

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

**КОМИССИЯ
ЕВРОПЕЙСКИХ СООБЩЕСТВ**

КОНФЕРЕНЦИЯ СТАТИСТИКОВ ЕВРОПЫ

ЕВРОСТАТ

**Совместная рабочая сессия ЭКЕ и ЕВРОСТАТа
по статистическим метаданным
(Люксембург, 6-8 марта 2002 г.)**

Рабочий доклад № 25

Тема 3: метаданные и качество

МЕТАДАННЫЕ И КАЧЕСТВО

Дополнительный доклад

Представлен Национальным статистическим комитетом Кыргызской Республики¹

I. ВВЕДЕНИЕ

1. В докладе рассматриваются вопросы, связанные с работой, проводимой Национальным статистическим комитетом Кыргызской Республики (Нацстаткомитет) относительно измерения качества данных и способах их представления пользователям статистической информации, использованием Интернета для интегрирования качества данных с данными, необходимостью разработки стандартов для метаданных качества данных. Анализируется общее состояние дел, в соответствующих случаях дается их увязка с работой, осуществляемой Нацстаткомитетом в соответствии с Общей системой распространения данных (ОСРД) МВФ, указываются возникшие проблемы, а также предполагаемые направления работы на ближайшую перспективу.

II. КАЧЕСТВО РАСПРОСТРАНЯЕМЫХ ДАННЫХ

A. Метаданные

2. Качество статистических данных определяет пригодность ее для использования. Общепринятого определения качества данных не существует, и оно трудно поддается оценке. Механизмом для отслеживания качества статистических данных, правильного толкования получаемой по запросам статистической информации служат метаданные, представляющие собой описание используемых в настоящее время методов подготовки и распространения статистических данных.

3. Первым шагом в составлении метаданных в Кыргызской Республике явилась их подготовка в соответствии с Общей системой распространения данных (ОСРД) МВФ. Использование ОСРД в качестве основы развития статистики создает условия для своевременного получения и распространения надежных, исчерпывающих и доступных экономических, финансовых и социально-демографических данных. В целях выполнения одного из обязательств по участию в ОСРД совместными усилиями МВФ и четырех учреждений Кыргызской Республики подготовлены следующие метаданные:

- Нацстаткомитетом по *реальному* (национальные счета, агрегаты национальных счетов, индексы производства, индексы цен, показатели рынка труда) и *внешнему* (экспорт/импорт товаров) *секторам, социально-демографическим данным* (население, здравоохранение, образование, бедность);

¹ Автор: Кудобаев Зарылбек Исалыевич, Председатель Национального статистического комитета Кыргызской Республики

- Минфином по *бюджетному сектору* (операции центрального правительства, долг центрального правительства, бюджетные агрегаты центрального правительства, внешний долг и график обслуживания внешнего долга);
- Национальным банком - *финансовому* (обзор широкой денежной массы, агрегаты широкой денежной массы и кредита, агрегаты Национального банка, процентные ставки) и *внешнему* (платежный баланс, агрегаты платежного баланса, международные резервы, обменные курсы) *секторам*;
- Государственной комиссией по рынку ценных бумаг – по фондовому рынку, входящему в финансовый сектор.

4. Каждым вышеуказанным ведомством подготовлены метаданные по достоверности данных и доступу общественности с изложением действующей политики и практики.

5. Каким же образом измерения качества данных следует предоставлять пользователю статистической информации? Подход Нацстаткомитета к решению данного вопроса, в соответствии с ОСРД, обеспечивается следующим образом.

6. Во-первых, это предоставление подробной информации об определениях, классификациях, источниках данных, методах составления и использования международных нормативов, помещаемой в метаданных по *комплексным основам*. Данная информация, содержащаяся в метаданных, не включает конкретной оценки качества, но позволяет пользователям делать выводы в отношении аспектов качества данных.

7. Во-вторых, применительно к *категориям* данным и показателям параметр качества предусматривает два способа для конкретной оценки: **предоставление документации по методологии и источникам**, используемой при подготовке данных с изложением порядка составления данных и источников данных, и **распространение данных о детализации компонентов, сверках со смежными данными, статистических основах, обеспечивающих статистические перекрестные проверки и гарантирующих обоснованность**.

8. Эти способы являются, на наш взгляд, наиболее приемлемыми. Пользователи, в распоряжении которых предоставляются лишь статистические оценки, как правило, не в состоянии оценить их точность на основе числовых значений. Наличие документации по методологии и источникам, лежащим в основе статистики, является ключевым фактором, позволяющим пользователям оценить преимущества и недостатки данных, судить о существенных отклонениях от международных методологических рекомендаций.

9. Так, во многих случаях выпуск статистических данных в Кыргызской Республике сопровождается техническими и краткими комментариями методологического характера о составляющих статистического показателя, содержит информацию об источниках, используемых при подготовке данных. Более подробные пояснения по методологии содержат годовые и тематические публикации.

10. При опубликовании предварительных данных указывается, что данные являются предварительными. Сообщение делается и в случае пересмотра данных. Заблаговременного уведомления о существенных изменениях в методологии, как правило, не делается, однако, при возникновении изменений выпуск данных сопровождается краткими пояснениями.

11. Детализация компонентов распространяемых данных позволяет производить перекрестные проверки и имеет самостоятельную ценность для углубленного анализа. При достаточном объеме опубликованных детальных данных на уровне ниже агрегатов, особенно в сочетании с документацией, пользователь имеет возможность проверить обоснованность данных. Степень сопоставимости повышается за счет интегрирования данных на общей основе, такой как СНС, и проверяется посредством проведения сопоставлений связанных рядов, которые

позволяют наблюдать за одним и тем же явлением под различными углами зрения. Например, в годовой публикации по национальным счетам Кыргызской Республики пользователь имеет доступ к информации по основным категориям конечного использования, доле основных отраслей экономики ко всему ВВП, региональным объемам валового продукта по отраслям и по компонентам конечного использования в разрезе регионов республики, счетам внешних операций. По индексу физического объема промышленной продукции наряду с агрегированным индексом публикуются индексы по основным видам деятельности и отдельным важнейшим продуктам. Приводятся данные по индексу физического объема как отдельно по видам продукции, по видам деятельности сельского хозяйства, так и в целом по сельскому хозяйству, что позволяет сопоставлять и контролировать данные пользователем. Индекс потребительских цен распространяется в разбивке по товарам и группам, что позволяет проводить проверки обоснованности. С целью сравнения по экспорту и импорту товаров предоставляется подробная географическая информация (по стране происхождения – для импорта, по стране назначения – для экспорта), информация по товарам публикуется по кодам Товарной номенклатуры внешнеэкономической деятельности (ТНВЭД) и Международной стандартной торговой классификация (СМТК).

12. Сопоставления и сверки в пределах республики позволяют проводить данные, получаемые из разных источников, например, данные по занятости, получаемые статистикой трудовых ресурсов, переписью населения и обследований домашних хозяйств.

13. В перспективе Нацстаткомитетом планируется выпуск «Методологических положений по статистике», которые будут содержать работы методологического характера, разработанные и используемые статистической службой республики, что даст дополнительную информацию о методологических основах подготавливаемых и распространяемых статистических данных.

14. Важным элементом качества является обеспечение своевременности, которое сводится к наличию заблаговременно намеченных сроков распространения данных и к соблюдению этих сроков. С целью обеспечения различных групп пользователей качественной статистической информацией осуществляется планирование статистической деятельности. Ежегодно Правительством Кыргызской Республики утверждается Программа статистических работ, в соответствии с которой Нацстаткомитет осуществляет свою основную деятельность. При ее составлении учитываются международные и межведомственные информационные обязательства, основополагающие принципы официальной статистики. Программа определяет состав информации, ее периодичность и сроки представления информации внешним пользователям, круг исполнителей.

15. Начиная с марта 1998г. Нацстаткомитет осуществляет новую **политику публикации**, которая предусматривает публикацию и распространение данных статистической информации Нацстаткомитета согласно графику публикаций. График публикаций помещается в годовых, тематических публикациях и Web-сайте Нацстаткомитета.

16. Пользователь статистических данных должен обеспечиваться информацией для оценки качества и улучшений качества. В качестве краткосрочных и долгосрочных планов совершенствования статистики, улучшения качества данных официально утверждена постановлением Правительства Кыргызской Республики в декабре 2000 года и реализуется «Программа развития статистики в Кыргызской Республике на 2001-2005гг.». Основные мероприятия программы включены в метаданные Кыргызской Республики в качестве планов совершенствования.

17. В целях качественного и своевременного выполнения вышеуказанных планов совершенствования министерствами и ведомствами, участвующими в ОСРД, подготовлены и Национальным координатором ОСРД по Кыргызской Республике утверждены мероприятия по их выполнению. Нацстаткомитет, выполняя координационные функции по ОСРД, ежеквартально подводит итоги выполнения данных мероприятий.

18. Метаданные, являясь основным средством предоставления информации о качестве данных, требуют улучшения качества самих метаданных. В связи с этим перед кыргызскими статистиками в настоящее время стоят проблемы создания централизованно управляемой системы метаданных, охватывающей не только показатели регламентируемые ОСРД МВФ, а полностью соответствующей статистическим базам данных, обеспечения непрерывного их обновления.

В. Интернет и использование его для интегрирования качества данных с данными

19. Расширяются каналы предоставления пользователям статистической информации на электронных носителях, в том числе через Интернет. Нацстаткомитет имеет возможность использовать Интернет для:

- приема/передачи электронной почты – E-mail;
- распространения статистических публикаций;
- распространения статистических данных;
- сбора данных;
- общения с пользователями и т.д.

20. Нацстаткомитет входит в список министерств, ведомств, учреждений, подключенных в Информационную систему Государственного управления (ИСГУ). Эта система обеспечивает доставку электронной почты во многие министерства, ведомства, учреждения и их подразделения на территории Кыргызской Республики за счет выделенных каналов связи. В настоящее время данная система находится еще в стадии развития. Имеется лишь один канал связи с одним областным управлением государственной статистики. Доставка статистической информации по выделенным каналам связи позволяет обеспечить надежность, качество и своевременность доставки корреспонденции. Остальные областные управления государственной статистики оснащены модемами для передачи введенной статистической информации посредством электронной почты E-mail в Главный вычислительный центр Нацстаткомитета по коммутируемым каналам связи.

21. С марта 1998 года в Интернете существует страничка Нацстаткомитета (www.nsk.bishkek.su), которая содержит общую информацию и данные по реальному, бюджетному, финансовому, внешнему и социальному секторам, а также содержит файлы в Excel из базы данных Нацстаткомитета. Ежемесячно обновляется информация по основным социально-экономическим показателям республики. Пользователи могут получить через Интернет метаданные на русском и английском языках. Открыт новый раздел “Вопросы методологии”, где помещаются работы методологического характера, разработанные Нацстаткомитетом и Институтом статистических исследований.

22. С введением Страницы к нам начали поступать запросы от библиотек, студентов, граждан, различных организаций, в том числе международных. Для изучения потребностей и требований пользователей мы планируем вести статистику запросов.

23. При участии советника – резидента МВФ была разработана Система статистических показателей и внедрена в качестве централизованной базы данных для подготовки публикаций. В ней содержится более 300 статистических таблиц. В дальнейшем мы планируем расширить базу данных, включив в нее все обработанные статистические данные Нацстаткомитета.

24. Конечно, бывают ситуации, когда потребителю необходима информация, которой нет в базе. У нас существует база первичных данных, но она находится в разрозненном виде. В перспективе планируется обеспечить через Интернет доступ пользователей к единой информационно-статистической базе данных, а создание специальной программы позволит формировать таблицы по запросу пользователей и удовлетворять возрастающий спрос.

25. Таким образом, информация, помещаемая на страничке Нацстаткомитета в Интернет, кроме статистических данных содержит информацию методологического характера, описание используемых в настоящее время методов подготовки и распространения данных, что позволяет пользователям судить не только о данных, но и их качестве. Повышается степень своевременности представления информации, обеспечивается беспрепятственное получение метаданных, осуществляется более приемлемая интеграция информации.

26. Велико значение Интернета для связи с международными организациями, статистическими службами других стран, как в части использования статистических данных, так и метаданных. С целью повышения эффективности использования метаданных, в метаданные, размещенные на Web-страницах международных организаций, могут быть включены ссылки на более подробные метаданные, находящиеся на национальных Web-сайтах. Примером тому могут быть метаданные на Доске объявлений по стандартам распространения (ДОСР) МВФ.

27. В целом же, статистиками Кыргызской Республики влияние использования Интернета для интегрирования качества данных с данными еще не достаточно изучено. Однако можно уже заранее предвидеть недостатки обследования через Интернет, таких как доступ не всех обследуемых единиц к Интернету. Так, при проведении обследований домохозяйств данный недостаток может оказать значительное влияние на все базовые компоненты обследования, например расслоение, пол, уровень образования, что в итоге скажется на качестве данных. В случае обследований предприятий эта проблема, как нам представляется, является менее острой. Применение сетевых средств для сбора информации находится у нас в стадии разработки.

28. Кроме того, одной из проблем в сборе данных через Интернет является использование единой терминологии в области метаданных и создание каталогов в качестве основы для баз статистических метаданных. С учетом маркетинговых соображений, а также новых средств программного обеспечения нам предстоит разработать новые подходы к электронному сбору данных непосредственно от обследуемых единиц.

С. Стандарты для метаданных качества данных

29. Бесспорным является тот факт, что статистики Кыргызской Республики уделяют первостепенное внимание вопросам достижения соответствия международным стандартам. Это дает возможность проводить международные сравнения и анализ экономической ситуации в стране всеми заинтересованными сторонами, наконец, это позволяет поддерживать свою независимость от правительственных органов, защищая их таким образом от соблазна манипулировать концептуальными положениями.

30. Термин «международные стандарты» охватывает концепции, классификации, систему национальных счетов, а также ряд статистических продуктов в различных областях деятельности. Но существуют ли общепринятые стандарты качества данных, директивного времени выпуска материалов и т.д.? В этой связи необходимо упомянуть стандарты для метаданных об основных статистических рядах Международного валютного фонда: Общая система распространения данных (ОСРД) для всех стран и Специальный стандарт на распространение данных (ССРД) для большинства развитых стран. Метаданные, представляемые в рамках этих стандартов, бесспорно, характеризуют качество распространяемых данных.

31. Кыргызская Республика в качестве стандарта распространения статистических данных в настоящее время, как уже было сказано выше, использует Общую систему распространения данных (ОСРД), следующим шагом будут принятие Специального стандарта распространения данных МВФ и статистических требований ЕС.

32. Метаданные охватывают весь производственный процесс, начиная со сбора данных и заканчивая их распространение. Пользователям статистической информации всех категорий, как правило, необходимы метаданные для осуществления следующих функций:

- установления наличия определенных данных;
- получения помощи в поиске информации;
- получения сведений об источниках данных, определениях, классификациях, методах составления и использования международных нормативов;
- получения сведений о методологии, используемой при подготовке данных;
- интерпретации информации;
- распространении информации о детализации компонентов, сверках со смежными данными, статистических основах, обеспечивающих статистические перекрестные проверки и гарантирующих обоснованность;
- в случае необходимости, последующей обработки информации.

33. Метаданные характеризуются значительными различиями в зависимости от страны и применяемой методологии. Многочисленные связи статистических служб с пользователями подчеркивают важность разработки общепринятой основы для определения качества данных с целью более эффективной оценки сопоставимости распространяемой статистической информации как на национальном, так и международном уровнях.

III. ЗАКЛЮЧЕНИЕ

34. Мы полагаем, что всеобъемлющая система оценки качества данных должна использоваться в первую очередь в интересах национальной статистической службы. Все запросы о предоставлении информации внешним пользователям, на обязательной или добровольной основе, должны удовлетворяться за счет информации из одной и той же базы статистических данных. Данные по качеству данных должны подвергаться всестороннему анализу и содержать не только количественные показатели, но и данные описательного характера в отношении применяемых процессов и процедур, которые позволяли бы заинтересованным сторонам определять степень качества данных.

35. Web-технологии обеспечивают возможности для повышения качества представления официальной статистики. Использование Интернета на национальном и международном уровнях требует согласования усилий в области разработки стандартов. Стандартизированные концепции статистических данных и метаданных, создание централизованно управляемых систем метаданных, должны стать одними из важных предварительных условий для удовлетворения ожиданий пользователей относительно качества статистических данных.