

Distr.
GENERAL

ECE/CES/SEM.54/11 (Summary)
11 April 2006

RUSSIAN
Original: ENGLISH and RUSSIAN ONLY

**СТАТИСТИЧЕСКАЯ КОМИССИЯ и
ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ
ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ**

ЕВРОПЕЙСКАЯ КОМИССИЯ

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ

**СТАТИСТИЧЕСКОЕ УПРАВЛЕНИЕ
ЕВРОПЕЙСКИХ СООБЩЕСТВ (ЕВРОСТАТ)**

**ОРГАНИЗАЦИЯ ЭКОНОМИЧЕСКОГО
СОТРУДНИЧЕСТВА И РАЗВИТИЯ (ОЭСР)
СТАТИСТИЧЕСКИЙ ДИРЕКТОРАТ**

**Совместный семинар ЕЭК ООН/Евростата/ОЭСР по вопросам управления
статистическими информационными системами (УСИС)**

София, Болгария, 21-23 июня 2006 года

Тема i): Изменения в статистических процессах

ИНТЕГРИРОВАННАЯ ТЕХНОЛОГИЯ ОБСЛЕДОВАНИЙ (ИТО) - НОВЫЙ ПОДХОД К ОБРАБОТКЕ СТАТИСТИЧЕСКИХ ДАННЫХ

**Вспомогательный документ, подготовленный Бранко Джиречеком и Мирой Никич,
Статистическое управление Сербии и Черногории**

Резюме

1. В течение длительного времени обработка данных в статистических системах была и до сих пор остается организованной по принципу "потока", результатом чего является создание набора различных и часто несовместимых между собой систем для решения схожих задач. Более общий подход был разработан и осуществлен в Статистическом управлении Сербии и Черногории; при этом подходе обработка всех статистических данных осуществляется совершенно иным образом.

I. КРАТКАЯ ИСТОРИЯ ВОПРОСА

2. Автоматическая обработка данных используется в нашей стране уже в течение длительного времени. В бывшей Югославии первым закупленным компьютером был компьютер, закупленный и установленный для обслуживания потребностей Союзного статистического управления (так оно раньше называлось). На протяжении длительного времени Союзное управление было лидером в области разработки и применения ИТ, поскольку инновационные технологии обработки данных всегда внедрялись в нем.

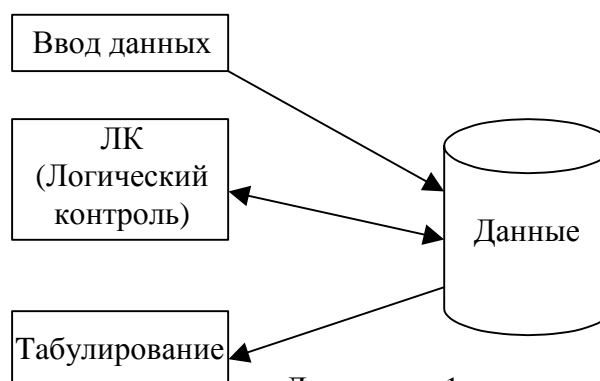
3. В последнее время произошло резкое свертывание деятельности как всего Управления, так и его департамента ИТ. В результате последнего "сокращения" в администрации союзного государства в департаменте ИТ осталось всего четыре сотрудника. В этой ситуации от этого подразделения вряд ли можно ожидать высоких результатов, но тем не менее определенные достижения у этого департамента имеются.

II. НЫНЕШНЯЯ СИТУАЦИЯ

4. Самым традиционным способом обработки данных статистических обследований является использование так называемого принципа "потока", который предусматривает осуществление всего процесса, начиная с ввода данных и кончая публикацией окончательных результатов, в отношении каждого обследования. Обработка данных каждого обследования состоит в основном из трех следующих этапов:

- i) Ввод данных
- ii) Логический контроль ("выверка" данных)
- iii) Табулирование (распространение, публикация)

5. С развитием ИТ у различных обследований появились разнообразные, "модные" на соответствующий момент времени версии. В зависимости от целого ряда факторов, таких, как важность некоторых обследований, желание или нежелание каких-либо изменений и т.д., возникали ситуации, когда различные обследования по-разному обрабатывались методами ИТ - в одних



случаях данные обследований хранились в "текстовых" последовательных файлах, а в других - (к сожалению) в различных других базах данных.

6. В течение длительного времени системы ИТ реализовывались исключительно на платформе больших вычислительных машин IBM. Естественно, что по мере развития персональных компьютеров некоторые обследования были переведены на эту новую платформу, и, естественно, мы имеем в наличии также смешанные варианты обработки данных, что еще больше осложняет ситуацию.

7. В итоге обработка статистических данных в нашей стране производится на различных платформах, реализованных с использованием различных программных средств; а данные хранятся различными способами, обособленно по каждому обследованию, и, что еще хуже, обособленно по каждому этапу, как показано на диаграмме 1. Это является весьма упрощенной картиной существующего положения, и это сделано умышленно, поскольку, по мнению авторов, именно в этом заключается самая большая проблема, присущая всем обследованиям.

IV. БАЗОВЫЕ КОНЦЕПЦИИ ИНТЕГРИРОВАННОЙ ТЕХНОЛОГИИ ОБСЛЕДОВАНИЙ (ИТО)

8. Следует отметить, что, хотя мы обращались к экспертам из ряда различных статистических управлений с просьбой помочь в решении вышеупомянутых проблем, похоже, никто на эту просьбу не откликнулся. Поэтому нам пришлось действовать "с чистого листа" и по собственному разумению. Единственным исключением была программа Euro Trace Евростата, но, поскольку эта программа не подходила для большинства наших обследований, то мы узнали об этой программе слишком поздно, и к тому времени мы уже разработали основные элементы нашей концепции. Тем не менее для нас было очень полезно обменяться опытом со специалистами, работающими с программой Euro Trace.

9. За все эти годы нами было написано в общей сложности только три прикладные программы, касающиеся ввода данных, логического контроля и табулирования. Было сочтено, что настало время наконец завершить написание этих трех программ, с тем чтобы никогда (по крайней мере, таким образом, как это делалось ранее) не возвращаться вновь к ним на предмет совершенствования.

10. Несмотря на мощное сопротивление использованию

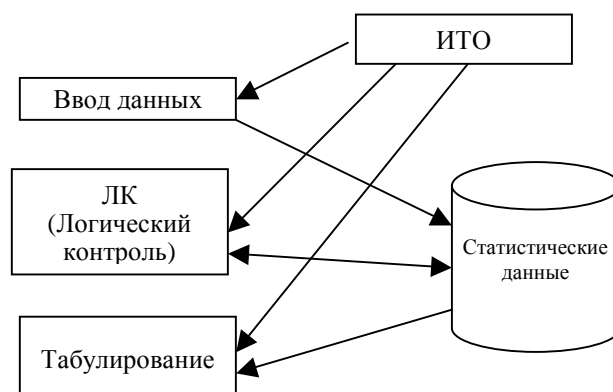


Диаграмма 2

универсальных программ, в прошлом мы довольно успешно применяли некоторые из них. Прежде всего речь идет о программе TABS - универсальной табулирующей программе, разработанной для платформы большой ЭВМ в Союзном статистическом управлении, и о программе GoDar, представляющей собой систему ввода и логического контроля данных для той же платформы, разработанной в Статистическом управлении Республики Сербии. Хотя оба эти решения давно устарели, существование такого рода традиции служит для нас вдохновляющим стимулом.

11. Было ясно, что данные необходимо хранить в реляционной базе данных. Проблема заключалась в том, что мы не знали, какую базу данных (DBMS) выбрать, поскольку в то время уже использовали две базы данных (DB2 IBM и SQL-Server Microsoft) и ничего не знали о возможностях увеличения их числа. По этой причине нам пришлось принимать решение независимо от физической реализации, что, с другой стороны, ограничило возможности использования лишь основным набором команд SQL, без какого бы то ни было усиления или совершенствования любой из баз данных DBMS.

12. Как показано на диаграмме 2, основная идея заключалась в том, чтобы "организовать" обследования таким образом, чтобы они сами "заботились" о своих данных и обеспечивали их обработку в отличие от существовавшего на тот момент положения, когда каждое обследование "было самостоятельным субъектом". Что касается технического решения, то показанный на диаграмме 2 блок ИТО представляет собой базу данных, в которой хранятся все соответствующие данные обследований, как операционные (такие, как описание данных, условия логического контроля, определения таблиц конечных данных), так и дескриптивные (методологии, описания полей и т.д.). Первым пилотным проектом стало обследование TU-11 по статистике туризма, и в настоящее время оно проводится на основе ИТО.

13. Хотя эта идея требует еще значительной доработки, а многие элементы необходимо усовершенствовать, написать или переписать, мы рады, что система доведена до операционного уровня, что свидетельствует о том, что идея в целом является правильной.

V. РЕАЛИЗАЦИЯ

14. Как уже упоминалось, ИТО работает с любой реляционной базой данных. Пожалуй, правильнее было бы сказать, что она работает с любой базой данных. Из-за отсутствия ограничений, требующих, чтобы данные находились только в одной базе данных, данные одного обследования могут находиться в различных (даже в физически различных) базах данных, чего вполне можно было ожидать, поскольку мы планируем иметь единые регистры, классификации и другие "консультационные таблицы" для всех обследований.

15. Возникла проблема с выбором языка программирования для реализации системы. В конечном счете выбор был остановлен на языке Microsoft Access, в версии VBA. Основным соображением в пользу принятия этого решения послужило то, что программа Microsoft Office уже установлена на всех ПК в Статистическом управлении, пользователи (статистики) знакомы с этой программой, и поэтому они могут выполнять некоторые "элементарные" операции в программе Access (что безусловно легче по сравнению, например, с DB/2).

16. Предполагается, что пользователь (статистик) выбирает обследование, с которым он хотел бы (и имеет право) работать, и осуществляет либо ввод данных, логический тест и выверку или производит расчет конечных результатов. Этап логического контроля является наиболее интересным, так как он позволяет пользователю либо проверить и изменить данные "локально" (на ПК пользователя) или "глобально" (непосредственно в базе данных, где эти данные хранятся, благодаря чему эту операцию могут одновременно выполнять несколько человек).

VI. ВЫВОДЫ

17. Хотя многое еще предстоит сделать, мы удовлетворены достигнутым прогрессом. Система в целом уже функционирует в Статистическом управлении Сербии и Черногории, и уже начата работа по ее реализации. Она предложена также Статистическому управлению Республики Черногория (МОНСТАТ), и мы надеемся, что вскоре она будет внедрена и в этом Управлении.

18. Мы открыты для любого рода обсуждений и хотели бы сотрудничать с другими статистическими управлениями с целью обмена опытом и приобретения новых знаний.
