензирования и ограничений;
- XML не от платформы;
- XML подмножеством (который с 1986 ). Уже накоплен опыт с языком и специализированные .
- XML не накладывает на расположение на строке

## 2.3.2 XML

- Синтаксис XML .
  - Размер XML существенно бинарного тех же данных. В оценках этого принимают за 1 (в 10 раз).
  - XML документа больше, чем в альтернативных форматах данных ( JSON) и в форматах оптимизированных для случая .
  - Избыточность XML повлиять на приложения. стоимость , обработки и данных.
  - Для количества не нужна вся синтаксиса XML и использовать более и производительные
- XML не содержит в язык типов . В нём нет понятий « чисел», «», «дат», « значений» и т.д.
- модель , предлагаемая XML, по сравнению с моделью и графами
  - не иерархических (например, ) требует усилий

- Дейт , что «...XML попыткой изобрести базы ...» (в 1980-е иерархические данных вытеснены базами ).
- Пространства XML сложно и их сложно в XML парсерах
- другие, сходными с XML , текстовые данных, обладают высоким чтением (YAML, JSON, SweetXML). в последнее очень распространение формат fb2.

## 2.3.3 Построения XML

В общем XML-документы удовлетворяют требованиям:

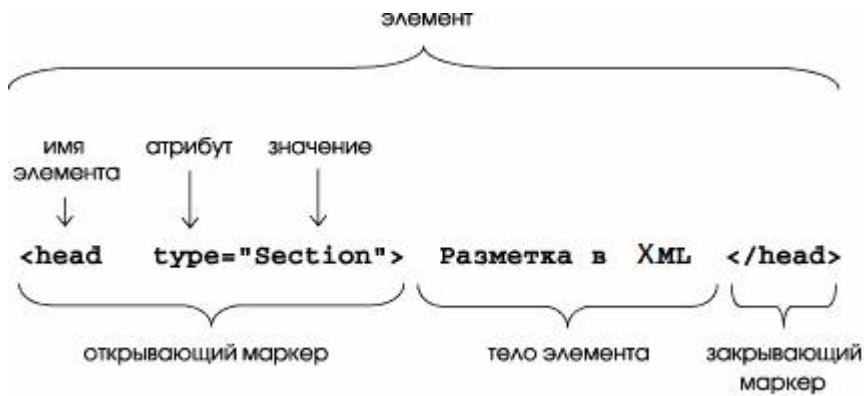
1. В заголовке помещается XML, в котором язык документа, его версии и информация
2. открывающий тэг, некоторую данных в обязательно иметь закрывающего "", т.е., в отличие от , нельзя закрывающие
3. В XML учитывается символ
4. Все атрибутов, в определении , должны заключены в
5. Вложенность в XML строго , поэтому следить за следования и закрывающих
6. Вся информация, между и конечными , рассматривается в XML как и поэтому все символы ( т.е. пробелы, строк, не игнорируются, как в )

Конструкции [17]

Содержимое XML- представляет набор , секций , директив , комментариев , , текстовых . Общая представлена на 2.

Пример -документа:

```
<>
<flower>>
<flower>>
<flower>>
</conservatory>
INCLUDEPICTURE "http://nknaromanova.narod.ru/sgml.files/image002.jpg" \*
MERGEFORMATINET
```



## 2- Общая структуры XML-

### Элементы

Элемент - это единица - документа. слово в в тэги `<> </flower>` , мы непустой , называемый `<>`, содержимым является . В общем в качестве элементов выступать как какой-то , так и другие , элементы , секции , инструкции по , комментарии , - т.е. любые XML- .

Любой элемент состоять из , конечного и данных, ними . Например, фрагменты являться :

`<flower>>`

`<city>>`

Набором элементов, в документе, его структура, и все иерархическое . Плоская данных с использованием в сложную систему с возможных между .

Производя в поиск в документе, клиента опираться на , заложенную в его - используя документа. То , если , требуется нужный в нужном , используя фрагмент , то необходимо просмотреть конкретного `<university>`, внутри элемента `<>`. Поиск при , естественно, гораздо эффективен, чем нужной по всему .

В XML документе, как , определяется бы один , называемый и с него начинают документа. В примере элементом `<country>`

В случаях могут и уточнять тех или иных документа, по определяя и ту же информацию и тем предоставляя этого сведения о использования данных. , прочитав `<city></city>` мы догадаться, что в этой документа о городе, а вот во `<restaurant></restaurant>` - о .

В случае, элемент не содержимого, то нет данных, он должен , то он называется . Примером элементов в могут такие HTML, как `<>` `<hr>`, `<>`. Необходимо помнить, что и конечные пустого как бы объединяется в , и надо ставить черту закрывающей скобкой (, `<empty/>`;) )

тегов и можно и по-русски. HTML , сколь тщательная и интернационализация аспектов , претендующего на роль в . Поэтому XML позаботились, в , о том, чтобы в тегов и можно пользоваться не латинскими , но и кириллицей, и вообще символами из Unicode, считаются "" хотя бы в языке или письменности.

CDATA для того, обозначить документа, не должны как разметка. CDATA со строки '`<![['` и заканчивается `']]>`'. самой не должна строка

Секция :

```
<example> <![ [ <aaa>bb&cc </example>
```

С XML-документа и его XML-процессором произвести простую того, что является оформленным. Для на этой специализированных необходимы средства этих . XML поддерживает два подобных : *определения документа* ( type , DTD) и (XML ).

## Глава 3. Гипертекст

В настоящий актуальным технологий для сайтов языки 5 и CSS3.

### 3.1 Язык разметки HTML5

HTML5 – это версия HTML.

5 вводит новых элементов и , которые типичное разметки на веб-сайтах. Некоторые из — семантические для использования блочных (`<>`) и строчных (`<>`) элементов, , `<nav>` ( навигации по ), `<footer>` ( относится к части или последней HTML ) или `<audio>` и `<video>` `<object>`. устаревшие , которые было в HTML 4.01, были , включая оформительские , такие как `<>` и `<center>`, чьи выполняются с каскадных таблиц стилей. Также в веб снова внимание на скриптов DOM( , javascript).[18]

HTML5 не базируется на SGML, на подобие его . Однако он был обратно с обычным более версий . В HTML5 новая строка, выглядит как типа в SGML, `<! html>`,

соответствующий режим . С 5 января года 5 также в себя Web 2.0, ранее отдельной WHATWG.

В дополнение к разметки 5 устанавливает API, может использоваться с JavaScript. DOM расширены и используемые задокументированы. добавлены API, например:

- элемент холст для непосредственного рисования в 2D. См. Canvas 2D API 1.0;
- контроль над медиафайлов, может , например, для субтитров с [34];
- хранение данных в браузере;
- документа: на страницу выбор ( <input ">) или перетаскиванием ( )
- Drag-and-drop: предоставляет событий для элемента DOM, как появление и в его зоне, которым может пользователя о действиях и перетаскиваемого , содержащего , имя, тип, размер и изменения;
- историей ;
- тип MIME и регистрация протокола;
- микроданные.

Не все перечисленные включены в W3C HTML5, они есть в WHATWG . Немного технологий, не являются ни одной из , следуют . W3C публикует для них отдельно:

- геолокация;
- данных SQL для Web, база (больше не );
- Индексированная данных ( ) API, индексирование по ключ-значение ( — WebSimpleDB);
- API, дескриптор файлов и ими;
- Работа с . Этот API для того, обеспечить информации со клиента без базами ;
- Запись в , использование API для в файл из приложения'

5 – на данный это уже стандарт в . В дополнение к применяются таблицы третьего – CSS3, о которых речь в подразделе.

## 3.2 Таблицы

Каскадные стили удобным , позволяющим разметку быстро, и красиво.

CSS создателями веб-страниц для цветов, шрифтов, расположения блоков и аспектов внешнего этих . Основной разработки CSS разделение логической веб-страницы ( производится с HTML или других языков разметки) от внешнего этой (которое производится с формального языка CSS). разделение увеличить

документа, большую и возможность его представлением, а уменьшить и повторяемость в содержимом. того, CSS представлять и тот же документ в стилях или методах вывода, таких как экранное представление, печатное представление, чтение голосом (специальным голосовым браузером или программой чтения с экрана), или при выводе устройствами, использующими шрифт Брайля.

Современное поколение каскадных таблиц стилей – CSS3, обладает огромным множеством возможностей для создания анимированного сайта без использования Javascript.

### **3.3 Выводы о будущем языков гипертекстовой разметки документов.**

Использование стека HTML5+CSS3 для верстки сайтов приобрело большую популярность. Данная комбинация технологий идеально подходит для разметки современных сайтов. Однако консорциум W3C на данном этапе не останавливается. Языки HTML5+CSS3 будут пока и дальше поддерживаться, развиваться, но, скорее всего, только в ближайшие 5-10 лет. Это связано с решением W3C создать новые, «идеальные» языки для разработки сайтов. Следующим поколением станут: HTML6 и CSS4. HTML6 предполагает создание одностраничных веб-приложений без использования технологий Javascript. Соответственно синтаксис данного языка будет значительно отличаться от предыдущего поколения – HTML5. Язык CSS4 будет создан с учетом синтаксиса нового HTML6. Но пока оба языка – HTML6 и CSS4 ещё в активной разработке и будут доступны для массового применения нескоро. Хотя, у W3C уже имеются готовые заготовки, которые веб-разработчики могут опробовать уже сейчас, но не все браузеры будут их поддерживать.[19]

## **Заключение**

В ходе курсовой работы был проведен всесторонний анализ языков гипертекстовой разметки документов. Были решены следующие задачи:

- 1) проведен анализ понятия «гипертекст»;
- 2) проанализирована история развития гипертекста;



- 3) проведен анализ моделей гипертекста;
- 4) изучены виды языков гипертекстовой разметки документов – синтаксис, структуру документов, основные элементы.
- 5) провести анализ будущего Web-программирования.

Таким образом, цель данной курсовой работы «обзор языков гипертекстовой разметки» была достигнута

## Библиография

1. Бройдо В.Л. Вычислительные системы, сети и телекоммуникации СПб, Питер 2012- 464 с.
2. Информатика /под редакцией С.В.Симоновича. СПб, Питер 2011- 400 с.
3. Кирмайер М. Информационные технологии. СПб.: Питер, 2013 – 443 с.
4. Мэтьюз Дж. Web – сервер. СПб.: Символ, 2008 – 356 с.
5. Олифер В. Г., Олифер Н.А. Компьютерные сети. СПб.: Питер, 2007 – 864 с
6. Олифер В. Г., Олифер Н.А. Сетевые операционные системы. СПб.: Питер, 2009 – 539 с.
- 7) Сайт «HyperText», What is HyperText. CERN. Проверено 20 октября 2018. URL: <http://info.cern.ch/hypertext/WWW/WhatIs.html>
- 8) Тед Нельсон. Curriculum Vitae: Theodor Holm Nelson, PhD (англ.). Сайт Теда Нельсона. Проверено 20 октября 2018. URL: <http://hyperland.com/TNvita>
- 9) Юлия Шатилова. Какой была бы альтернативная Сеть? Грезы о цифровой вселенной знаний (рус.)(недоступная ссылка — история) (13 августа 2012 года, 16:02). Проверено 20 октября 2018. Архивировано из первоисточника 25 августа 2012.
- 10) Тед Нельсон. What's On My Mind (англ.). Сайт проекта Xanadu (Тед Нельсон — автор проекта). Проверено 20 октября 2018. URL: <http://www.xanadu.com.au/ted/zigzag/xybrap.html>
- 11) Ted Nelson. Literary Machines. — Edition 87.1. — 2007.
- 12) "Complex information processing: a file structure for the complex, the changing and the indeterminate" in Association for Computing Machinery: Proceedings of the 20th

National Conference. Ed. Lewis Winner: 84-100, Cleveland (Canada): ACM.

DOI:10.1145/800197.806036

13) Дуванов А.А., История гипертекста // Информатика – 1 сентября. - 2014. - №4. - С.23-24.

14) Костов Д.А., История гипертекста // Эврика – 2013. - №7. – С. 56-60

15) Квинт И.. HTML, XHTML и CSS. СПб.: Питер, 2011 – 382 с.

16) Брайн М. SGML and HTML Explained. Addison Wesley, 1997 – с. 584

17) Холзнер С. XML Энциклопедия. СПб.: Питер, 2010 – с. 1092

18) Сухов К. HTML5. Путеводитель по технологии.

19) Сайт консорциума W3C. URL: <https://lists.w3.org/Archives/Public/public-whatwg-archive/2018Mar/0071.html>

## Приложение А

Таблица 1 - Историческая последовательность

| <b>Год</b> | <b>Событие</b>   |
|------------|--|
| 1945 г.    | выход статьи Ванневары Буша "As We May Think".   |
| 1963 г.    | Энгельбарт публикует "A Conceptual Framework for the Augmentation of Man's Intellect".   |
| 1965 г.    | Нельсон вводит термин "гипертекст".  |
| 1967 г.    | в Брауновском Университете Анди ван Дам и Нельсон разрабатывают Hypertext Editing System (HES), за которой последовал выпуск FRESS в 1968 г. |

- 1968 г. Энгельбарт демонстрирует NLS на FJCC (Fall Joint Computer Conference), часть проекта Augment, начатого в 1962 г.
- 1972 г. начинается разработка ZOG в Университете Карнеги-Мэллона группой, возглавляемой Робертсоном.
- 1979 г. Нельсон приступает к проекту Xanadu.
- 1981 г. начинается разработка KMS в Knowledge Systems.
- 1981 г. Нельсон публикует "Literary Machines", где подробно описан проект Xanadu.
- 1982 г. система ZOG установлена на американском атомном авианосце Carl Vinson.
- 1982 г. начинается разработка Питером Брауном системы Unix Guide в Университете Кента.
- 1983 г. начинается разработка HyperTIES в Университете Мэриленда.
- 1983 г. Рэндол Тригг защищает первую диссертацию по гипертексту в Университете Мэриленда.
- 1984 г. начинается разработка Notecards в Xerox PARK.
- 1985 г. начало разработки Intermedia в Брауновском Университете.
- 1986 г. в Университете Северной Каролины приступают к разработке WE (Writing Environment).

- 1986 г. Office Workstation Inc. (OWL) выпускает Guide для Макинтоша.
- 1987 г. Apple Computers выпускает HyperCard (Бил Аткинсон) – первую гипермедиа авторскую систему, бесплатно устанавливаемую на каждом продаваемом "Макинтоше".
- 1987 г. Джефф Конклин публикует свой выдающийся обзор "Hypertext: An Introduction and Survey".
- 1987 г. ACM организует первую конференцию по гипертексту Hypertext'87 (Chapel Hill, North Carolina).
- 1987 г. выпуск системы Guide для MS Windows.
- 1989 г. Шнейдерман и Керсли разрабатывают Hypertext Hands-On! – первую электронную гипертекстовую книгу.
- 1989 г. Autodesk, производитель систем CAD, начинает поддержку Xanadu.
- 1989 г. Тим Бернерс-Ли выдвигает проект World Wide Web.
- 1989 г. коммерческая реализация IRIS Intermedia 3.0.
- 1989 г. опубликован "Afternoon, A Story" М. Джойса – первое произведение гипертекстовой беллетристики.
- 1989 г. вторая конференция ACM Hypertext'89 (Pittsburgh, Pennsylvania).
- 1990 г. первая Европейская конференция по гипертексту ECHT'90.

- 1990 г. декабрь основание Научно-технического центра гиперинформационных технологий (ГНТЦ "Гинтех") Министерства связи РФ.
- 1991 г. WWW в ЦЕРНе становится первым глобальным гипертекстом.
- 1991 г. разработана первая версия пакета ГиперМетод (для DOS) в Ленинградском электротехническом институте.
- 1991-1999 гг. проводятся международные гипертекстовые конференции.
- 1991-1996 гг. серия конференций Восток-Запад "Взаимодействие человека с компьютером" (EWHCI).
- 1992 г. Autodesk отказывается от работ по проекту Xanadu.
- 1992 г. создается группа новостей alt.hypertext.
- 1993 г. Международная конференция по гипермедиа и гипертекстовым стандартам в Амстердаме.
- 1993 г. NCSA (The National Center for Supercomputing Applications) выпускает Mosaic 1.0 (Marc Andreesson, Eric Bina).
- 1993 г. первая конференция, посвященная World Wide Web в Женеве.
- 1994 г. информационный поток в WWW (Web-трафик) впервые превышает другие трафики Интернет.