**УПРАВЛЕНИЕ ДАННЫМИ**

1) НЕОБХОДИМОСТЬ УПРАВЛЕНИЯ ДАННЫМИ

2) ДИЗАЙН БАЗЫ ДАННЫХ

3) УПРАВЛЕНИЕ ДАННЫМИ И ЭКСПЛУАТАЦИЯ

4) ДОСТУП К ДАННЫМ И РАСПРОСТРАНЕНИЕ

Данные должны храниться надежно, но должны быть легко доступны для анализа. База данных должна хранить исходные необработанные данные. Система управления данными должна быть максимально интегрирована с системой сбора данных. Интерфейс «человек-компьютер» должен помочь пользователю получить максимальную отдачу от системы. Ввод данных должен включать функции импорта и средства проверки, обработка должна использовать встроенные функции для общих процедур, а отчетность должна быть гибкой и включать средство экспорта. Доступ должен контролироваться, чтобы обеспечить целостность и конфиденциальность базы данных.

**1) НЕОБХОДИМОСТЬ УПРАВЛЕНИЯ ДАННЫМИ**

Принятие решений при разработке, планировании и управлении, в значительной степени зависит от обработанной информации. Данные должны быть интерпретированы, прежде чем они могут быть использованы. Объем необработанных первичных данных часто очень велик, и поэтому его можно эффективно использовать только в том случае, если он хранится в системе управления базами данных (СУБД).

Функции СУБД:

· Обеспечить соответствие данных стандартным классификациям;

· Обеспечить достоверность данных;

· Обеспечить целостность данных и внутреннюю согласованность;

· Обеспечить и поддерживать первичные данные;

· Обеспечить легкий доступ к первичным данным;

· Эффективно обрабатывать данные по мере необходимости;

· Разрешить интеграцию различных наборов данных, тем самым увеличивая их общую полезность.

Основополагающим принципом является хранение всех данных в том виде, в котором они были собраны, в их первичной форме. Это обеспечивает гибкость в способах обработки данных (например, отфильтрованных, агрегированных, преобразованных) и гарантирует, что все вычисления воспроизводятся из исходных данных, включающих все ревизии. Учитывая значительные инвестиции в сбор данных и низкие затраты на хранение и обработку, нет особых причин не хранить полные данные в их первичной форме.

**2) ДИЗАЙН БАЗЫ ДАННЫХ**

2.1) Методология

2.2) Человеко-машинный интерфейс

2.3) Компьютеризированная документация

2.4) Ввод данных

2.5) Обработка данных

2.6) Представление данных

2.1) Методология

Следует рассмотреть децентрализованную структуру базы данных, чтобы упростить управление базой данных и проверку данных. В распределенной системе данные вводятся и проверяются локально, но связаны с другими базами данных для анализа.

При рассмотрении подхода к созданию новой системы сбора данных предлагаются различные варианты. Они включают:

* Взять коммерчески доступное программное обеспечение и адаптировать его к новым требованиям;
* Соединение системы с различными программными компонентами;
* Создание собственной системы с нуля.

Преимущества и недостатки различаются для каждого подхода и должны тщательно взвешиваться перед выделением ресурсов.

Важным преимуществом пользовательской разработки является то, что ее можно настроить так, чтобы она точно соответствовала методологии выборки данных, поэтому система будет более эффективной и легко принятой. Другим возможным преимуществом является то, что проект базы данных также может быть использован в качестве инструмента для содействия разработке программы сбора данных. Если две фазы разработки происходят одновременно, использование общей терминологии (то есть идентификации видов, методов отбора проб) и инструментов (то есть диаграмм потоков данных, анализа задач) может быть взаимовыгодным для обеих систем.

2.2) Человеко-машинный интерфейс

Важным для общей приемлемости СУБД является интерфейс «человек-компьютер» (HCI). Ниже приведены некоторые основные принципы, которые могут быть использованы для разработки эффективных интерфейсов HCI:

* Автоматизированные процедуры для руководства пользователей о том, как действовать при использовании системы;
* Использование графических структур, таких как командные кнопки в HCI, предпочтительно с обычно применяемыми значками, для облегчения доступа к часто используемым функциям;

2.3) Компьютеризированная документация

Онлайновая помощь, документация, учебные пособия и обучение являются факторами, способствующими устойчивости базы данных. Особое внимание следует уделить разработке этих компонентов в системе. Предпочтительно, разработка этих компонентов должна идти параллельно с разработкой ПО / пользовательских интерфейсов. Однако это не устраняет необходимость в печатных копиях документации.

2.4) Ввод данных

Должна быть доступна функция «Импорт» для включения данных, обычно хранящихся в альтернативных форматах (например, текстовый процессор или электронная таблица). Эта функция должна обеспечивать целостность и качество данных.

Когда это применимо, следует разработать специальные структуры или программные ссылки для облегчения поиска данных из других компьютерных источников, таких как электронные журналы.

Проверка данных может осуществляться на различных уровнях, включая сбор, компиляцию, ввод данных в СУБД, обработку и анализ данных. Пользовательские интерфейсы ввода данных должны быть структурированы, чтобы обеспечить соблюдение наборов правил, применяемых для проверки входных данных.

2.5) Обработка данных

Особенностью технологии СУБД, которую следует использовать при разработке или изменении системы сбора данных, является возможность встраивания управления и обработки в базу данных с использованием хранимых процедур и запросов. Этот подход имеет следующие преимущества:

* Уменьшение количества необходимой внешней обработки;
* Обеспечение более немедленной проверки данных;
* Повышение гибкости для будущих модификаций системы.

Важным фактором при обработке данных является необходимость ведения контрольного журнала всех действий, выполняемых с данными, чтобы обеспечить последующий анализ качества информации.

Следует использовать параметры, чтобы сделать систему более гибкой. Параметры - это легко изменяемые значения, которые изменяют структуру и функции системы. Зачастую требования меняются в течение срока службы системы, и возможность расширения и модификации без существенных изменений конфигурации может сохранить жизнеспособность системы сбора данных.

2.6) Представление данных

Гибкость при создании отчетов из данных очень важна. Часто потенциальное использование данных не полностью распознается до того, как система заработает.

Для обеспечения гибкости отчета должна быть предусмотрена универсальная функция «экспорта». Функции, которые должна иметь эта функция:

* Определение тегов имен для всех экспортируемых атрибутов данных;
* Сводка типов данных и форматов;
* Записи переменной длины с выбранными пользователем разделителями полей (т.е. файлы ASCII с запятыми или табуляцией).

**3) УПРАВЛЕНИЕ ДАННЫМИ И ЭКСПЛУАТАЦИЯ**

3.1) Архивы

3.2) Переоценка проекта

3.1) Архивы

База данных должна регулярно создавать резервные копии. Система всегда должна быть готова к серьезным аппаратным или программным сбоям и потере данных. Процедуры должны быть сделаны как можно более простыми, чтобы обеспечить регулярное резервное копирование.

Поскольку база данных развивается со временем и происходят изменения в информационных технологиях, архивирование данных необходимо для извлечения исторических данных, сохраненных в прежней структуре или дизайне.

3.2) Переоценка проекта

В результате установленных механизмов обратной связи и для обеспечения того, чтобы система управления данными соответствовала своим целям, следует проводить периодические оценки.

Рекомендуется постоянная программа оценки дизайна, чтобы гарантировать, что система использует преимущества последних разработок в области информационных технологий. Особое внимание следует уделить установлению процедур обновления архивных данных, чтобы данные в старом формате по-прежнему были доступны.

**4) ДОСТУП К ДАННЫМ И РАСПРОСТРАНЕНИЕ**

4.1) Владение данными и контроль

4.2) Сети связи

4.3) Компьютеризированная публикация

4.1) Владение данными и контроль

Государство или учреждение, из которого были получены данные, является основным владельцем данных. Признавая, что данные являются ресурсом и, следовательно, имеют ценности, экономические или иные, правительство должно осуществлять свое право на поддержание, безопасность и контроль доступа.

Контроль - это ограничение, налагаемое на способность отдельного лица, группы лиц, организаций или другого государства иметь частичный или полный доступ к данным, содержащимся в базе данных. Частичный доступ к данным - это невозможность выполнения любого из следующих действий: просматривать все данные, введенные и сохраненные системой, добавлять данные, редактировать данные, копировать данные или распространять / делиться данными любым способом. Элементы управления должны использоваться для ограничения доступа в соответствии с любыми требованиями конфиденциальности и защиты данных от несанкционированных изменений. Наибольшее значение имеет защита первичных данных от случайного повреждения. Основная копия данных всегда должна быть защищена от записи. Однако, хотя контроль и безопасность важны, они не должны препятствовать законному доступу. В частности, функции безопасности и контроля СУБД никогда не должны мешать признанным государством научным учреждениям получать доступ к данным для исследования управления ресурсами.

4.2) Сети связи

Разработки в области коммуникационных технологий открывают новые возможности для распространения данных. Когда это возможно и уместно, при разработке СУБД следует учитывать структуры, которые будут способствовать распространению или позволят осуществлять прямой доступ к данным из удаленных мест.

4.3) Компьютеризированная публикация

Разработка ПО для учебных пособий, демонстраций и связанных с ними документов имеет важное значение для долгосрочной жизнеспособности базы данных. Эти документы могут храниться локально или, предпочтительно, на национальном уровне в форме, обеспечивающей доступ к сети.

Использование цифровых медиа также следует учитывать для распространения статистики. Например, Интернет предлагает недорогой способ обмена информацией, обеспечивающий безопасный доступ к данным и аналитическим результатам.

Андрианов Н. Я.

НТУУ «КПИ им. Игоря Сикорского»