

NATIONS UNIES

**COMMISSION ECONOMIQUE
POUR L'EUROPE**

ОБЪЕДИНЕННЫЕ НАЦИИ

**ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ
КОМИССИЯ**

UNITED NATIONS

**ECONOMIC COMMISSION
FOR EUROPE**

SEMINAIRE

СЕМИНАР

SEMINAR



Distr.
GENERAL

CES/SEM.43/10 (Summary)
3 March 2000

RUSSIAN
Original: ENGLISH

СТАТИСТИЧЕСКАЯ КОМИССИЯ и
ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ
КОМИССИЯ

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ
СТАТИСТИКОВ

Семинар по интегрированным статистическим информационным
системам и связанным с ними вопросам (ИСИС-2000)
(Рига, Латвия, 29-31 мая 2000 года)

Тема I: Создание хранилищ данных и разработка и использование
статистических баз данных в сетевой среде

СОЗДАНИЕ КОРПОРАТИВНОГО ХРАНИЛИЩА МЕТАДААННЫХ В БЮРО ПЕРЕПИСЕЙ США

Вспомогательный документ,

представленный Бюро переписей США¹

Резюме

1. На протяжении многих лет Бюро переписей США (БП), как и большинство других статистических организаций, приобретает и разрабатывает компьютерные системы для обработки результатов обследований. В результате была создана компьютерная среда для обработки результатов обследований, состоящая из значительного числа несоотносимых друг с другом систем, из которых лишь некоторые могут быть взаимосвязаны. Таким образом, в настоящее время на многих системах оценка или обработка их собственного(ых) набора(ов) данных производится путем использования конкретной

¹ Авторы: Самюэль Н. Хайсмит и Даниэль В. Гиллман.

не пригодной для совместного применения документации для этих наборов данных и процессов. Это приводит к возникновению ряда типичных сопутствующих проблем:

- Для преобразования файла, используемого в одной системе, в формат, предусмотренный для другой системы, требуется значительное время.
- Тот факт, что используется лишь весьма незначительный объем общей документации или процедур, приводит к естественному распространению различных систем, предназначенных для решения одной и той же проблемы.
- Стоимость подготовки нового обследования или переписи является очень высокой, если не удастся воспользоваться преимуществами решений, разработанных для ранее созданных систем.

2. На протяжении многих лет мы занимаемся разработкой решения этой проблемы на основе использования модели общего типа. В настоящее время это решение получило название корпоративного хранилища метаданных (КХМД) в БП. Следует надеяться, что использование КХМД позволит обеспечить такое описание обследования или переписи на протяжении всего экономического цикла, при котором с помощью любой прикладной системы, которую можно согласовать с КХМД, можно будет получать доступ к любой информации об обследовании или переписи, зарегистрированной в КХМД, и немедленно использовать ее. Мы поставили перед собой цель обеспечить возможности для полного документирования любого обследования и создать систему, способную понимать эту документацию, предоставлять несложный доступ к информации об обследовании и использовать ее без каких-либо изменений в прикладной системе.

3. В настоящем документе приводится описание способа решения вышеуказанной задачи и многих других проблем.

4. КХМД создана на основе значительного числа научных исследований и широкого сотрудничества. Сотрудничество БП со Швецией, Канадой и Австралией, а также участие в рабочем совещании ЕЭК ООН по метаданным привело к мысли о необходимости разработки модели коммерческих данных для обработки результатов обследований и переписей. С этой целью можно было бы использовать ряд методов, включая

организацию полностью распределенной и централизованной системы, которую можно было бы применять для создания хранилища метаданных. КХМД в БП должно выполнять такие же функции в отношении метаданных, как и карточки в каталоге библиотеки. Оно содержит данные об адресе и характеристиках любой зарегистрированной в нем информации.

5. В сотрудничестве со многими заинтересованными подразделениями Бюро переписей и под руководством компетентного консультанта по вопросам технологий использования хранилищ данных и метаданных была создана модель коммерческих данных (МКД), позволяющая описывать процесс обработки результатов обследований и переписей. Итоги этой совместной деятельности показали, что процессы, используемые в наших экономических организациях и организациях, занимающихся проведением переписей населения и демографических обследований, являются весьма сходными. В конечном итоге была обеспечена удачная разработка одной модели, позволяющей описывать процесс обработки результатов обследований и переписей в БП.

6. Одновременно мы заключили с частной компанией "Метадейта менеджмент инкорпорейтед" договор на создание формального регистра элементов данных на основе стандарта ИСО/МЭК 11179. Затем с помощью инструментального программного средства "Эрвин" этот регистр элементов данных был включен вместе с МКД в единую модель отношений объектов. На основе этой модели мы смогли автоматически создать базу данных для реализации данной модели.

7. Нам повезло, что возникла реальная необходимость в практической реализации этого проекта. Сотрудники Отдела переписей, проводимых с десятилетним интервалом, получив задание создать систему распространения данных для переписи населения 2000 года, решили развернуть модель КХМД, а не разрабатывать свою собственную систему. Они решили расширить нашу модель с целью обеспечить уникальные функциональные возможности и увеличить эффективность в соответствии с потребностями их прикладной системы. Действующая в настоящее время прикладная программа American FactFinder, доступ к которой можно получить с Web-сайта www.census.gov, является программой, полностью управляемой метаданными из хранилища данных, основывающегося на модели КХМД.

8. В декабре 1996 года Отдел статистических исследований решил разработать программу прототипного применения КХМД, создание которого намечалось в то время. Приложениями к сопутствующему интерфейсу являлись прикладная программа American FactFinder, инструментальное средство для распространения данных FERRET и система управления документацией Экономического директората.

9. Нами было использовано прототипное приложение с тем, чтобы продемонстрировать его ценность для программных областей деятельности БП. В результате был подписан ряд меморандумов о понимании с целью совместного осуществления деятельности по созданию КХМД.

10. На этом этапе Экономический директорат использовал КХМД для хранения метаданных, необходимых для его электронного АСУВ. Было предложено обеспечить опытное использование хранилища метаданных. В случае Экономического директората в ходе опытного использования хранилища метаданных основное внимание можно было бы сосредоточить на двух этапах процесса проведения переписи. Эта программа могла бы стать дополнительным средством для ввода метаданных с целью одновременного охвата таких аспектов, как сбор и распространение данных. Осуществление опытной программы использования хранилища метаданных для целей деятельности Экономического директората началось летом 1999 года и было завершено в декабре 1999 года. Эта программа получила широкое признание, и в настоящее время аналитики приступили к вводу данных, используя в этих целях указанный программный продукт.

11. Следующий шаг заключался в разработке архитектуры, определяющей КХМД, все интерфейсы и все вспомогательные инструментальные средства, которые мы предложили создать. Основные требования, предъявляемые к этой операционной среде, заключаются в следующем:

- обеспечить открытую архитектуру;
- обеспечить строгое соблюдение открытых стандартов;
- обеспечить строгое соблюдение требований БП по защите информации;
- обеспечить поддержку Web-браузеров для основанных на Web приложений КХМД;
- использовать комплексные системы программного обеспечения;

- обеспечить интеграцию с системами данных для новых отраслей промышленности;
- использовать, по возможности, COTS с целью поддержания расходов на низком уровне и сведения заказного проектирования к минимуму;
- использовать программное обеспечение с общекорпоративной лицензией БП, или программное обеспечение с лицензией, предусматривающей большое число инсталляций в БП, с целью поддержания расходов на низком уровне;
- обеспечить открытый ППИ;
- обеспечить обмен метаданными на основе открытых стандартов;
- обеспечить поддержку обмена метаданными между КХМД и другими системами и средствами программного обеспечения БП;
- обеспечить поддержку признанных в КХМД форматов ввода метаданных: XML и AFF;
- обеспечить использование расширяемой модели метаданных КХМД, согласующейся со стандартом ИСО/МЭК 11179;
- обеспечить средства для совместного использования модели метаданных КХМД в рамках БП;
- обеспечить средства для объединения неструктурированных метаданных с КХМД.

12. К началу 2001 года следует завершить процесс создания КХМД, вспомогательных инструментальных средств и интерфейсов и ввести их в эксплуатацию. В настоящее время Экономический директорат БП фактически уже использует их для обработки информации, собираемой в ходе проведения ежегодных обследований продукции обрабатывающей промышленности. Он создаст инфраструктуру, с помощью которой будут осуществляться многие прикладные программы изучения данных, манипулирования данными и распространения данных. После регистрации наборов данных и сопутствующих метаданных любая прикладная система, способная взаимодействовать с КХМД, сможет полностью использовать этот набор данных и все сопутствующие метаданные без каких-либо изменений в этой системе. Короче говоря, соответствующую деятельность можно будет осуществлять на основе принципа "определить один раз и использовать всегда".
