

Distr.  
GENERAL

CES/SEM.47/31  
18 January 2002

RUSSIAN ONLY

**КОМИССИЯ ПО СТАТИСТИКЕ и  
ЭКОНОМИЧЕСКАЯ КОМИССИЯ ДЛЯ ЕВРОПЫ**

**КОМИССИЯ  
ЕВРОПЕЙСКИХ СООБЩЕСТВ**

**КОНФЕРЕНЦИЯ СТАТИСТИКОВ ЕВРОПЫ**

**ЕВРОСТАТ**

**Совместный рабочий семинар ЭКЕ и ЕВРОСТАТа  
по интегрированным статистическим  
информационным системам и смежным вопросам  
(ИСИС 2002) (Женева, 17 - 19 Апрель 2002)**

Тема I: применение Web-технологии для интегрирования статистических данных

### **Использование WEB-технологий для интеграции статистических данных**

Дополнительный доклад

Представлен Агентством Республики Казахстан по статистике <sup>1</sup>

1. Бурное развитие Internet и Web-технологий значительно повлияли на способы и методы сбора и обработки статистической информации. Технология WWW за последнее время приобрела столько функциональных возможностей, что сейчас возникла необходимость пересмотреть многие методы обработки статистической информации на всех этапах статистического производственного процесса. Так, например, новый класс приложений - так называемые Web-приложения стали одним из наиболее эффективных инструментов, обеспечивающих доступ к данным через Интернет или интрансет к информационным системам и базам данных.

2. При построении корпоративной информационной системы статистики РК было принято решение максимально использовать стратегии Internet и Intranet. С целью повышения эффективности функционирования системы статистического производства осуществлена модернизация корпоративной сети статистики. Корпоративная сеть статистики использует публичные каналы связи, входящие в Internet, в сети обеспечивается защита передаваемых данных и приняты меры по пересечению проникновения извне на корпоративные узлы. Корпоративная сеть органов статистики Республики Казахстан включает в себя сетевые узлы нижнего уровня (удаленные офисы) и верхнего уровня (в центральном офисе). В ее состав входят:

- Локальная вычислительная сеть республиканского уровня;
- Локальные компьютерные сети областных управлений по статистике;
- Компьютеры отдельных абонентов районного уровня .

---

<sup>1</sup> Автор: Алихан Асханович Смаилов

3. На республиканском уровне введена в строй локальная вычислительная сеть на 450 клиентских мест. В состав сети включены два сервера - IBM RS/6000 на платформе RISC, функционирующий под управлением сетевой операционной системы AIX 4.3.3, и сервер- кластер IBM NetFinity 7000, включающий в себя два четырехпроцессорных сервера PentiumPRO 200, а также множество серверов, на которых развернут пакет MS BackOffice 2000 (MS Exchange Server, MS Windows 2000 Server, MS Proxy Server, MS Internet Server, MS SQL Server). Проведена работа по ускорению доступа к Web-сайту, который подключен к узлу провайдера Интернета скоростным каналом связи. Обеспечен быстрый и безопасный доступ в Internet через аппаратный и программный Firewall. Приняты меры к обеспечению максимальной безопасности Intranet от несанкционированного доступа как со стороны Internet, так и со стороны Intranet. Эти меры включают в себя как технические (организация сети, разграничение прав доступа и т.д.), так и административные.

4. При разработке информационной системы осуществляется интеграция технологий баз данных, текстовых (документальных) систем, технологий Java, Web-технологий на основе различных подходов. Это ведет не только к обогащению функциональных возможностей создаваемых крупных подсистем, но и к обеспечению интеграции неоднородных информационных ресурсов, созданных и поддерживаемых средствами различных технологий. Внедряются технологии обеспечивающие сопряжение различных типов информационных серверов , например, SQL серверов и Web-серверов как в составе одного информационного узла, так и находящихся в различных информационных узлах компьютерной сети системы Статагентства. Используется технология от Microsoft ASP (Active Server Pages) , позволяющая легко разрабатывать приложения для WWW. ASP работает на платформе Windows 2000 и IIS (Internet Information Server). ASP – технология, позволяет подключать программы к Web-страницам. Главные преимущества ASP – простой скриптовый язык (Visual Basic Script или Java Script) и возможность использования внешних COM (Component Object Model). При разработке некоторых подсистем обработки статотчетности реализуется типичное Intranet - решение, реализованное на базе программных продуктов Microsoft, клиентские части на базе Internet Explorer , сервер приложений - на основе Internet Information Server, и, наконец, сервер базы данных на основе Microsoft SQL Server 2000. Благодаря связям с корпоративными базами данных, файл-серверами и хранилищами документов Web-серверы предоставляют сотрудникам статистики самые различные виды информации через единый интерфейс - Web-браузер. Организована корпоративная работа с некоторыми базами данных , всем специалистам, работающим в системе статистики, предоставлен быстрый и эффективный доступ к статистическим данным, хранящимся в базах данных. Сотрудники имеют доступ к общим информационным ресурсам Агентства РК по статистике благодаря внедрению и широкому использованию Web и офисных приложений, систем электронной почты и систем коллективной работы.

5. С точки зрения разработки прикладного программного обеспечения Web оказал огромное влияние на разработку средств доступа к информационным ресурсам. Основным здесь является отказ от программирования интерфейсов пользователя, что составляло львиную долю времени работы программистов. Теперь таким универсальным интерфейсом пользователя стал браузер. Требования к квалификации разработчиков резко снизились. Фактически, нужно уметь только работать со стандартными потоками ввода/вывода. Используя Web-технологии , как универсальный способ построения сетевых информационных систем, независимых от платформ при создании приложений по обработке статистических данных, мы получили ряд преимуществ-упростились разработка информационных приложений и методы администрирования приложений и рабочих мест, приложения могут работать на разнообразных платформах, мы можем добиться унификации диалогового интерфейса и рабочих мест и осуществлять интеграцию офисных приложений и приложений, работающих с базами данных. Эффективным механизмом решения проблемы обеспечения “взаимопонимания” различных документационных систем является стандартизация обменных форматов с использованием расширяемого языка разметки XML(eXtensible Markup Language), де-факто, ставшего стандартом для передачи структурированных данных между разнородными системами, и автоматизации электронного взаимодействия различных систем делопроизводства.

### **Использование intranet-сети с внутренними Web-сайтом**

6. В локальной сетевой среде все преимущества Internet реализованы с использованием стандартных средств Internet: корпоративным WWW-сервером с комплектом необходимого программного обеспечения и Web-браузеров. Информация внутри организации распределена по множеству компьютеров и хранится в виде разнообразных файлов, отчетов и сообщений электронной почты. Поэтому чтобы облегчить своим сотрудникам доступ к разнообразной информации в системе статистики создана intranet-сеть с внутренними Web-сервером.

7. В системе статистики ПК реализованы Intranet приложения, которые позволяют разместить информацию в локальной вычислительной сети и обеспечить доступ к ней сверху вниз. Различные отделы или департаменты размещают свою информацию на корпоративном сервере, обеспечивая, таким образом, максимально быстрый доступ к ней.

8. Ведется работа по созданию транзакционных приложений, обеспечивающих “двухстороннюю обратную связь” с пользователями системы. В случаях, когда сотрудникам необходимы разработанные отчеты, получение обновленных версий программных продуктов, отсылка и получение локальной почты, использование WEB-технологий позволяет значительно сократить время и уменьшить “фактор восстановления”. Разрабатываются комплексные приложения с групповым способом общения. Они включают группы новостей с возможностью прямого обмена информацией между различными членами группы и предоставление доступа к определенной информации для пользователей вне группы. Эти типы WEB приложений значительно увеличили продуктивность совместной работы отдельных отраслевых отделов, а также значительно улучшили связь между департаментами и филиалами, позволили изменить доступ к информации, получаемой из различных источников, в различных форматах.

9. Внедрение Web-технологий позволило эффективно выполнять работы, связанные с коллективной подготовкой и обработкой информации, управлением электронными документами и документооборотом. Web-интерфейс стал привычным для многих приложений автоматизации статистической деятельности. Почтовые, новостные и другие сервисы Internet встраиваются в приложения для коллективной работы. Осуществляется сближения поисковых технологий Internet и офисных систем управления электронными документами, осуществляется интеграция Internet с корпоративными офисными приложениями. Благодаря функционированию сети Intranet, отпала необходимость в тиражировании многочисленных бумажных документов т.к. все нужные для работы документы сосредоточены в одном месте, на Web-сервере, и специалисты различных управлений и отделов могут легко найти их, прочитать, изучить и внести свои корректировки в случае необходимости.

10. В частности, использование интранет-технологий позволило создать и предоставить всем пользователям локальной сети внутренний сайт со статистической информацией, выпускаемой Агентством (ежегодники, сборники, справочники, журналы, бюллетени и экспресс-информация, аналитические данные). Сайт позволяет просмотреть и использовать информацию из подсистемы "Динамика статистических показателей", и еще массу разнообразной справочной информации. В дальнейшем по мере необходимости количество Web-сайтов будет увеличиваться, т.к. будут создаваться дополнительные специализированные сайты по определенной теме в каждом управлении Агентства РК по статистике.

### **Использование Web-технологий для разработки приложений подсистемы «Динамика статистических показателей»**

11. В 2002 году в Информационном Вычислительном Центре (ИВЦ) Агентства Республики Казахстан(РК) по статистике завершается внедрение в промышленную эксплуатацию подсистемы корпоративного хранилища данных «Динамика статистических показателей», разработанной с использованием системы управления базами данных Oracle8. Данная подсистема позволяет ежемесячное накопление и постоянное, надежное хранение статистических показателей по всем

отраслям статистики. В основе базы данных - преобразованные статистические данные, которые формируются из первичных данных, импортированных из локальных баз комплексов первичной обработки данных статотчетности, полученных в результате проведения обследований. Формирование данных происходит на основе существующих статистического регистра и классификаторов. Основная масса информации не просто копируется в БД, а подвергается интеллектуальной обработке - данные преобразовываются, очищаются, к ним добавляются новые атрибуты. Попав в БД, данные проходят второй этап обработки, в ходе которого, с точки зрения конечного пользователя, повышается их ценность: они по мере необходимости консолидируются, агрегируются, разбиваются на группы и пр., становясь сопоставимыми и легкодоступными. Статистические данные - это главный стратегический актив органов статистики. Но известно, что истинной ценностью обладают не только сами данные, а та информация, которую нужно уметь извлечь из них и правильно ее проинтерпретировать и проанализировать. Для того, чтобы это сделать достаточно быстро и эффективно, необходимо правильно организовать информацию в виде базы данных, где данные, сгруппированы по информационным объектам и отношениям с использованием определенных систем кодирования и стандартов, и легко доступны пользователю. Такие функции успешно выполняет новая версия подсистемы корпоративного хранилища данных «Динамика статистических показателей».

12. Для быстрого и несложного доступа к данным, просмотра и манипулирования данными в различных разрезах, создан привлекательный, понятный и удобный интерфейс пользователя, с помощью которого динамику показателей можно представить в виде отчетов и графиков в разрезе областей, форм собственности, видов экономической деятельности, секторов экономики и т.д. Есть возможность формирования запроса к БД наиболее простыми для неподготовленного пользователя средствами, провести быстрый анализ данных, выполнить сопоставление данных по разным территориям РК и за разные годы. Например, всегда можно узнать, в каком объеме производится тот или иной продукт в разных областях в разные периоды времени, а также выяснить, где и в каких количествах он потребляется. Обеспечена аутентификация пользователя для разграничения доступа.

13. Применение Web-технологий послужило основой для создания интерфейса с хранилищем данных «Динамика статистических показателей». Удаленные пользователи будут работать с системой, используя единый интерфейс - хорошо знакомый Web-браузер. Затрат на обучение работе с системой практически не требуется, т.к. большинство пользователей уже умеют работать с WEB-браузерами. При этом предприняты меры по защите передаваемой по сети информации. В настоящее время специалисты ИВЦ продолжают работать над проблемами - обеспечить отображение интерфейса пользователя в формате HTML для отображения программой просмотра клиента и обеспечить обработку запроса и возврат результата в HTML формате для отображения программой просмотра пользователя.

### **Проблемы совершенствования корпоративного Web-узла**

14. Всемирная паутина используется повсюду, и благодаря этому она оказывается идеальным средством представления информации. Публикация данных на Web-странице - эффективный, простой и недорогой способ передачи информации всем, кто в ней заинтересован.

15. Специалисты ИВЦ поставили задачу максимально использовать основные компоненты Web-технологии при работе над совершенствованием современного сложного корпоративного Web-узла, который является основой Intranet-публикации статистических данных. Непременными атрибутами Web-узла являются наличие виртуальных страниц, доступ к корпоративной базе данных, наличие поисковой машины, средств разграничения доступа к ресурсам и т.п. Узел должен быть и распределенным. При этом учитывается, что в системе статистики для работы с информационными ресурсами узла применяется неоднородное программное и аппаратное обеспечение. Ведется работа по решению вопросов настройки и оптимизации программного обеспечения Web-узла, сбора статистики посещений отдельных страниц, использование на своем

Web-узле поисковых машин. Расширение возможностей Web-узла с использованием универсальных СУБД.

16. Важно рассматривать большой Web-сайт не просто как базу данных, а как информационную систему, построенную на основе одной или более баз данных со сложной навигационной структурой. При такой точке зрения Web-сайт имеет много сходных характеристик с информационными системами других типов. Использование этих принципов при построении Web-сайтов будет также оказывать влияние на способы осуществления запросов к Web-сайтам и интеграции данных из множественных Web-источников.

17. Итак, наиболее нужными, важными и привлекательными свойствами Web-технологии являются возможность интеграции различных типов данных, использование механизмов связывания информации, распределенной по территориально разнесенным узлам компьютерной сети и наличие простого в установке и использовании браузера, претендующего на роль универсального клиента.