



**ЭКОНОМИЧЕСКИЙ
И СОЦИАЛЬНЫЙ СОВЕТ**

Distr.
GENERAL

ECE/CES/GE.41/2009/9
19 August 2009

RUSSIAN
Original: ENGLISH

ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ

Группа экспертов по переписям населения и жилищного фонда

Двенадцатое совещание

Женева, 28-30 октября 2009 года

Пункт 4 предварительной повестки дня

КОНТРОЛЬ КАЧЕСТВА ПЕРЕПИСИ И ПРЕДОТВРАЩЕНИЕ ИДЕНТИФИКАЦИИ

**Вопросы конфиденциальности при проведении переписи населения
и жилищного фонда в 2011 году**

Записка Статистического управления Европейских сообществ и Статистического
управления Нидерландов

I. ВВЕДЕНИЕ

1. В 2008 году Совет Европейского союза и Европейский парламент во исполнение рекомендаций Конференции европейских статистиков по проведению переписей населения и жилищного фонда 2010 года приняли Постановление (ЕС) № 763/2008¹

¹ Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses (Text with EEA relevance). OJ L 218, 13.8.2008, p.14.

относительно переписей населения и жилищного фонда. В Постановлении говорится, что "государства-члены принимают все меры, необходимые для выполнения требований о защите данных. Настоящее постановление не распространяется на национальные положения государств-членов по защите данных". Передача конфиденциальной статистической информации регулируется отдельными постановлениями ЕС, обеспечивающими физическую и логическую защиту конфиденциальной информации и не допускающими незаконного разглашения или использования не в статистических целях статистических данных, собираемых и распространяемых Сообществом. В частности, в новом "Статистическом законе ЕС"² вопросам конфиденциальности статистической информации посвящена целая глава. Иными словами, данные, считающиеся конфиденциальными на национальном уровне, остаются конфиденциальными и после передачи в Евростат, и если страна хочет передать конфиденциальные данные, она должна делать это в соответствии с действующими постановлениями ЕС.

2. В основе разработанной Евростатом программы распространения данных раунда переписей 2011 года лежит новаторский подход. Исходные данные вводятся в форме гиперкубов – многомерных таблиц, имеющих несколько измерений. Размер этих гиперкубов важен с точки зрения конфиденциальности: в отличие от двух- или трехмерных таблиц общей заданной формы, которые позволяют (сравнительно легко) предотвращать идентификацию на основе данных переписей, в случае многомерных таблиц такой контроль становится реальной проблемой.

3. Для сведения к минимуму риска идентификации используются гиперкубы как можно меньших размеров. Однако даже при этом вероятность того, что потребуются прибегнуть к тем или иным методам защиты конфиденциальности информации, остается очень высокой. Если страны будут использовать разные методы, может пострадать сопоставимость данных. Кроме того, пользователям проще понять недостатки применения одного метода всеми странами, чем влияние разных национальных методов на конфиденциальность данных. Поэтому было сочтено целесообразным изучить возможность использования единого подхода к предотвращению идентификации на основе данных переписей в масштабе всего ЕС.

² Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics (Text with relevance for the EEA and for Switzerland). OJ L 87, 31.3.2009, p. 164.

4. Для анализа современных методов предотвращения идентификации на основе данных переписей в Европе в мае 2008 года был проведен небольшой опрос.

В разосланных 45 странам вопросниках задавались следующие вопросы:

- a) Существует ли в стране конкретная норма о конфиденциальности данных переписей?
- b) Как обеспечивается выполнение национальных требований к защите конфиденциальности информации?
- c) Какое учреждение отвечает за определение/применение методики предотвращения идентификации?
- d) Существует ли методология обеспечения конфиденциальности на основе данных переписей 2011 года?
- e) Находят ли поддержку усилия по выработке единой методологии на уровне ЕС?

5. С полным текстом ответов, полученных от 32 стран, можно ознакомиться в рубрике "Population" на сайте CIRCA³.

6. Обычно не существует каких-либо особых положений относительно конфиденциальности данных переписей, так как на них распространяются те же ограничения, что и на все другие статистические данные. Если такие положения упоминаются в законе, то речь, как правило, идет об общенациональных нормах по данному вопросу. Ответственность за практическое осуществление этих общих положений обычно возлагается на национальные статистические управления. Однако немногие из них разработали методологию предупреждения идентификации на основе данных следующей переписи. Вот почему решение Евростата проанализировать возможность использования единой методологии предупреждения идентификации встретило горячую поддержку.

³ <http://circa.europa.eu/Public/irc/dsis/population> (доступ только для зарегистрированных членов).

II. ЗАДАЧИ, СВЯЗАННЫЕ С РАЗРАБОТКОЙ УНИФИЦИРОВАННОГО ПОДХОДА К ПРЕДОТВРАЩЕНИЮ ИДЕНТИФИКАЦИИ НА УРОВНЕ ЕС

7. Исходя из результатов опроса, Евростат создал целевую группу по предупреждению идентификации на основе данных переписей (ЦГПИ). В нее вошли эксперты по вопросам предотвращения идентификации из Германии, Эстонии, Италии, Нидерландов, Португалии и Соединенного Королевства. Ее конкретная цель заключается в выявлении и устранении трудностей, касающихся защиты конфиденциальности данных переписей населения и жилищного фонда, а также в принятии или выработке унифицированной методологии на основе учета национальных требований. ЦГПИ провела два совещания и должна представить результаты своей деятельности Рабочей группе по демографии и переписям населения в первой половине 2010 года.

8. ЦГПИ придется решить ряд концептуальных задач. Во-первых, сравнительно большое число измерений гиперкубов осложняет применение стандартного метода предупреждения идентификации. Во-вторых, с учетом практикуемого Европейским центром переписей подхода необходимо принять решение о том, на каком уровне должна решаться эта задача – на уровне национальных баз данных или в процессе обработки. В-третьих, следует обеспечить согласованность результирующих данных таблиц между гиперкубами и между извлекаемыми данными. В-четвертых, поскольку опыт и инструментарий предупреждения идентификации, имеющийся в распоряжении государств-членов, может различаться, единый метод должен быть как можно более простым в использовании. В-пятых, этот метод должен быть легко понятен обычному пользователю. В-шестых, поскольку одним из плюсов переписи является получение подробной информации, следует свести к минимуму потерю данных. Наконец, каждая страна имеет свои собственные нормы защиты конфиденциальности информации, которые должны соблюдаться.

9. Для следующего раунда переписи намечена амбициозная программа распространения собранных данных. Благодаря единому интерфейсу – центральной информационной системе Евростата – пользователи из государств - членов ЕС (и других стран, желающих участвовать в программе распространения данных переписей) получают по сравнению с предыдущими раундами переписи более широкий доступ к информации. Эта система распределяет запросы пользователей между национальными базами данных, собирает и представляет систематизированную информацию. С точки зрения конфиденциальности обратная сторона медали заключается в том, что, с учетом большого числа измерений и имеющихся у пользователей широких возможностей для составления интересующих их таблиц (включая возможность направлять повторные запросы), следует

очень внимательно оценивать риски использования стандартных методов предупреждения идентификации.

10. Один из основных вопросов заключается в том, следует ли проверять на предмет конфиденциальности используемые национальные данные заранее или же в процессе обработки, т.е. непосредственно перед их представлением пользователям. Последний вариант может быть оправдан тем, что число измерений, доступных пользователям, гораздо меньше общего числа измерений соответствующего гиперкуба, что снижает риск идентификации. С другой стороны, такой подход означает, что конфиденциальные данные так или иначе поступают из национальных баз данных (со всеми вытекающими отсюда последствиями для защиты ИТ), что повышает риски, связанные с множественными запросами со стороны потенциально недобросовестных пользователей. Данные могут проверяться на предмет конфиденциальности на трех уровнях – на уровне микроданных, гиперкубов или извлеченных данных.

11. Независимо от того, на каком этапе происходит предупреждение идентификации, считается важным, чтобы представляемые пользователям результаты согласовывались друг с другом на уровне как отдельных таблиц, так и извлеченной информации. Широкое использование данных переписей и тот факт, что к ним прибегают в течение длительных периодов времени, являются достаточными основаниями для того, чтобы данные, получаемые пользователями, не различались в зависимости от момента извлечения или используемого в качестве источника гиперкуба. Хотя некоторые методы обеспечения конфиденциальности данных могут быть весьма эффективными, необходимо удостовериться в том, что они не оказывают отрицательного влияния на сопоставимость данных.

12. Методы и инструменты предупреждения идентификации стали довольно сложными. С учетом этого был разработан ряд европейских проектов, таких, как проект CFSC (2000-2003 год, см. <http://neon.vb.cbs.nl/casc/>), проект CENEX (2006 год) и проект ESS (2008-2009 годы), а ЕЭК ООН и Евростат регулярно организуют совещания, посвященные вопросам конфиденциальности статистических данных. В то же время не следует считать, что все статистические управления обладают одинаковыми возможностями. Если предупреждение идентификации на основе статистических данных будет вестись на национальном уровне на основе единой методики, в процессе унификации методов защиты конфиденциальности данных должны учитываться дополнительные требования, такие, как простота использования и доступность необходимого инструментария (желательно без чрезмерных расходов). В периоды дефицита ресурсов от стран нельзя требовать дополнительных крупных расходов.

13. Помимо вышеуказанных трудностей, связанных с выработкой единой методики, следует учитывать и другие факторы вне зависимости от того, будет ли методика применяться во всех странах одинаково или с учетом национальной специфики. При использовании методов предупреждения идентификации пользователь должен информироваться об их характеристиках и влиянии на данные. Чем проще метод, тем легче обычному пользователю понять его последствия (и, по всей видимости, тем проще передавать эту информацию). В интересах прозрачности и общего качества данных этот фактор не следует игнорировать.

14. Одно из важнейших достоинств переписи заключается в том, что она позволяет получить сведения на всех географических уровнях и/или о небольших группах лиц и иногда является единственным источником информации. Этим следует в максимальной степени дорожить, признавая в то же время необходимость обеспечения конфиденциальности данных. Хотя широкое применение жестких методов предупреждения статистической идентификации и помогает сохранить их конфиденциальность, оно также способно значительно ограничить доступность информации. Сведение к минимуму последствий применения того или иного метода предупреждения идентификации для доступности данных отвечает интересам пользователей. Инструменты защиты конфиденциальности информации должны применяться разумно с учетом неизбежной потери части (подробных) сведений.

15. Поиск удовлетворительного решения для каждой из вышеперечисленных проблем вполне может стать безнадежной задачей, которую учет национальных барьеров будет лишь осложнять. Однако, несмотря на многочисленность национальных требований к защите конфиденциальности информации, поиск общего решения все же имеет смысл, поскольку в национальных нормах зачастую устанавливаются лишь общие принципы, а многие практические вопросы (и связанный с ними методологический выбор) оставляются на усмотрение национальных статистических управлений. Если эти статистические ведомства согласуют единый подход, он не будет идти вразрез с национальными требованиями, поскольку именно они решают технические задачи предупреждения идентификации на основе национальных данных. Широкую поддержку совместной деятельности по защите конфиденциальности данных переписей на уровне ЕС со стороны статистических управлений можно рассматривать как признание необходимости налаживания обмена опытом и/или оказания помощи по техническим вопросам. Важную роль в этих вопросах, безусловно, может сыграть ЦГПИ.

III. ТЕХНИЧЕСКИЕ АСПЕКТЫ МЕТОДОЛОГИИ ЕС

16. Постановление (ЕС) № 763/2008 ориентировано на итоговые материалы, т.е. допускает использование разных источников данных, требуя при этом уважать основные принципы переписей населения и жилищного фонда и использовать унифицированные определения, технические спецификации, темы и уровни детализации. Это постановление предусматривает единый график переписей (начиная с 2011 года), единую программу распространения данных на уровне ЕС, единые технические стандарты передачи данных и контроль качества для европейских потребностей. Что касается конфиденциальности статистических данных, то особое значение имеют следующие аспекты:

- a) Пункт 2 статьи 4 предусматривает, что "государства-члены принимают все меры, необходимые для выполнения требований о защите данных. Настоящее постановление не распространяется на национальные нормы государств-членов по защите данных". Это означает, что защита конфиденциальности данных относится к ведению государств-членов и должна осуществляться на их уровне, а не на уровне Комиссии. В пункте 2 статьи 4 говорится далее, что Европейская комиссия не уполномочена издавать нормативные акты, касающиеся предупреждения идентификации данных на основе постановления относительно проведения переписей. В то же время в пункте 4 статьи 6 говорится, что "Комиссия (Евростат) в сотрудничестве с компетентными органами государств-членов готовит методологические рекомендации, призванные обеспечить качество представляемых данных и метаданных, исходя, в частности, из Рекомендаций Конференции европейских статистиков по проведению переписей населения и жилищного фонда 2010 года". В пункте 3 вступительной части постановления поясняется, что "в свете методологических и технических изменений следует выявлять примеры передовой практики и стремиться к совершенствованию источников данных и методологий, используемых при проведении переписей в государствах-членах".
- b) В пункте 2 статьи 5 постановления относительно переписей предусматривается, что "государства-члены представляют в Комиссию (Евростат) окончательные, выверенные и агрегированные данные (...)". Это не относится к передаче микроданных. Хотя агрегированные данные необязательно защищаются на предмет разглашения чувствительной информации, в духе пункта 2 статьи 5 конфиденциальные данные не должны передаваться Евростату.

- c) В пунктах 5 и 7 вступительной части говорится, что Закон о статистике и, соответственно, Европейский кодекс статистической практики представляют собой рамочную основу постановления относительно переписей и в обоих из них содержатся положения, касающиеся конфиденциальности статистической информации.
- d) В пункте 6 вступительной части напоминаются требования постановлений, регулирующих передачу конфиденциальной статистической информации. В них говорится, что, если государства-члены передают статистические данные, которые они считают конфиденциальными, Евростат должен обеспечить их физическую и логическую защиту и не допустить незаконного разглашения или использования не в статистических целях статистических данных, собираемых и распространяемых Сообществом. Однако постановление относительно переписей не предусматривает передачи конфиденциальных данных переписей государствами-членами Евростату. Пункт 6 вступительной части в широком смысле напоминает, что необходимо делать все возможное для предупреждения непреднамеренного разглашения конфиденциальной информации.

17. В принципе ЦГПИ руководствуется в своей деятельности двумя основными рекомендациями:

- a) Рекомендацией о защите от шумов на этапе до составления таблиц на уровне микроданных. Это, по-видимому, имеет смысл в контексте распространения результатов переписи 2011 года как на национальном, так и на общеевропейском уровне. В то же время такая защита может быть обеспечена лишь НСИ (национальными статистическими институтами), и у Евростата отсутствуют возможности хотя бы проверить, обеспечивалась такая защита или нет.
- b) Рекомендация о защите на этапе после составления таблиц (на уровне гиперкубов). На сегодняшний день эта работа делится на "отсев элементов данных" и "защиту от шумов на этапе после составления таблиц данных". Простым выходом из положения стал бы отбор тех элементов данных, которые не могут быть опубликованы (первичный отсев), дополняемый защитой ряда элементов, с тем чтобы не прибегать к пересчету на основе пограничных значений (так называемый вторичный отсев). В то же время целевая группа могла бы рассмотреть возможность получения синергического эффекта от увязывания этих двух методологий, поскольку цель заключается в

предупреждении идентификации субъектов, т.е. недопущения определенности в отношении значений в ячейках частотных таблиц. Такая мера должна в идеале осуществляться с минимальной потерей данных.

18. Поскольку реальные данные переписей 2011 года пока отсутствуют, ЦГПИ использует тестовые гиперкубы по некоторым странам. В двух нижеследующих примерах изображены двухмерные таблицы многомерных пробных итальянских гиперкубов⁴. По очевидным причинам названия переменных были заменены на обозначения типа var2, var3 и т.д. В первом примере (таблица 1) некоторые элементы данных требуют защиты, однако проблемы конфиденциальности выглядят вполне решаемыми. Даже если данные регистрации будут отсутствовать в большем числе ячеек, использование правильной стратегии защиты, по всей видимости, позволит иметь большое число ячеек без реальных показателей частотности.

19. Во втором примере (таблица 2) даже все внутренние ячейки имеют нулевые значения или должны быть отброшены. Очевидно, что с этой проблемой чаще других будут сталкиваться страны, использующие результаты обследований (вместо данных по всей совокупности или данных регистров) для расчета некоторых переменных переписи 2011 года.

20. Какой урок позволяют нам извлечь эти таблицы? Все гиперкубы должны проверяться на предмет возможности опубликования большинства элементов данных. Не имеет смысла составлять и публиковать гиперкубы в основном или исключительно с нулевыми или отбрасываемыми значениями. Хотя страны могут использовать разные правила защиты конфиденциальности данных, эта проблема актуальна для всех европейских стран. Крупные страны имеют больше территорий и поэтому сталкиваются с теми же проблемами на более детализированном уровне, что и малые страны.

⁴ Эти таблицы были составлены Федеральным статистическим управлением Германии.

Таблица 1

Table: Region x var1 x var2 x var3 x var6 | frequenza

var2	var3	var6																
tot	tot	tot	tot	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
tot	46.191	2.292	2.064	1.655	1.821	4.785	7.895	7.234	5.237	3.160	2.089	1.622	1.297	1.142	880	723	7	
8101	114	6	6	3	3	9	14	26	20	5	3	4	4	3	4	2		
8102	213	9	19	12	11	24	43	23	31	9	4	2	3	3	7	1		
8103	86	1	4	2	1	11	20	20	11	4	3	1	2	2	1	2		
8104	228	5	9	10	21	25	30	30	25	9	11	3	6	7	6	4		
8105	131	9	7	5	7	18	24	17	13	5	7	4	3	5	1	4		
8106	402	12	17	9	9	49	65	55	37	38	22	13	18	20	4	4		
8107	86	2	3	4	6	15	14	10	9	6	2	5	1	1	3	-		
8108	88	-	1	3	3	17	14	12	11	9	4	5	2	4	1	1		
8109	72	6	2	2	3	8	17	12	8	3	2	1	-	-	4	1		
8110	47	3	1	3	4	7	5	7	10	4	-	1	-	-	1	-		
8111	238	13	12	10	12	15	38	23	33	20	14	12	12	9	3	2		
8112	601	33	25	29	25	71	107	97	56	39	29	18	19	7	9	6		
8113	51	3	2	1	1	8	8	10	3	1	2	4	4	-	1	-		
8114	160	1	5	5	8	15	24	22	12	13	8	7	5	4	6	4		
8115	1.100	75	45	50	38	102	202	194	146	77	49	30	20	12	20	6		
8116	402	19	23	24	23	34	87	56	37	24	20	18	7	6	7	1		
8117	394	24	19	20	14	43	72	61	43	26	11	6	12	11	10	8		
8118	401	22	16	18	16	41	58	63	47	26	14	15	11	11	11	7		
8119	237	11	8	11	12	32	46	36	21	14	15	5	8	6	2	2		
8120	89	2	6	1	3	10	12	12	10	1	4	2	3	2	4	-		
8121	739	40	41	27	39	76	127	109	76	47	39	31	25	20	8	7		
8122	555	38	32	20	18	51	91	92	73	38	13	22	14	19	12	6		
8123	243	14	8	13	16	24	38	43	24	11	10	5	4	2	3			
8124	32	-	-	-	1	3	8	4	2	2	-	1	1	-	-	1		
8125	731	28	27	22	34	85	131	108	65	40	29	31	20	14	12	18		
8126	151	7	4	8	2	18	30	21	16	9	6	13	2	6	3	1		
8127	405	22	23	13	22	46	69	54	58	29	16	6	9	5	12	6		
8128	431	26	15	25	25	47	81	60	37	38	27	9	16	5	8	6		
8129	28	1	2	-	1	7	6	4	4	-	-	-	-	-	2	-		
8130	186	8	9	5	6	20	34	34	21	13	12	4	4	5	4	2		
8131	164	13	10	6	3	10	33	28	17	10	5	4	9	3	2	3		
8132	54	3	2	4	3	4	9	5	10	3	3	-	-	3	1	-		
8133	31	1	1	1	1	8	3	4	-	4	2	4	-	1	-	-		
8134	18	-	-	-	-	2	-	2	2	-	3	-	1	2	3	3		
8135	230	15	6	11	10	19	35	37	30	10	13	9	7	11	4	3		

3 dig. separator Select Table Change View Write table
Output View Table Summary Close

Cell Information

Value: 0

Status: Empty

Cost: 0

Shadow: 0

contributions: 0

Top n of shadow:

Holding level: ☐

Request: 0

Change status

Set to Safe

Set to Unsafe

Set to Protected

A priori info

All Non-StructEmpty

Recode

Suppress

HyperCube

Modular

Network

Optimal

Marginal

Rounding

InverseWgt

Singleton

Undo Singleton

Suppress

Undo Suppress

Audit

Таблица 2

Table: Region x var1 x var2 x var3 x var6 | frequenza

var2	var3	var6																				
9	tot	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
tot	218	4	5	3	6	8	12	26	34	20	26	11	12	8	15	10	5	6	5	2	-	-
8101	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8102	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8103	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8104	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8105	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8106	1	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
8107	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8108	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8109	2	-	-	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8110	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8111	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8112	3	-	-	-	-	-	-	2	-	1	-	-	-	-	-	-	-	-	-	-	-	-
8113	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8114	1	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
8115	9	-	1	-	2	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-
8116	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8117	1	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
8118	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8119	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8120	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8121	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8122	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8123	2	-	-	-	-	-	-	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
8124	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8125	1	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
8126	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8127	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8128	3	-	-	-	-	-	-	-	-	-	1	-	1	1	-	-	-	-	-	-	-	-
8129	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8130	4	-	-	-	-	1	-	-	1	-	-	-	-	1	-	1	-	-	-	-	-	-
8131	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	1	-	-	-	-
8132	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8133	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8134	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8135	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-
8136	3	-	-	-	-	1	-	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-

3 dig. separator Select Table Change View Write table
Output View Table Summary Close

Cell Information

Value: 218

Status: Safe

Cost: 218

Shadow: 218

contributions: 218

Top n of shadow:

Holding level: ☐

Request: 0

Change status

Set to Safe

Set to Unsafe

Set to Protected

A priori info

All Non-StructEmpty

Recode

Suppress

HyperCube

Modular

Network

Optimal

Marginal

Rounding

InverseWgt

Singleton

Undo Singleton

Suppress

Undo Suppress

Audit

21. Целевая группа сопоставит различные правила защиты конфиденциальности информации, которые применяют страны, и проанализирует их влияние с точки зрения первичного отсева данных. Для того чтобы не прибегать к пересчету первоначально отброшенных данных на основе общих пограничных значений, необходимо отбросить ряд дополнительных элементов данных. Такой дополнительный отброс именуется вторичным отсевом. Общеизвестным оптимальным методом вторичного отсева является так называемый метод гиперкубов Tau-ARGUS, который применяется многими странами. Программное обеспечение Tau-ARGUS можно загрузить бесплатно с сайта <http://neon.vb.cbs.nl/casc/>. На этом сайте можно также найти и скачать пробные данные и руководство по использованию программного обеспечения.

IV. НЕКОТОРЫЕ ПРЕДВАРИТЕЛЬНЫЕ ВЫВОДЫ

22. С учетом итогов работы первых двух совещаний ЦГПИ, а также графика проведения переписей в настоящее время предпочтительным считается использование упрощенной методики. Если государства-члены не придут к явному и полному согласию по вопросу предлагаемой методологии, в постановление ЕС, регулирующее передачу статистических данных Евростату, не будет включено никаких положений о методе предупреждения идентификации данных на основе переписи. Это означает, что по соображениям конфиденциальности страны смогут оставлять ячейки гиперкубов пустыми.

23. Однако следует помнить, что государства-члены могут считать конфиденциальными не все темы. Например, следует подумать над тем, подлежат ли защите от идентификации в соответствии с национальными нормами такие признаки, как "пол" и "возраст". Вполне может оказаться, что одни признаки являются более "чувствительными", чем другие. Поскольку признаки, перечисленные в постановлении ЕС, т.е. обязательные для всех государств-членов, по сути являются теми же, что и основные признаки КЕС, которые не затрагивают характеристик, касающихся, например, сферы здравоохранения или доходов, возможно, следовало бы еще на национальном уровне и с учетом национальных требований задуматься над тем, имеют ли соображения конфиденциальности отношение ко всем признакам и всем единицам наблюдения⁵.

24. Максимально учитывая вышеупомянутые ограничения, ЦГПИ будет продолжать разработку унифицированного подхода, который будет рекомендован для использования всеми странами. Ответ на вопрос о том, поможет ли это повышению сопоставимости

⁵ Если вопросы конфиденциальности, безусловно, актуальны для физических лиц, они не столь очевидны применительно к другим единицам наблюдения, таким как домашние хозяйства, жилые помещения и т.д.

данных переписей, будет зависеть от того, с какой степенью гибкости национальные статистические органы будут относиться к единой для всех методике, которую предложит целевая группа. В любом случае, ее работа поможет обсуждению данной темы.
