



**ЭКОНОМИЧЕСКИЙ
И СОЦИАЛЬНЫЙ СОВЕТ**

Distr.
GENERAL

CES/2003/12
7 April 2003

RUSSIAN
Original: ENGLISH

СТАТИСТИЧЕСКАЯ КОМИССИЯ и ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ

Пятьдесят первая пленарная сессия

(Женева, 10-12 июня 2003 года)

ДОСТУП К МИКРОДАНЫМ - ВОПРОСЫ, ОРГАНИЗАЦИЯ И ПОДХОДЫ

Документ представлен Статистическим бюро Австралии (СБА)¹

I. СПРАВОЧНАЯ ИНФОРМАЦИЯ

1. Обеспечение конфиденциальности имеет важное значение не только по юридическим и этическим причинам, но также и для поддержания доверия общественности. Увеличение спроса на детализированные данные в совокупности с ростом технологических мощностей и возможностей, а также наличие потенциально увязываемых наборов данных значительно затрудняют решение задачи обеспечения конфиденциальности микроданных. Сегодня практически невозможно публиковать микроданные, содержащие полезную информацию без потенциального риска идентификации. Данная проблема усугубляется наличием наборов продольных данных. Мы более не можем полагаться только на использование различных видов модификации данных с целью устранения риска идентификации.

2. С другой стороны, постоянно увеличивается спрос на доступ к микроданным в целях проведения различных видов исследований и вторичного анализа данных. Рост вычислительных мощностей расширяет возможности исследователей проводить анализ такого рода.

¹ Автор: Денис Тревин.

3. Существует целый ряд мотивов для поиска ответа на вопрос о том, как наиболее рационально обеспечить доступ исследователей к микроданным.

- Ценные (и высококачественные) данные не используются в полной степени.
- Исследователи могут попытаться собрать замещающие наборы данных для получения микроданных, что приведет к нерациональному использованию государственных ресурсов (для получения данных возможно более низкого качества).
- Правительственные ведомства могут обратиться к альтернативным поставщикам данных с целью получения результатов обследований для исследовательских и аналитических целей, что приведет к получению низкокачественных данных, которые вряд ли будут содействовать поддержанию престижа национальной статистики.

4. Еще одним важным аспектом, требующим изучения, является невероятный по своим возможностям ценный аналитический потенциал связанных наборов данных, включая связи с наборами данных СБА.

5. Вышеперечисленные факторы заставили нас переосмыслить наш подход к предоставлению доступа к микроданным. Это касается и различных других НСУ, многие из которых находятся в процессе пересмотра своих практических методов. В настоящем документе описываются осуществляемые или планируемые нами меры. Для обеспечения доступа к данным обследований домашних хозяйств и предприятий могут потребоваться различные стратегии. В настоящем документе приводится описание положения дел в СБА, которое, как мы надеемся, может быть полезным для решения проблем, стоящих перед другими НСУ.

6. До того, как приступить к данному описанию, необходимо подчеркнуть, что принимаемые меры должны быть приемлемыми с точки зрения как законодательства, так и интересов общественности. Закон можно изменить, однако это - длительный и сложный процесс, чреватый появлением лишних проблем. Исходя из этого, наш подход заключается в проведении работы в рамках действующего законодательства.

II. КРАТКОЕ ОПИСАНИЕ ПОТРЕБНОСТЕЙ

7. В идеале пользователи хотели бы иметь:

- возможность работать в интерактивном режиме с данными;
- доступ к экспертам СБА и надлежащей документации по описанию данных;
- доступ к большему числу наборов данных;
- высококачественные данные о представляющих для них интерес совокупностях и переменных и определенную информацию об источниках погрешностей;
- своевременные публикации данных; и
- более широкий доступ к связанным наборам данных (включая данные, увязанные во времени).

8. Исследователи неоднократно указывали на то, что наш подход к публикации микроданных является избыточно консервативным. С учетом этого в последнее время мы занимались изучением вопроса о том, каким образом мы можем расширить доступ к микроданным, не нанося ущерба нашей высокой репутации в области защиты конфиденциальности (и оставаясь при этом в рамках закона), что имеет большую важность для поддержания высокого уровня сотрудничества в рамках наших обследований.

9. Использование связанных наборов данных предполагает возможность осуществления СБА функций хранителя не принадлежащих ему наборов данных в целях обеспечения надлежащей защиты конфиденциальности. Хотя это полностью согласуется с задачами национальной статистической службы, выполнение таких функций потребует разработки соответствующей политики и оперативных процедур. Эти вопросы подробно рассматриваются в пунктах 24-28.

III. СРЕДСТВА УДОВЛЕТВОРЕНИЯ ПОТРЕБНОСТЕЙ

10. Существует большое число вариантов или каналов распространения, которые обладают различной степенью защиты от посягательств на принцип конфиденциальности. Нижеприводимый перечень начинается с вариантов, которые в значительной мере

опираются на безопасные данные, а заканчивается вариантами, опирающимися в большей степени на безопасную среду, включая использование юридических обязательств со строгими санкциями за их нарушения.

11. Доступность и удобство для исследователей также являются различными в зависимости от варианта.

12. Публикация микроданных, которая является отдельной темой данного заседания, является одним из ключевых элементов обеспечения доступа в исследовательских целях. Статистическое законодательство позволяет нам публиковать микроданные, но только "таким образом, который бы делал маловероятной идентификацию конкретных лица или организации, к которым они относятся". Кроме того, пользователи должны также брать на себя определенные обязательства. Тем не менее существует ряд способов получить доступ к микроданным, соблюдая при этом юридические ограничения. Они поясняются ниже и кратко описываются в таблице 1.

13. Для оказания нам содействия в оценке приемлемости риска идентификации (т.е. "маловероятности идентификации") применительно к этим вариантам распространения, связанным с публикацией микроданных, была создана группа по анализу микроданных. Она занимается изучением двух ключевых зон риска:

- предупреждение спонтанной идентификации; и
- предупреждение риска увязки.

14. Правовое заключение гласит, что принятие пользователями на себя юридических обязательств не предпринимать определенные действия согласуется с принципом "маловероятности", и такая возможность должна учитываться при проведении вышеупомянутой оценки.

15. В нижеследующих разделах подробно описываются преимущества и недостатки каждого варианта.

Таблица 1: Каналы распространения данных в исследовательских целях

Канал распространения	Примечания
1. Стандартные статистические материалы	Как правило, в форме таблиц. Ограничивают возможности пользователей проводить различные виды анализа.
2. Кубы данных	Обеспечивают более высокий уровень детализации и гибкости с точки зрения разработки исследователями своих собственных таблиц.
3. Специальные информационные услуги	По запросу пользователей, как правило, за незначительную плату.
4. Обезличенные файлы единичных записей (ОФЕЗ)	Данные не поддаются идентификации. Публикуются на КД-ПЗУ. Эквивалент тому, что обычно называют публикациями микроданных.
5. Лаборатория дистанционного доступа к данным (ЛДДД)	Возможен доступ к более подробным ОФЕЗ, поскольку обеспечивается более строгий контроль с целью предупреждения увязки с внешними базами данных.
6. Лаборатория данных СБА	Также обеспечивает доступ только к неидентифицируемым данным.
7. Сотрудничество	Означает взаимодействие с одним из сотрудников СБА, а не прямой доступ к данным.
8. Анализ на месте	На деле работа в помещениях СБА в качестве сотрудника СБА. Это возможно лишь в том случае, если исследователь оказывает помощь СБА в выполнении его функций.

16. Стандартные статистические материалы

- Что это предусматривает? - Публикация статистических материалов, обычно в форме таблиц, в печатном и/или электронном формате.
- Защита конфиденциальности - Речь идет о безопасных данных. Применяются стандартные процедуры обеспечения конфиденциальности СБА.
- Преимущества - Удобные и дешевые. Позволяют составить впечатление о полной совокупности данных. Расширение возможностей загрузки электронных данных повышает удобство их использования для последующего анализа. Легкодоступные для широкого круга исследователей. Низкие издержки для СБА.
- Недостатки - Ограничивают число возможных типов анализа. Невозможно проводить анализ, опирающийся на микроданные.
- Текущее состояние - Увеличение объема данных в этом формате на вебсайте. Повышение доступности вспомогательных метаданных.
- Возможности использования - Стандартные материалы не следует недооценивать в качестве удобного средства поддержки исследований. Они должны рассматриваться в качестве одного из ключевых компонентов стратегии распространения в отношении всех статистических материалов.

17. Кубы данных

- Что это предусматривает? - Публикация подробных статистических матриц, которые уже были обезличены. Это является более подходящей формой публикации в том случае, когда защита конфиденциальности может быть автоматизирована, в частности в отношении малых ячеек (например, перепись населения). Специальные положения о конфиденциальности в отношении данных о торговле также позволяют публиковать данные в этой форме.
- Защита конфиденциальности - Речь идет о безопасных данных. Применяются стандартные процедуры обеспечения конфиденциальности СБА (если только не существуют специальные положения в отношении некоторых данных, например о торговле).
- Преимущества - Разумно удобный доступ к более подробным данным по сравнению со стандартными статистическими материалами.

- Недостатки - Те же, что и в случае стандартных статистических материалов. Кроме того, проректирование качественных кубов данных является непростой задачей. Некоторые исследователи находят кубы данных трудными для использования. Невозможность подготовки обезличенных кубов данных по многим статистическим материалам.
- Текущее состояние - Мы медленно увеличиваем число предоставляемых кубов данных. Разработчики кубов данных накапливают опыт.
- Возможности использования - Как правило, кубы более полезны в отношении данных о лицах, чем данных о предприятиях. В случае некоторых статистических материалов они должны рассматриваться в качестве составного компонента стратегии распространения.

18. Специальные информационные услуги

- Что они предусматривают? - Публикация статистических материалов, не обязательно таблиц, по запросу исследователей.
- Защита конфиденциальности - Речь идет о безопасных данных. Условием предоставления их исследователям является обеспечение защиты конфиденциальности.
- Преимущества - Данные и форма доставки могут определяться с учетом потребностей пользователей.
- Недостатки - Могут являться дорогостоящими для некоторых исследователей (и СБА с точки зрения обслуживания запросов). Возможности анализа являются ограниченными в силу невозможности работы в интерактивном режиме. Исследователь не может применять свои собственные процедуры корректировки микроданных (например, резко выделяющихся значений). Переход к различным процедурам анализа данных может быть медленным.
- Текущее состояние - Предлагается в качестве услуги, но спрос является незначительным. Попытки развивать такого типа услуги не предпринимаются, за исключением ключевых клиентов и избранных областей (например, региональная статистика).

- Возможности использования - Как правило, в отношении табличных материалов, которые не предоставляются в рамках стандартных материалов, и когда доступ к микроданным невозможен. Другие виды анализа, по всей видимости, могут проводиться в рамках сотрудничества (см. ниже).

19. Обезличенные файлы единичных записей (ОФЕЗ)

- Что это предусматривает? - Публикация файлов микроданных на КД-ПЗУ, в которые были внесены изменения для исключения вероятности идентификации отдельных лиц или организаций.
- Защита конфиденциальности - Группа по анализу микроданных представляет консультационные заключения по корректировкам, необходимым для защиты конфиденциальности данных. Они могут касаться рекомендаций относительно использования методов корректировки данных, таких, как исключение некоторых переменных, снижение уровня детализации некоторых переменных (в частности, географической), исключение некоторых характеризующихся высокой степенью идентифицируемости лиц и внесение случайные помех. Дополнительной мерой защиты конфиденциальности является взятие на себя всеми исследователями, получающими доступ к микроданным, соответствующих юридических обязательств, за нарушение которых предусматривается применение санкций (включая лишение права доступа к микроданным) к исследователю и, возможно, его учреждению. Возможно также привлечение к судебной ответственности.
- Преимущества - Большая гибкость и удобство для исследователей.
- Недостатки - Не всегда обеспечивается искомый уровень детализации. Как правило, ОФЕЗ отсутствуют в случае данных о предприятиях. Было отмечено небольшое число случаев нарушения обязательств (однако без идентификации индивидуальных данных).
- Текущее состояние - ОФЕЗ будут оставаться важным каналом распространения данных в целях исследований и вторичного анализа. Спрос на них является высоким. Мы стремимся повысить оперативность наших публикаций.
- Возможности использования - ОФЕЗ рассматриваются в качестве одного из стандартных материалов обследований домохозяйств. Используются избирательно в отношении других обследований, данные которых остаются

полезными в исследовательских целях после применения процедур защиты конфиденциальности.

20. Лаборатория дистанционного доступа к данным (ЛДДД)

- Что это предусматривает? - Выполнение заданий, представленных зарегистрированными пользователями через Интернет, на обработку ОФЕЗ СБА и предоставление им результатов анализа после осуществления в значительной степени автоматизированных проверок для обеспечения конфиденциальности. Услуга, аналогичная ОФЕЗ, за исключением возможности предоставления доступа к более подробным данным, поскольку риск увязки может контролироваться в силу того, что данные не покидают СБА. Возможности ограничены набором аналитических программных средств, предоставляемых ЛДДД (например, SAS и SPSS). Материалы просматриваются вручную до передачи пользователям.
- Защита конфиденциальности - Консультативные заключения группы по анализу микроданных. Ручная проверка материалов, усиленная автоматическими присоединенными процедурами, для выявления материалов, которые могут требовать более тщательной проверки. Хранение данных и результатов проверки. Предусматривается подписание юридических обязательств. Санкции против нарушителей.
- Преимущества - Доступ к более подробным данным. Доступ к аналитическому программному обеспечению, которое может отсутствовать у исследователя. Бесплатные мощности для обработки данных.
- Недостатки - Неудобства по сравнению с ОФЕЗ. Некоторые задержки с публикацией материалов. Большие издержки по администрированию для СБА.
- Текущее состояние - Введение в эксплуатацию в апреле 2003 года. Будет совершенствоваться с учетом замечаний пользователей. Количество предлагаемых в рамках этого сервиса наборов данных будет возрастать со временем.
- Возможности использования - Будет использоваться вместо ОФЕЗ в тех случаях, когда риск увязки данных ОФЕЗ будет чрезмерно высоким, а использование обязательств/санкций сопряжено с риском. Например, ЛДДД

может использоваться в отношении связанных файлов данных, в особенности, если один из связанных файлов является внешним.

21. Лаборатория данных СБА

- Что это предусматривает? - Сервис, аналогичный ЛДДД, за исключением отсутствия возможности загрузки данных единичных записей (в случае ЛДДД возможна загрузка до 30 записей в целях выявления выбросов и т.д.). Отметим, что ситуация является отличной во многих других странах, в которых подписка о неразглашении позволяет получить доступ в помещениях статуправления к необезличенным данным единичных записей. У нас это возможно только в том случае, если исследователь действительно помогает статистику осуществлять его функции, а его статус позволяет рассматривать его в качестве сотрудника СБА. Это означает оплату его услуг.
- Защита конфиденциальности - Аналогична ЛДДД, за исключением более тщательного контроля за материалами; данные единичных записей не могут покидать пределы СБА.
- Преимущества - Доступ к данным, которые могут не предоставляться в рамках сервисов ОФЕЗ или ЛДДД (например, продольные файлы данных). Прямой доступ к экспертам СБА.
- Недостатки - Неудобства, связанные с работой в помещениях СБА. Значительные трудозатраты для сотрудников СБА по администрированию, в особенности с учетом существования девяти отделений.
- Текущее состояние - Уже функционирует. Главным образом используется в отношении продольных файлов данных, в особенности, когда единица выборки или некоторые данные могут быть взяты из административной системы другого агентства.
- Возможности использования - Используется только в тех случаях, когда сервисы ОФЕЗ или ЛДДД рассматриваются в качестве неподходящих для определенного набора данных или когда исследователь предпочитает данную форму работы, а СБА готово оказать такую поддержку.

22. Сотрудничество

- Что это предполагает? - Совместную работу с исследователем по подготовке материалов (часто публикуемых материалов), представляющих интерес для СБА. Может как являться, так и не являться статистическим материалом, публикуемым СБА. Действующие правила обычно не запрещают исследователям публиковать или представлять результаты этой работы вне СБА, в том числе в научных журналах.
- Защита конфиденциальности - Исследователи не имеют прямого доступа к данным единичных записей. Прямой доступ к данным имеет сотрудник СБА, сотрудничающий с исследователем.
- Преимущества - Взаимные выгоды от сотрудничества. Реальная передача знаний. Исследователь может осуществлять основную часть своей работы вне помещений СБА. Такое сотрудничество может содействовать предоставлению финансовых ресурсов исследователю на его исследовательские цели. Затраты исследователя как правило являются небольшими. Потенциальный доступ к грантам Австралийского совета по исследованиям.
- Недостатки - Отсутствие прямого доступа к данным. Возможности сотрудничества ограничиваются совместными проектами, представляющими интерес для СБА.
- Текущее состояние - Разработана политика в области проведения совместных работ. Был создан Отдел анализа, который действует уже на протяжении более четырех лет и насчитывает около 30 сотрудников. Это позволило начать реальное сотрудничество с исследователями. Ранее такие работы осуществлялись на специальной основе.
- Возможности использования - Данный вариант используется в тех случаях, когда результатом сотрудничества являются материалы, представляющие взаимный интерес. В случае некоторых высокоприоритетных проектов СБА может само искать сотрудничества с исследователями. Что касается конфиденциальности, то принципы, призванные регулировать совместную работу, должны согласовываться с принципами государственных закупок, содействовать повышению полезности статистических данных и обеспечивать беспристрастность и транспарентность и защиту репутации СБА.

- Что это предусматривает? - СБА может производить наем исследователей в качестве "сотрудников", если они решают задачи, оказывающие помощь СБА в его деятельности. В этом случае они могут иметь доступ к единичным записям при условии соблюдения правил защиты конфиденциальности, применимых к другим сотрудникам СБА. Данный вариант может быть уместным в том случае, когда СБА желает подготовить материал, с помощью которого исследователь восполнит выявленный пробел в экспертизе. Как правило, могут заключаться соглашения, позволяющие исследователям публиковать некоторые части своей работы с разрешения СБА в других изданиях.
- Защита конфиденциальности - Применяются те же положения о защите конфиденциальности, что и к сотрудникам СБА. Их нарушение предусматривает суровые санкции.
- Преимущества - Обеспечивает исследователям доступ к данным единичных записей. Взаимные выгоды от совместной работы.
- Недостатки - Большая часть работы должна осуществляться в помещениях СБА. Тематика ограничивается вопросами, представляющими непосредственный интерес для СБА. Некоторые ограничения на материалы исследований. Не всегда существует возможность найма в качестве сотрудника СБА.
- Текущее состояние - Этот вариант используется редко. Последние изменения в порядке деятельности государственных служб упрощают его использование.
- Возможности использования - В тех случаях, когда СБА выступает с инициативой найма исследователя для оказания ему помощи в его статистической деятельности. (Несомненно, возможны и взаимные выгоды.)

IV. СВЯЗАННЫЕ НАБОРЫ ДАННЫХ

24. Связанные наборы данных являются особым типом наборов микроданных, к которым пользователи могут пожелать получить доступ. В данном случае речь идет об использовании методов увязки для объединения единичных записей с целью формирования набора сводных данных. Сводные данные могут опираться на жесткую увязку с использованием идентификаторов или статистическую увязку с использованием определенной комбинации переменных (например, географических признаков, возраста, пола, характеристик домохозяйства). Оба этих метода создают риски с точки зрения

защиты конфиденциальности. Методы жесткой увязки представляют, естественно, большую опасность, однако проведенное нами исследование свидетельствует об удивительно высокой пропорции точных совпадений с использованием методов статистической увязки, в особенности в случае файлов, которые содержат данные о структуре домохозяйства.

25. Связанные наборы данных могут включать в себя:

- a) увязку наборов данных СБА;
- b) увязку набора данных СБА и данных, не принадлежащих ему; или
- c) увязку наборов данных, не принадлежащих СБА.

В случаях a) и b) СБА должно выполнять функции надзора и предоставлять доступ через упомянутые в настоящем документе каналы распространения. В случае c) функции надзора необязательно должно выполнять СБА, однако выполнение таких функций сулит определенные выгоды. Мы располагаем законодательством, которое может быть использовано для заключения соглашений о предоставлении доступа к таким наборам данных и обеспечении защиты их конфиденциальности. Кроме того, наша репутация такова, что мы пользуемся высоким доверием со стороны общественности в качестве надежного хранителя информации. Мы также располагаем инструментами и системами для обеспечения доступа.

26. Связанный набор данных может открывать значительные аналитические возможности, которые иллюстрируют нижеприведенные примеры:

- изучение связей лица с различными учреждениями, скажем, на основе увязки данных о медицинских услугах, оказанных врачами, больницами, частными лечебницами и т.п.;
- изучение связей между затратами, выпуском и результатами за счет сведения воедино информации о деятельности правоохранительных органов, судов и тюрем; и
- изучение динамики во времени за счет формирования продольной базы данных.

27. В целях создания таких наборов данных, а также работы с ними необходимо установить некоторые дополнительные принципы. В качестве ключевых можно выделить следующие принципы:

- Соответствие задаче СБА использовать статистическую информацию для эффективной поддержки процесса обоснованного принятия решений, исследований и обсуждений.
- Наглядные статистические выгоды.
- Честность и открытость в отношении прикладных видов использования.
- Публикация статистических материалов на основе каждого связанного набора данных.
- Поддержание доверия общественности за счет обеспечения соблюдения законодательства, регулирующего деятельность СБА, законодательства в области конфиденциальности и других соответствующих норм.

28. Мы занимаемся изучением вопроса о создании комитета по этике для оказания нам помощи в принятии решений в этой области.

V. НАШИ ПЛАНЫ

29. До недавнего времени подход к использованию каждого из каналов распространения в целях оказания поддержки внешним исследователям являлся следующим:

- Стандартные статистические материалы - в качестве стандартной услуги предоставлялись во всех областях статистики.
- Кубы данных - в состоянии разработки, но имелись в наличии по некоторым статистическим рядам (например, в области демографии, рабочей силы).
- Специальные информационные услуги - предлагались, но использовались лишь от случая к случаю. Обычно для разработки подобных таблиц.
- ОФЕЗ - регулярно готовились на основе результатов обследований домохозяйств, а в некоторых случаях - на основе результатов обследований предприятий, что, однако, требовало снижения уровня детализации публикуемых данных в силу значительного риска увязки. Также требуют принятия строгих юридических обязательств в отношении публикации данных.

- ЛДДД - сервис отсутствует.
- Лаборатория данных СБА - использовалась время от времени, но не пропагандировалась.
- Сотрудничество - использовалось только от случая к случаю в прошлом, однако в последний год было заключено около 40 соглашений о сотрудничестве. В их число входят 12 соглашений о сотрудничестве, заключенных в рамках Аналитической программы австралийской переписи.
- Анализ на месте - использовался редко.

30. В будущем для оказания поддержки внешним исследователям будут использоваться все восемь каналов. С учетом их стоимости мы попытаемся ограничить оказание специальных информационных услуг (канал 3) ключевыми пользователями. По причине возросшего риска увязки необходимо будет снизить уровень детализации данных, предоставляемых в ОФЕЗ. Тем не менее они будут оставаться основным инструментом предоставления доступа исследователям к микроданным.

31. Ключевыми областями совершенствования будут являться стандартные статистические материалы (канал 1), ЛДДД (канал 5) и сотрудничество (канал 7). Мы надеемся опубликовать большее число кубов данных (канал 2), однако на практике эта форма распространения данных пригодна только в отношении ограниченного набора статистических рядов.

32. Наша цель в области стандартных статистических материалов будет заключаться в увеличении количества данных, публикуемых в такой форме через наши специальные вебсервисы (например, AUSSTATS). Всем статистическим подразделениям будет предложено произвести пересмотр своих стратегий распространения с целью сокращения масштабов использования бумажных публикаций и расширения практики предоставления материалов в электронном режиме.

33. ЛДДД является новым сервисом, который был введен в действие незадолго до начала сессии КЕС. Мы планируем продолжить разработки в этой области на основе опыта использования первой версии. Они будут направлены на решение следующих задач:

- предоставление доступа к более подробным наборам микроданных; и

- предоставление доступа к связанным наборам данных, в особенности в тех случаях, когда один из наборов данных ведется сторонней организацией.

34. В настоящее время мы более активно занимаемся развитием варианта "Сотрудничество", поскольку мы располагаем полностью работоспособным и имеющим высокую репутацию Отделом анализа. Мы попытаемся начать сотрудничество в тех областях, которые представляют для нас наибольший интерес, в частности когда существует потенциал разработки новых статистических материалов. Особый интерес представляет повышение полезности существующих наборов данных за счет применения к ним аналитических методов. На практике некоторые исследователи будут обращаться к нам в первую очередь. Мы изучим вопрос о том, какие взаимные выгоды могут принести нам соглашения о сотрудничестве. Канал распространения 8, возможно, будет использоваться в целях некоторых совместных проектов, однако мы будем использовать его выборочно.

VI. ОРГАНИЗАЦИОННЫЕ МЕХАНИЗМЫ

35. Инициатива разработки таких механизмов должна исходить от руководства СБА, особенно в период, когда такие механизмы претерпевают значительные изменения. Важное значение имеет связь как с внутренними, так и внешними пользователями. Мы ощущаем поддержку со стороны Отдела СБА, отвечающего за политику и координацию.

36. Фактическим ведением и администрированием этих механизмов занимается наш Отдел информационного обслуживания. В этом Отделе существует подразделение, отвечающее за администрирование и распространение материалов ОФЕЗ, ЛДДД и Лаборатории данных СБА. Отдел также отвечает за пропаганду этих услуг и связь с пользователями.

37. Во многих отношениях эти каналы доступа по-прежнему находятся в стадии разработки. В целях укрепления нашего исследовательского потенциала, в том числе за счет исследований, проводимых сторонними организациями, и придания нашей работе большей целенаправленности мы создали в Отделе методологии группу по методам доступа к данным и обеспечению конфиденциальности. Ее возглавляет старший методист.

38. Для поддержки развития ЛДДД была создана специальная проектная группа (деятельность которой курирует Совет по проектам). Текущие функции по ведению этих систем были переданы Отделу информационного обслуживания.

39. Отдел анализа отвечает за разработку и осуществление большинства соглашений о сотрудничестве, которые предусматривают доступ к микроданным. Некоторые из них могут осуществляться тематическими подразделениями, однако это является исключением. Мы используем Отдел анализа для обеспечения более согласованного подхода. Кроме того, они обладают техническими знаниями для наиболее эффективной работы с исследователями. Надзор за этой работой осуществляет Совет по проектам, в состав которого входят ведущие специалисты в различных отраслях статистики.

40. И наконец, необходимо обеспечить тесное сотрудничество со всеми отраслевыми отделами. Отраслевые отделы отвечают за предоставление базовых наборов данных для всех каналов распространения. Кроме того, время от времени исследователи будут обращаться к ним за специальными консультациями.

VII. КЛЮЧЕВЫЕ ВОПРОСЫ

41. Предоставление по-настоящему "безопасных данных" становится все более трудной задачей, в связи с чем становится неизбежным более широкое использование "безопасных схем", включая юридические механизмы, для вторичного анализа данных. Это является более трудоемкой задачей, требующей дополнительных ассигнований в период, когда НСУ приходится во многих случаях работать в условиях дефицита ресурсов. Однако мы считаем целесообразным такое перераспределение ресурсов, если оно позволит эффективно использовать наши данные.

42. Одной из возможных проблем может стать отношение исследователей к этим вариантам. С точки зрения исследователей, они могут создать излишние ограничения или неудобства. Они спрашивают, с чем связано наше недоверие к ним по поводу того, что они могут сделать что-то не так, как надо. С учетом этого особое значение приобретает стратегия связей с пользователями. Мы должны информировать исследователей не только об этих новых вариантах, но также и о причинах, обуславливающих необходимость ограничений. Есть больше шансов того, что они будут работать в системе, логику которой они понимают.

43. Мы на практике осуществляем переход от парадигмы избежания рисков к парадигме управления рисками. Существует значительный риск утраты доверия общественности к обеспечиваемому нами уровню защиты конфиденциальности предоставляемых ею данных. Данный риск может быть незначительным и оправданным с учетом повышения полезности наших статистических данных, но он тем не менее существует. Система ценностей исследователей отлична от системы ценностей официальных статистиков. В ней преобладает исследовательский императив, в связи с чем исследователи могут быть

разочарованы тем, что может показаться им избыточными препятствиями или проявлениями бюрократизма. Некоторые неизбежно попытаются перейти линию дозволенного. Мало вероятно, что исследователи будут пытаться идентифицировать какое-то отдельное лицо, поскольку для этого у них отсутствуют мотивы. Исходя из нашего опыта, скорее можно предположить, что они попытаются обойти правила с целью обеспечения прогресса в осуществлении своих исследовательских программ (например, мы имели случаи, когда наши микроданные предлагались для дальнейших исследований, хотя и добавленной стоимостью). Важно, чтобы мы не оставляли без внимания такие случаи. В некоторых случаях уместным может быть применение юридических санкций. Данный процесс может быть трудным и длительным. Лишение права на обслуживание, в том числе и всего учреждения, по вине одного из его исследователей, может быть легко применимой и весьма эффективной мерой, в особенности если среди исследователей распространится информация о том, что СБА готово применять такую меру.

44. И наконец, исследователи активно участвуют в международном сотрудничестве. Они могут сообщать друг другу о том, что они могут делать в стране А по сравнению со страной В. Нам известно об этом из личного опыта. Более тесное согласование наших подходов несет в себе значительные выгоды. Если мы пришли к согласию по Основным принципам официальной статистики, то почему мы не могли бы сделать то же самое в отношении основных принципов использования микроданных внешними исследователями? Этому вопросу посвящается нижеследующий раздел.

VIII. ВЫВОДЫ

45. Предоставление внешним исследователям наших статистических данных является одним из важных способов повышения полезности нашей статистической деятельности. Мы рассматриваем данное сотрудничество в качестве одной из важных задач СБА. Кроме того, наше законодательство разрешает нам осуществлять такого рода деятельность.

46. В прошлом при толковании данного законодательства мы придерживались более консервативного подхода, главное место в котором уделялось обеспечению безопасности данных. Непрерывное развитие технологии и наличие внешних баз данных затрудняют публикацию по-настоящему "безопасных" микроданных (или безопасных кубов данных в этой области). Все более широкое распространение баз данных частного сектора, возможно, является наиболее острой проблемой, поскольку вопросы их использования и злоупотреблений в этой области не являются объектом жесткого регулирования.

47. Следовательно необходимо использовать методы распространения, опирающиеся на "безопасную среду". В этом отношении благоприятным для нас известием стало подтверждение того, что юридические обязательства, подписанные исследователями, могут приниматься во внимание при оценке степени соблюдения нами определяющего наши права законодательства. Это означает, что мы не должны опираться только на безопасные данные.

48. Мы продолжим придерживаться несколько консервативного подхода при интерпретации наших установленных законом полномочий в области публикации микроданных, хотя и не столь консервативного, как ранее. Это объясняется тем, что, по нашему мнению, один крупный инцидент может нанести значительный ущерб нашей репутации и нашей способности поддерживать доверие общественности к уровню обеспечиваемой нами защиты конфиденциальности предоставляемых ею данных. Это отрицательно скажется на коэффициентах предоставления ответов в ходе сбора данных и качестве производимой нами статистики.

49. Наиболее многообещающим из новых подходов к представлению доступа к микроданным является ЛДДД. Особый интерес представляет ее использование в качестве средства расширения доступа к связанным наборам данных.

50. Как и в случае многих других статистических начинаний, в данной области мы можем многому научиться друг у друга. Речь идет как о положительном, так и неудачном опыте. Доступ к микроданным является сферой деятельности, в которой мы, возможно, пожелаем согласовать некоторые базовые принципы. Исследователи работают в различных странах и проводят сопоставления. На практике уже существуют механизмы (например, Люксембургское исследование доходов), в рамках которых микроданные из нескольких стран сводятся воедино для обеспечения удобного доступа к ним исследователям.

51. Естественно, что юридические и административные механизмы будут являться различными в зависимости от страны. Однако это не мешает согласованию определенных базовых принципов. Для начала я предлагаю обсудить следующие вопросы:

- Микроданные, собираемые в целях официальной статистики, должны использоваться в исследовательских целях и в целях вторичного анализа при условии соблюдения установленных требований, препятствующих злоупотреблениям.

- Использование микроданных вне стен НСУ, за исключением исследовательских и статистических целей, не допускается.
- Необходимо разработать юридические или другие механизмы использования микроданных с целью повышения уверенности общественности в их надлежащем использовании.
- Виды использования микроданных должны быть прозрачными, а информация о них должна широко распространяться с целью повышения уверенности общественности в том, что микроданные используются надлежащим образом.
- НСУ могут производить наем внешних исследователей в качестве сотрудников только в том случае, если это будет содействовать работе, результатом которой станет разработка новых статистических материалов.
- Механизмы предоставления доступа к микроданным должны быть одобрены национальными органами по вопросам защиты конфиденциальности.
