

Халабия Мария Леонидовна

магистрант

ФГБОУ ВО «Московский технологический университет»

г. Москва

СРЕДСТВА ДОСТИЖЕНИЯ ПОИСКОВОЙ ДОСТУПНОСТИ В РАСПРЕДЕЛЕННОЙ БАЗЕ ДАННЫХ

***Аннотация:** в данной статье даются определения поисковой доступности, распределенной вычислительной сети и базе данных. Автором представлена объектно-ориентированная модель библиографических метаданных, которая является средством достижения поисковой доступности.*

***Ключевые слова:** поисковая доступность, распределенная база данных, объектно-ориентированное моделирование, библиографические метаданных.*

DOI: 10.21661/r-112054

Начало XXI в. характеризуется распространением Интернет и Web-технологий для построения корпоративных информационных систем (КИС). Благодаря появлению беспроводной Интернет-технологии и высокой скорости передачи данных стал возможным стремительный переход от централизованных к распределенным системам хранения и передачи данных. В настоящее время не существует академического определения «распределенной системы». Э. Таненбаум в монографии, посвященной распределенным вычислительным системам (РВС) описал эту технологию так: «распределенная система – это набор независимых компьютеров, представляющий их пользователям единой объединенной системой» [1, с. 23]. Архитектура РВС позволяет легко соединять пользователей с ресурсами, скрывая тот факт, что они размещены далеко друг от друга.

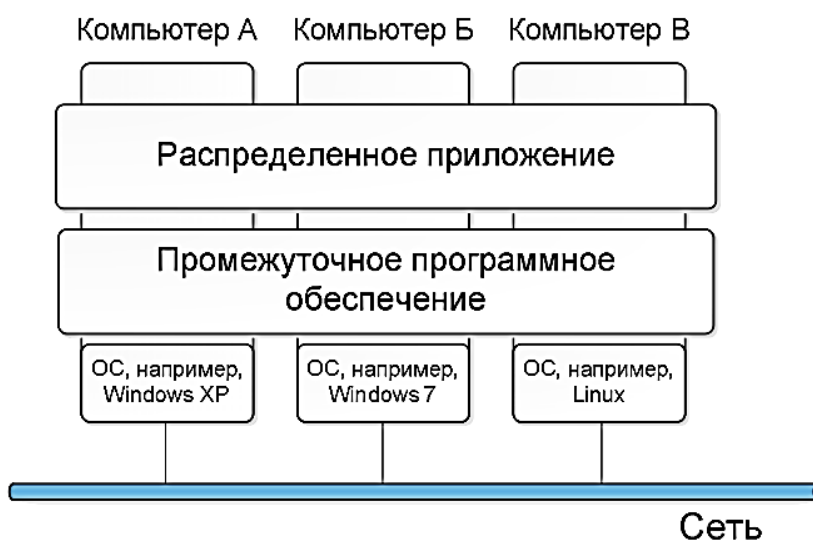


Рис. 1. Слои программного обеспечения в РВС

Исходя из классического понимания распределенной вычислительной системы, определим понятие распределенные базы данных (РБД) как набор множества логически связанных данных (баз и банков данных), представляющихся их пользователям единой объединенной системой для поиска необходимой информации.

В крупном холдинге информационный центр представляет собой сложную структуру, состоящую из множества подразделений. В этой связи на первый план выдвигается своевременная обработка данных и доступ к информации, необходимой для принятия бизнес-решений. Целью создания, обновления и хранения информации в распределенных базах данных является совершенствование поисковой доступности [2] за счет создания общего распределенного пространства.

Понятие «поисковой доступности» является многозначным, но для нашего исследования под этим термином подразумевается «степень легкости обнаружения конкретного объекта» [2, с. 20]. По нашему мнению, для информационных центров, удовлетворяющих информационные потребности корпорации, именно от скорости и качества передачи данных зависит своевременная реакция на события, происходящие в мире.

Библиографические метаданные являются описательными и извлекаются из титульного листа документа. РБД содержит описание ресурсов, которые могут

быть логически связаны отношениями путем ссылок с разными объектами, при этом в представлении пользователя они находятся в единой базе данных. В этой связи актуальным является объектно-ориентированное моделирование элементов данных, представляющихся составной частью распределенной базы данных. Оно характеризуется индивидуальностью, таксономией, полиморфизмом и наследованием. Следует отметить, что перечисленные характеристики объектно-ориентированной технологии могут рассматриваться обособленно, как части большого и сложного конструктора, которые вместе представляют единый и исчерпывающий взгляд на конкретную предметную область, содержащуюся в РБД.

Основное назначение моделирования, состоит в том, чтобы рассмотреть библиографические метаданные с точки зрения прикладной лингвистики как сложную систему со своим метаязыком. Такой подход позволяет сократить сложность представления библиографических данных, выделяя ограниченный набор важнейших свойств документа. Следовательно, под моделью понимается абстракция, которая создается с целью создания поиска информационного ресурса.

В нашем исследовании мы акцентируем внимание на модели классов, поскольку именно эта модель является самой важной и позволяет сформировать поисковый образ документа. Кроме того, она улучшает поисковую доступность объектов и ресурсов путем создания связей между разными свойствами описываемого объекта. Модель классов описывает структуру ресурсов, содержащихся в библиографических метаданных, их индивидуальность, отношения с другими объектами, атрибуты и операции.

В РБД под объектами имеют в виду «записанную информацию, имеющую определенные реквизиты и соответствующую требованиям жанра и вида документа, зафиксированную на (в) вещевом изделии» [4]. Это определение, на наш взгляд, очень точно характеризует поисковую доступность любого объекта, так как для поиска нужно задать координаты, чтобы найти ресурс.

Внутренняя индивидуальность ресурса определяется формой произведения, а также видом и типом описываемого документа. Благодаря этой особенности

объекты можно отличить друг от друга даже при наличии одинаковых реквизитов (совпадающие ФИО автора и названия произведений). Таким образом объектом моделирования в объектно-ориентированной модели является информация о документах, представленная в форме библиографических метаданных, которые в сжатой форме представляют основной предмет содержания документа.

Отдельный объект, представленный в модели, является экземпляром класса. Класс описывает группу объектов с одинаковыми атрибутами, одинаковыми операциями, типами отношений и семантикой. В случае с библиографическими метаданными за основу класса принята форма произведения. В зависимости от объёма и сложности представления библиографических данных «форма произведения» может являться суперклассом, объединяющим произведения или ресурсы путем наследования определенных характеристик. В случае простого описания «форма произведения» представляет собой класс, к которому относится описываемый или разыскиваемый объект.

Класс характеризуется общими атрибутами, операциями и связями и ассоциациями. Подклассы, в свою очередь, добавляют к ним собственные атрибуты, операции и связи. Следовательно, под классификацией здесь понимается группировка объектов библиографического описания с одинаковыми структурами данных (атрибутами) и операциями в классы. Поэтому процессы обобщения информации, представленной в произведении, играют центральное место [3].

Под обобщением понимают отношения между классом (суперклассом) и одной или несколькими его вариациями (подклассами). Часто эту связь можно представить словом «является». Например, «форма произведения является романом, пьесой, кинофильмом и т. п.». Каждый экземпляр подкласса (роман, пьеса и т. п.) одновременно является экземпляром в суперклассе. Замечено, что простые обобщения упорядочивают классы в рамках определенной схемы. Иерархия со сложной структурой является библиографическим родом, который может быть представлен, помимо традиционных изданий кинофильмом, музыкальными и другими произведениями. В этом случае каждый подкласс может иметь

одного непосредственного предка (его суперкласс). Термины «родитель» и «потомок» используются для описания классов, находящихся далеко друг от друга по уровням, но связанные отношениями обобщения (например, роман (класс) и снятый по его мотивам кинофильм). Такая структура иерархии данных называется наследованием. Следует отметить, что под этим термином понимают «полное или частичное копирование внутренней структуры и набора операций заданного класса (суперкласса) для подчиненного класса (подкласса)» [3].

Атрибуты характеризуют особенности суперкласса (класса) и его подклассов. Они являются «наименованным свойством класса, описывающим значение, которое может иметь каждый объект класса» [3, с. 45]. Атрибуты считаются прилагательными, которые раскрывают особенности класса. В нашей модели библиографических метаданных определены 24 атрибута, которые раскрывают суть классов, представленные в таблице 1.

Таблица 1

Атрибуты модели элементов библиографических метаданных

<i>Суперкласс / класс</i>	<i>Атрибут</i>	<i>Значение атрибута</i>
Форма произведения	Форма произведения (formofwork)	Называет форму произведения и определяет общую структуру представления записи. Является родителем для всех последующих произведений.
Лицо	Имя лица (nameofperson)	Имя лица
	Даты (dateofperson)	Даты, связанные с именем лица
	Титул (titleofperson)	Титул лица
	Коллективный автор (corporatename)	Коллективный автор
Конкретная форма произведения (роман, пьеса, кинофильм) как подкласс или самостоятельный класс	Название рукописи (nameofmanuscript)	Характеризует рабочее название произведения
	Форма произведения (formofwork)	Обозначает форму произведения
	Автор (author)	Автор произведения, обозначенный на титульном листе
	Дата создания произведения (dateofcreation)	Дата создания произведения

	Язык (language)	Язык произведения
	Версия (version)	Версия произведения
	Заглавие произведения (title of work)	Окончательное заглавие произведения
Издание	Издатель (publishing house)	Издатель, издательство
	Место издания (place of edition)	Место издания
	Страна публикации (country of publication)	Обозначает страну, в которой произведение впервые опубликовано
	Год издания (year)	Год издания
	Том (value)	Том, если издание многотомное или многочастное
	Серия (series)	Заглавие серии
	Описание (annotation)	Аннотация, краткое описание издания
	Носитель информации (holder of information)	Носитель информации, на котором представлено издание
	Формат (format)	Формат, в котором доступен электронный ресурс
	Объем (pagination)	Пагинация, количество страниц
Классификационный индекс	Классификационный индекс (classification index)	Классификационный индекс
	Система классификации (system classification)	Система классификации, принятая в библиографирующей организации
	Предметная рубрика (subject)	Предметная рубрика, к которой относится классификационный индекс
	Ключевые слова (keywords)	Неконтролируемые или контролируемые ключевые слова
Экземпляр	Организация (organization)	Организация – держатель экземпляра
	Отдел (department)	Отдел, в котором храниться экземпляр издания
	Доступ (access)	Условия доступа
	Шифр хранения (pressmark)	Шифр хранения

Приведенные в таблице 1 атрибуты описывают значения, которые принадлежат отдельным физическим единицам. Тем самым они не могут обладать индивидуальностью, присущей отдельным экземплярам, которые являются одновременно и объектами описания, и объектами поисковой доступности [5].

В описанной модели связь (link) – концептуальное и единичное соотношение между объектами. Ассоциации, возникающие между объектами, представляют собой группу отношений, обладающих общей структурой и семантикой. Замечено, что связи и ассоциации присутствуют в моделировании в качестве глаголов. Они могут быть различными и зависеть в некоторой степени от владения пользователем информационной грамотностью.

Таким образом создание поискового образа документа в распределенной системе базируется на объектно-ориентированной модели библиографических метаданных [5]. Эта модель позволяет создавать поисковой образ документа, соответствующий запросу пользователя, что в свою очередь способствует совершенствованию поисковой доступности, что в конечном итоге экономит время на принятие бизнес-решений сотрудниками крупных корпораций.

Список литературы

1. Таненбаум Э. Распределенные системы [Текст]: принципы и парадигмы / Э. Таненбаум, М.В. Стеен. – СПб.: Питер, 2003. – 877 с.
2. Морвиль П. Тотальная видимость / Питер Морвиль; пер. С. Иноземцева. – СПб.: Символ-Плюс, 2008. – 272 с.
3. Рамбо Дж. UML 2.0 [Текст]: объектно-ориентированное моделирование / Дж. Рамбо, М. Блаха. – 2-е изд. – СПб.: Питер, 2007.
4. Швецова-Водка Г.Н. Общая теория документа и книги [Текст]: Учеб. пособие / Г.Н. Швецова. – М.: Рыбари; Киев: Знання, 2009. – 487 с.
5. Халабия Р.Ф. Объектно-ориентированный подход к моделированию элементов библиографического описания в динамических распределенных базах данных / Р.Ф. Халабия, М.Л. Халабия // Информатизация образования и науки. – 2012. – №2. – С. 140–149.