

Distr.
GENERAL

CES/SEM.40/4
8 September 1998

RUSSIAN
Original: ENGLISH

СТАТИСТИЧЕСКАЯ КОМИССИЯ и
ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ

СТАТИСТИЧЕСКОЕ УПРАВЛЕНИЕ
ЕВРОПЕЙСКИХ СООБЩЕСТВ (ЕВРОСТАТ)

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ

Совместная рабочая сессия ЕЭК-Евростат
по переписям населения и жилищ 1/
(Дублин, Ирландия, 9-11 ноября 1998 года)

Тема для исследования 1

ПРИМЕНЕНИЕ ГЕОГРАФИИ ПРИ ПЕРЕПИСЯХ: ОПЫТ ФРАНЦИИ

Документ, представленный НИСЭИ, Франция 2/, в соответствии
с высказанным предложением

География всегда играла ключевую роль в проведении переписей и других процессах, связанных со сбором данных. В наши дни эта роль меняется вследствие существования ГИС.

В настоящем документе представлены различные формы использования географии, зависящие от методик сбора информации при переписях, и, в частности, охарактеризован французский опыт. В нем также показаны новые перспективы, которые открываются благодаря новым географическим методам, применяемым во Франции.

Документ будет служить основой для выработки ряда рекомендаций, которые должны быть обсуждены на дублинской конференции, организуемой ЕЭК и Евростатом.

I. Методика сбора информации при переписях и география

1. Используемые в настоящее время в различных странах методики сбора информации, как и существующие географические базы данных, весьма разнообразны. Это обуславливает потребность в различных географических инструментах. В настоящем документе рассматривается применение географии для разнообразных целей, относящихся к сбору данных, в том числе организация работы с почтовыми адресами в тех случаях, когда этого требует методика сбора. Для примера выбрано несколько стран, хотя можно было бы рассмотреть и многие другие примеры.

1/ Документы, подготовленные для этой рабочей сессии, будут обрабатываться таким же образом, как и документы, подготавливаемые для семинаров.

2/ Подготовлен Мишелем Жако.

2. При подготовке настоящего документа не ставилась цель акцентировать внимание на географических аспектах распространения данных. Методики распространения данных во многих отношениях не зависят от методик их сбора и сильно различаются в разных странах. Распространение данных связано со следующими аспектами: географическая привязка статистических результатов; подготовка тематических карт; картографирование связей между территориальными единицами; распространение информации через систему Интернет; электронные носители информации и бумажная документация; графическая семиология.

Тем не менее географические аспекты проблем распространения не следует отделять от проблем сбора по причине затрат и качества.

3. Всего лишь несколько примеров достаточно для того, чтобы показать, что в его географических аспектах вопрос о разнообразии методологии сбора информации при переписях не может быть должным образом освещен, если ориентироваться на проведение различия между двумя обычными категориями:

- переписями, проводимыми на основе регистров;
- традиционными переписями.

4. Действительно, накопленный в Дании и Финляндии опыт получения информации "переписного типа" во многочисленных взаимосвязанных административных источниках свидетельствует о необходимости широкого использования географии при ведении файлов данных о жилищных единицах с географической привязкой. Данные о месте работы или месте учебы также берутся из файлов данных с географической привязкой.

5. Бельгийская методика является примером смешанной стратегии, предусматривающей использование регистра и проведение обследований для проверки и дополнения существующей информации. Она также предусматривает использование почтовых адресов для рассылки форм по почте.

6. Что касается традиционных переписей, то методика "мейл-ин", применяемая, например, в Бюро переписей США, требует проведения до дня переписи большой предварительной работы по составлению списка адресов жилищных единиц и их географической привязке. Концептуально она не очень сильно отличается от бельгийской, если не считать того, что эта работа носит единовременный характер. Если Бюро переписей США получит разрешение и начнет вести этот "генеральный файл адресов" на долгосрочной основе, концептуальное различие в географических аспектах практически исчезнет, хотя процедура по-прежнему останется иной.

7. Методика "мейл-аут" (Канада) внешне существенно отличается. При ней также должен составляться список адресов, в частности для определения доли ответивших, но это может быть сделано и после возвращения переписных форм. При этом не ставится цель повторно использовать этот список в следующий раз. Правда, работа по определению границ переписных участков и подготовке местных карт и их выдаче счетчикам на всей территории страны должна быть завершена до дня переписи.

II. Некоторые географические аспекты опыта подготовки к переписи 1999 года во Франции

8. В этой главе после короткого обзора различных географических проблем, обусловленных разнообразием методик сбора информации, будут более глубоко рассмотрены конкретные формы географической работы и приведены примеры текущей подготовки к проведению переписи 1999 года во Франции. Будет исследовано несколько уровней – от коммунального и до характеризуемого наибольшей степенью детализации – места нахождения строения. Конкретные задачи отличаются в зависимости от географического уровня, хотя общие цели всегда заключаются в том, чтобы гарантировать исчерпывающий сбор информации и создать условия для получения максимально возможного количества результатов для каждого географического уровня.

II.1 География коммун: организация работы по внесению немногочисленных изменений

9. Французские статистики используют очень стабильную и очень детализированную систему административного деления, включающую в себя 100 департаментов и, главное, 36 000 коммун (по европейской терминологии – NUTS 5), многие из которых имеют очень маленькие размеры (в среднем около 300 жителей, при этом 50% населения живет в коммунах с численностью населения менее 10 000 человек).

10. На этих двух уровнях с начала 80-х годов (коммуны и департаменты) для целей распространения применяются карты, основанные на картах НГИ 3/, и используется программное обеспечение SAS. В нескольких случаях часть атласов была изготовлена на договорной основе специализированными фирмами. В более позднее время в связи с картами коммун возникла значительная потребность в электронном распространении данных и НИСЭИ попытался выйти на этот рынок с французским распространителем "MapInfo".

3/ НГИ – Национальный географический институт (картографическое агентство Франции).

11. Случаи изменения границ между коммунами весьма редки, при этом многие из них связаны всего лишь с обменом сельскохозяйственными землями без жилых строений. Некоторые случаи слияния или разделения территории имеют место в промежутках между переписями. Например, в период с 1990 по 1998 год было зарегистрировано 16 слияний и 25 разделений. Они немногочисленны, да и большинство из них приходится на несколько районов Франции. НИЭСИ регистрирует эти изменения, поскольку они отражаются на численности населения и влияют на ретроспективные ссылки. Он регулярно публикует официальный регистр географических кодов и ежегодно распространяет его с помощью электронных средств информации. При переписях в слившихся частях коммуны из предосторожности производится отдельный подсчет: через несколько лет они вновь могут быть разделены из-за изменения консенсуса на местном уровне.

12. Для 28 300 коммун с численностью населения менее 1 000 человек 4/ (в 1990 году общая численность их населения составляла 9,3 млн. человек) предусматривается лишь один счетчик на коммуны без подразделения их территории на несколько переписных участков. Разделение территории таких коммун на более мелкие части для проведения переписи и распространения полученных результатов не очень разумно: люди имеют друг о друге больше информации, чем ее можно получить при любой переписи.

13. Для 7 400 коммун с численностью населения от 1 000 до 10 000 человек (общая численность населения 19,5 млн. человек) нужно нанимать по несколько счетчиков и формировать переписные участки. В целях ограничения рабочей нагрузки и концентрации ресурсов в крупных городах, а не в мелких городах и деревнях, где проводить перепись объективно легче, НИСЭИ разрешает муниципалитетам проводить географическую разбивку территории коммун без строгого предварительного контроля.

14. НИСЭИ участвует в географической работе главным образом применительно к 800 коммунам с числом жителей более 10 000 человек (общая численность их населения – 27,8 млн. человек). В этом случае работа ведется не на уровне коммун, а на более низком уровне – квартала.

4/ Это и следующее пороговое значение применяются не строго.

II.2 География на внутрикоммунальном уровне: проблемы большой важности

15. Как и в большинстве стран, где проводятся традиционные переписи, во Франции методика сбора данных является источником двух групп проблем:

- подразделения территории на единицы, могущие использоваться на практике;
- обеспечения того, чтобы зона действия каждого счетчика была определена правильно и чтобы перепись на каждом участке территории фактически проводилась лишь один раз.

II.2.1 Подразделение территории

16. Его цель заключается в формировании небольших участков территории, подходящих для целей распространения и сбора информации 5/. Опыт Франции показал трудность решения этих двух видов проблем в некоторых случаях и подвел нас к определению двух географических понятий, которые очень часто на местах имеют идентичное значение:

- квартал
- округ.

При проведении переписи 1990 года в коммунах с численностью населения более 10 000 человек было выделено 220 000 кварталов и 230 000 округов.

17. Информация никогда не предоставляется по более низкому уровню, чем квартал. Этому препятствует ограничение возможностей географической идентификации при распространении файлов на этом уровне. Проблема заключается в определении полезных географических единиц, т.е. таких единиц, которые имели бы небольшой размер 6/ и границы которых совпадали бы с какими-то существующими административными границами и некоторыми видимыми ограничителями типа улиц, рек, железнодорожных путей... На практике квартал, как правило, представляет собой наименьшую составную часть территории коммуны, ограниченную улицами. Другим требованием к нему являются минимально возможные временные изменения в промежутках между переписями. Наконец, он должен давать возможность соблюдать правила конфиденциальности и обеспечения качества. Что касается конфиденциальности, то, согласно НКИС (французский орган по вопросам охраны конфиденциальной информации), информация передается государственным органам (включая местные) после официального принятия ими на себя обязательства не допустить ее дальнейшего распространения.

5/ В некоторых странах эти две цели четко разделены.

6/ А иногда они бывают слишком маленькими по размеру.

18. Округ – географическая единица, используемая для сбора информации. Он характеризуется ограниченностью размера (менее 800 жителей 7/) и видимостью границ для низкоквалифицированного счетчика. Как общее правило, обе единицы – округ и квартал – идентичны, но иногда квартал может быть разбит на несколько округов ввиду его размера; это делается в тех случаях, когда очень крупные жилые здания на какой-то территории не разделены между собой улицами. На месте обычно не возникает трудностей с разграничением различных округов, но нынешние ГИС НИСЭИ, за некоторыми исключениями, характеризуют не строения, а лишь улицы.

19. НИСЭИ сотрудничает с местными муниципалитетами в работе по выделению кварталов и округов до переписи. Их заинтересованность объясняется результатами, которые они ожидают получить в будущем. Процесс длится чуть менее года: он начался в январе 1998 года и должен завершиться в ноябре 1998 года, а день переписи назначен на 8 марта 1999 года.

20. Наличие ГИС не является чем-то абсолютно необходимым: кварталы можно выделять на местных картах, предоставляемых, в частности, муниципалитетами; однако ГИС очень помогает в обеспечении единообразной регистрации этих границ после их определения для целей дальнейшего сбора или распространения.

II.2.2 Один округ – один счетчик

21. К каждому округу прикреплен счетчик, но на одного счетчика при проведении переписи может приходиться несколько округов. Обычная рабочая нагрузка на одного регистратора составляет 500–600 жителей. Она может быть чуть более 600, когда слишком трудно разделить округ на две единицы из-за отсутствия видимых внутренних границ. В городе она никогда не бывает значительно меньше 500, так как в этом случае зарплата счетчика не стимулировала бы его в достаточной степени 8/.

7/ И в этом случае данная пороговая величина применяется не строго.

8/ Зарплата счетчиков напрямую зависит от числа заполненных форм, возвращаемых ими в НИСЭИ.

22. Счетчиков для ориентации снабжают картой каждого округа и перечнем отрезков улиц, подлежащих охвату при проведении переписи 9/. Ввиду особенностей нынешнего организационного устройства НИСЭИ источники у этих двух информационных документов разные. В настоящее время нами поддерживаются две географические базы данных:

- ⇒ "REPLIC" 10/ - перечень участков улиц по каждому кварталу с указанием почтовых адресов в его крайних точках. Эта буквенно-цифровая база данных была создана после переписи 1982 года для целей распространения; вторая версия была создана перед 1990 годом для целей сбора информации, а третью намечается подготовить к ноябрю 1998 года.
- ⇒ "Base-îlot" - ГИС, построенная в сотрудничестве с НГИ и существующая с 1993 года. Программное обеспечение - ArcInfo. Она включает в себя характеристику осей ориентации улиц и почтовые адреса в конечных точках отрезков улиц, ограничивающих квартал. Помимо улиц, характеризуются и другие ограничители. Короче говоря, НГИ снабжает НИСЭИ информацией о геометрии улиц, НИСЭИ добавляет в ГИС границы кварталов, и вдвоем они служат источником информации о почтовых адресах.

23. К переписи 1999 года обе системы были дополнены характеристикой округов, не совпадающих с кварталами.

24. В обеих базах содержание данных об идентификационных номерах кварталов и округов и их структуре на местности идентично; есть некоторые различия в названиях улиц в тех случаях, когда, например, в одной базе они были обновлены, а в другой - еще нет. Основной проблемой здесь являются почтовые адреса. Это объясняется отсутствием физической связи между двумя базами: ни НГИ, ни НИСЭИ пока еще не сумели преодолеть эту техническую трудность. Но базовая информация собирается только один раз в местных органах самоуправления и других источниках, и эта информация обычно используется для поддержания обеих баз, но с использованием разных процедур.

25. Другая серьезная проблема обусловлена концептуальной трудностью определения почтовых адресов в некоторых случаях - когда, например, на краю отрезка улицы нет здания или когда там находится магазин, но отсутствуют единицы жилья, номер последнего подъезда здания, подлежащего регистрации при переписи, и номер последнего участка в разных по охвату источниках не совпадают, нарушая статистические правила.

9/ См. образец в приложении.

10/ "REPERTOIRE de Localisation Infra-Communale" (Реестр данных о местоположении объектов внутри коммун).

26. Вследствие этого в материалах, предоставляемых счетчикам, имеются определенные несовпадения. Проверки показывают, что при сборе сведений в районе они успешно решают проблему таких расхождений.

27. Программные средства "Base-îlot" скоро будут обладать новыми функциональными возможностями. Кроме того, в 2000 году на основе базы данных "Base-îlot" будет создана новая версия "REPLIC" после исправления почтовых адресов в первой посредством их автоматического сличения с данными "REPLIC".

28. Когда в 1993 году было принято решение о замене обычных кадастровых карт цифровыми картами, в НИСЭИ возникла еще одна проблема. На существовавших кадастровых картах были очень детально изображены участки и здания, но они обновлялись не часто. Несмотря на частые задержки с обновлением, эти карты раньше использовались специалистами НИСЭИ, большинство из которых считало такую подробную информацию об участках и зданиях необходимой для счетчиков. Без такой информации счетчик может потерять ориентацию и попасть не в свой, а в соседний округ или пропустить второстепенные здания, расположенные не на улице, а, например, во втором ряду.

29. В конце 1993 года была проведена комплексная проверка с охватом 108 000 единиц жилья в 1 075 округах. Каждый счетчик обслуживал две группы округов. В первом он использовал кадастровые карты, а во втором – цифровые. Кроме того, на место для ориентировки выходил специально подготовленный инструктор со всеми имеющимися материалами. В целях сведения к минимуму влияния личности счетчика был проведен тест Вилкоксона по сопоставлению числа единиц в округах. В результате его не было выявлено никаких значительных различий между двумя методами, за исключением тех случаев, когда план строения отличался большой сложностью; было решено в таких случаях проводить оцифровку некоторых дополнительных объектов, чтобы помочь счетчикам хорошо ориентироваться на месте [FANOUILLET].

II.2.3 Наем и назначение счетчиков

30. Формально счетчиков во Франции нанимают муниципалитеты, получающие необходимые деньги от НИСЭИ. НИСЭИ должен только проверять компетентность кандидатов, представляемых местными органами самоуправления, обучать их и обеспечивать им разумную рабочую нагрузку. НИСЭИ отвечает за всю бумажную работу, проводимую инструкторами и, конечно, за их подготовку и постановку им задач.

31. Применительно к переписи 1999 года эта задача начала выполняться в июне 1998 года, при этом никакие средства ГИС не использовались нигде, кроме некоторых местных управлений, где с ее помощью составлялись задачи инструкторов. В этой области имеются явные возможности для улучшений положения в будущем.

II.2.4 Организация сбора информации

32. Во Франции период сбора данных длится четыре недели. Затем в течение примерно шести месяцев результаты подсчетов на местах проверяются одновременно местными органами самоуправления и НИСЭИ, после чего публикуются предварительные данные (июль 1999 года) и результаты окончательного подсчета (конец 1999 года). Весь этот процесс децентрализован: на период сбора данных на территориальном и местном уровнях создаются иерархически организационные структуры, а обработкой результатов подсчетов занимаются 22 постоянных территориальных управления НИСЭИ.

33. Географические аспекты работы по сбору данных – от низшего уровня до общенационального – определяются до дня переписи, если не считать некоторой неизбежной работы по обновлению информации, проводимой в локальном масштабе в начальной стадии работ на местах. Из ГИС извлекается буквенно-цифровой файл, используемый для подготовки карт счетчиков "Base-îlots"; он включает в себя полный список всех географических единиц. Этот файл является стержневым элементом программных средств управления, которые должны использоваться в период обработки данных.

34. Мы решили не использовать для облегчения организационной работы средства ГИС. В то время преимущества ГИС не представлялись очевидными из-за необходимости единообразной обработки всех единиц и неиспользования топографических связей между этими единицами, делающих ГИС уникальным инструментом.

II.2.5 Дополнительные обследования

35. Во Франции каждое лицо должно заполнить единую для всех форму: коротких и длинных форм не существует (в отличие, например, от США). Следовательно, нет и необходимости в проведении различия между домохозяйствами. Но зато традиционно проводится дополнительное обследование для получения информации о составе семьи, фертильности, работе женщин и ее влиянии на воспитание детей. Выборочная совокупность для этого обследования составляется на основе списка округов; но точнее было бы сказать, что это – выборка переписанных участков. Если не считать этого, то обследование проводится по той же процедуре и в то же время, что и обычная перепись. С учетом процедуры сбора форма специально задумана короткой (4 стр.). По этой же причине она должна быть ясной. Несмотря на это объем выборки может быть очень большим при весьма низком уровне расходов (в 1990 году – 340 000 респондентов).

36. В 1999 году будет проведено три обследования такого рода:

- "исследование" истории семьи (формы должны быть заполнены 250 000 женщин);
- "исследование" истории семьи (формы должны быть заполнены 150 000 мужчин – нововведение 1999 года);

- обследование "повседневная жизнь и здоровье" (полное обследование 200 000 домохозяйств, которое позволит всесторонне изучить проблему инвалидности по состоянию на конец 1999 года на примере 20 000; это второе нововведение 1999 года).

37. Структура этих выборочных обследований определялась исходя из двух целей - обеспечения региональной репрезентативности и минимизации действия факторов, препятствующих работе счетчиков и их инструкторов. Эти препятствия побудили нас использовать на обоих уровнях - округов (разумеется) и групп счетчиков, инструктируемых одним и тем же лицом - своего рода гнездовой метод. Ограничения такой системы выборки частично компенсируются возможностью использования многочисленных социально-демографических переменных, получаемых благодаря одновременности переписи.

38. В случае переписи 1999 года программные средства ГИС для выборочных обследований не используются, хотя, как представляется, эта работа имеет явно географический характер. И в данном случае мы могли бы повысить за счет них производительность и, возможно, достоверность результатов.

II.3 Проверка существования строений и их местоположения

39. Опыт предшествующих переписей показывает, что примерно половина случаев недоучета объясняется пропуском жилищных единиц. В 1990 году впервые для проверки полноты их учета использовались файлы, полученные со стороны. Налоговая администрация подготовила файлы данных по налогу на занятые помещения с некоторой задержкой, но для этой цели интерес представляют даже предварительные данные. В 1990 году эти файлы были использованы в городах с населением свыше 20 000 человек после сортировки по названию улиц. Такая сортировка не дала удовлетворительных результатов из-за нередко очень большой длины улиц и их прохождения по территории нескольких округов, поэтому последствия для сбора данных были значительными: послеперписное обследование показало, что в городах, где процесс проверки был успешно применен, доля пропусков снизилась с 0,6% до 0,2% [COEFFIC].

40. В отношении переписи 1999 года было принято решение о дальнейшем расширении этой системы и ее распространении на все города с населением свыше 10 000 человек и продолжении подготовки таких файлов с той целью, чтобы счетчики и инструкторы легко могли их использовать: в январе 1999 года на уровне округов будет произведена

географическая привязка жилищных единиц 11/ и строений с использованием разработанной НИСЭИ системы, называемой "îlotage". В качестве справочного файла для целей определения местоположения используется база данных "REPLIC" (см. выше). Очень эффективно автоматическое кодирование, при этом охват кодированием почти во всех городах превышает 95%. Эти цифры не должны удивлять, так как для ведения этих файлов местные органы самоуправления и налоговые службы поддерживают между собой тесные связи 12/.

41. Вместо системы "REPLIC" с буквенно-цифровыми файлами мы могли бы использовать ГИС "Base-îlot" с программными средствами ГИС для геокодирования. Так, например, в настоящее время проводятся испытания с использованием программы "MapInfo", но их цели идут дальше нынешних целей, определенных для переписи 1999 года (см. ниже). Поскольку цель заключается лишь в обеспечении счетчиков и инструкторов местными списками, средства ГИС на практике не являются чем-то абсолютно необходимым.

II.4 География и обработка данных

II.4.1 Производственный процесс

42. Обработка данных переписи включает считывание, кодирование, условное исчисление, выборку данных, если не вся информация обрабатывается одинаково детально или в одном месте, а в некоторых странах – и проведение оценки с использованием полученных со стороны данных на районном или провинциальном или же местном уровне. Последняя операция этого перечня сейчас во Франции не используется и не будет рассматриваться в этом документе, но нам придется столкнуться с этой проблемой в ближайшем будущем при решении проблем, связанных с проведением непрерывной переписи (см. ниже).

11/ Эти файлы характеризуют как жилищные, так и другие единицы (гаражи, отдельно расположенные помещения для хранения, квартиры, используемые в профессиональных целях, ...). До использования они должны быть по возможности выверены.

12/ Налог на занятые помещения обеспечивает значительную часть поступлений в бюджеты муниципалитетов.

43. Обработка данных при таких масштабах переписи – вопрос производственного характера, поэтому ее следует организовывать соответствующим образом. Одним из ключевых аспектов здесь является разделение всего массива данных на подмножества. Поскольку процесс обработки у нас разбивается на три важных компонента, критерий, используемый для формирования подмножеств, не будет одинаковым на протяжении всего процесса.

44. Применительно к переписи 1999 года речь идет о трех следующих важных компонентах:

- официальный подсчет: работа будет организована по коммунам, в основном независимо от их размера, и должна быть проведена в октябре 1999 года;
- первый этап статистической обработки: сканирование всех форм и обработка изображений (ОРС и считывание данных с изображений в случае вопросов, не имеющих конфиденциального характера) без выделения вопросов, касающихся работы (коды приносящих доход занятий и отраслей); весь процесс обработки организуется по группе коммун сразу после получения форм для сканирования; по одной такой группе обрабатываются формы, заполненные 250 000 человек. Завершить этот процесс и поочередно обнародовать результаты по департаментам планируется во втором и третьем кварталах 2000 года;
- второй этап статистической обработки: производится обработка данных выборки, охватывающей примерно 1/4 всех форм, которая посвящается вопросам работы. И в этом случае обработка организуется по группам коммун, но не по тем же, что и раньше, так как их размер больше в силу двух причин: 1) меньшего объема работы на каждую форму; 2) разработки плана выборки для контроля за качеством для более крупных единиц в целях ограничения затрат. Результаты ожидаются в третьем и четвертом кварталах 2000 года.

45. Этот выбор можно обсуждать – даже в условиях Франции. Например, мы могли бы сэкономить время за счет улучшения координации двух вышеизложенных этапов статистической обработки. Сейчас, когда после недавнего решения об использовании оптических методов считывания 13/, весь процесс видоизменен, такой подход к организации работы может показаться несколько консервативным. В любом случае географический фактор может в значительной мере отразиться на обработке данных.

13/ Это решение было принято в июне 1998 года, но подготовка этой альтернативы традиционному способу считывания данных была начата за два года до этого.

II.4.2 Выборочная совокупность для обработки

46. Выборка (1/4), упоминавшаяся в связи со вторым этапом обработки данных, представляет собой выборочную совокупность домохозяйств, отбираемую по принципу "округ за округом". Этот метод может быть улучшен при проведении переписи 1999 года благодаря использованию данных, полученных на первом этапе статистической обработки, например по таким показателям, как размер домохозяйств и социально-экономическое положение главы домохозяйства. Влияние географического фактора весьма однородно, так как он призван лишь поддерживать максимально высокий уровень репрезентативности результатов на местах.

47. На раннем этапе этого процесса отбирается еще одна выборочная совокупность, чтобы проверить весь процесс практически в реальном масштабе и в итоге подготовить несколько первых демографических и социально-экономических отчетов на уровне районов и страны. Отношение выборки к генеральной совокупности составляет 1 КПО.

48. Традиционные методы, существовавшие до появления методов оптического распознавания, не позволяли нам использовать альтернативные способы составления такой выборочной совокупности: по причине затрат она должна была представлять собой выборку округов, где для дальнейшей обработки данных выборки, охватывающей 1/4 генеральной совокупности, необходимо было кодировать одно домохозяйство из четырех. К переписи 1990 года были проведены методологические исследования и составлена сбалансированная выборка округов с целью оптимизации ее качества [DEVILLE]. При проведении переписи 1999 года сканирование всех форм будет завершено до декабря 1999 года; это даст возможность непосредственно отбирать выборочную совокупность домохозяйств без значительных расходов.

49. Таким образом, географический аспект в этой работе уже не присутствует, за счет чего повышается качество статистической информации.

II.4.3 Условные исчисления и география

50. География, напротив, является одним из основных элементов методов условного исчисления, используемых во Франции наряду с "методом вмененных ответов", который широко применяется при отсутствии ответов и непоследовательных ответах, когда их невозможно заменить детерминистическим условным исчислением на основе остальной части вопросника. Метод "вмененных ответов" предполагает близость социально-демографических характеристик и географических данных.

51. Влияние географического фактора глубоко не изучалось, но на местах требуется иметь в виду, что оно, вероятно, существует.

III. Сбор данных при проведении переписей во Франции и география: перспективы

52. В сентябре 1994 года из-за проблем с бюджетом очередная перепись, первоначально подготавливавшаяся к 1997 году, была перенесена на 1999 год. Целый ряд причин побудил НИСЭИ подумать о применении при ее проведении новой методики [JC DEVILLE, M JACOD]. Главная причина связана с межпереписным периодом во Франции (в последние десятилетия он длился 7-9 лет), который пользователи считают слишком длинным и неопределенным (взять хотя бы перенос переписи на 1999 год). Другая причина - возможность нарушения конфиденциальности при сотрудничестве с местными органами самоуправления. Третья причина связана с проблемой бюджета и имеет внутренний и внешний аспекты.

53. Решение еще не принято, но уже преодолено несколько препятствий: внутри организации создана структура для реализации этого проекта в короткие сроки (дата начала сбора данных может быть намечена на конец 2001 года), с экспертами правительства обсуждены последствия на уровне для нормативного регулирования в подведомственных им сферах, через Государственный совет 14/ проверена конституционность такого рода переписи.

54. Следующими шагами являются:

- организация широких консультаций пользователями в апреле 1999 года после завершения периода сбора данных в рамках переписи 1999 года;
- подготовка парламентской процедуры, которая может состояться в конце 1999 года;
- формирование нескольких сценариев, в зависимости от возможных результатов этих двух процедур, с той целью, чтобы своевременно начать проект и уложиться в крайний срок, предусмотренный в 2001 году.

III.1 Зачем прибегать к географии?

55. В этом документе предполагается лишь в общих чертах описать некоторые формы использования географии, которые мы намерены применять, учитывая возможность ее превращения в один из основных инструментов при новом подходе к организации 15/.

14/ В апреле 1998 года правительство обратилось за консультацией по этому вопросу к Государственному совету и в июле получило утвердительный ответ.

15/ Более полное описание содержится в работе [ISNARD].

Вторая причина отражения этого вопроса в настоящем документе состоит в том, что эти новые задачи будут выполняться в очень скором будущем. Например, сейчас уже идет разработка нового продукта - "répertoire d'immeubles localisés 16/", сокращенно называемого RIL, поскольку его применение представляется целесообразным даже в том случае, если новая методика проведения переписи применяться не будет.

56. Не углубляясь в детали процедуры, отметим, что вся совокупность, состоящая из 36 000 коммун, будет разделена на два подмножества:

- в небольших коммунах (пороговый уровень может составлять 10 000 жителей) могут проводиться традиционные переписи, но с повторной выборкой на основе пятилетнего цикла;
- в каждом городе, где численность населения превышает этот пороговый уровень, могут ежегодно проводиться выборочные обследования с охватом 1/20 генеральной совокупности.

57. Годовые оценки будут представляться практически каждой коммуной на основе такого частичного сбора данных наряду с внешними данными. Эти внешние данные могут быть номинативными или неноминативными. На сегодняшний день этот вопрос еще не прояснен, и формы их использования могут несколько варьировать, в зависимости от его решения. Такими данными, вероятно, будут данные по налогу на занимаемые помещения, которые мы уже используем в переписи 1999 года (см. выше раздел II.3), и данные по социальному страхованию, к которым мы также имеем доступ при проведении конкретных операций с участием НИСЭИ, связанных с контролем избирательного процесса.

58. При проведении выборочных обследований в небольших коммунах можно было использовать средства ГИС для учета их некоторых топологических особенностей. Но наибольший интерес ГИС, по-видимому, представляли бы при организации процесса сбора в городах, особенно при организации выборочных обследований и сбора данных.

III.1.1 Проведение выборочных обследований в городах

59. Годовое выборочное обследование с охватом 1/20 генеральной совокупности 17/ должно быть репрезентативным для города. В целях обеспечения репрезентативности на местном уровне, который предстоит определить, выборочная совокупность должна охватывать всю территорию города. Кроме того, пять выборочных обследований,

16/ Реестр строений с характеристикой их местоположения.

17/ Этот параметр может быть средним. В маленьких городах он может составлять где-то 1/5, а в самых крупных - 1/25.

проводимых за пятилетний цикл, должны быть независимыми друг от друга, так как собираемые данные с целью улучшения их репрезентативности будут на протяжении цикла суммироваться. Эти условия, естественно, побуждают к более широкому использованию ГИС.

60. Во Франции не существует исчерпывающего перечня жилищных единиц (упоминавшийся до этого файл данных по налогу на занимаемые помещения может использоваться для увеличения полноты охвата, но не для проведения выборочных обследований. Составление такого перечня, и особенно его последующую актуализацию, мы считаем слишком сложной задачей. В этой связи могут также возникнуть трудноразрешимые проблемы с конфиденциальностью. Поэтому мы не можем формировать выборочную совокупность жилищных единиц после добавления в такой перечень необходимых географических характеристик.

61. С учетом этого предполагается использовать территориальные выборочные обследования типа проводившегося нами ранее обследования рабочей силы (ОРС) или германских микропереписей. Но в отличие от ОРС оба процесса (разграничение участков, выборка на основе перечней участков) должны носить как бы индустриальный характер, т.е. быть автоматизированными и сопровождаться проверками качества ввиду их масштаба и влияния на результаты.

62. В отличие от списка жилищных единиц, составление и ведение перечня строений представляется осуществимым как в техническом плане, так и с точки зрения возможностей бюджета. В этом случае не возникает проблемы с конфиденциальностью, да и сама эта совокупность является весьма стабильной. Этот перечень будет материализован в виде дополнительного уровня нашей нынешней ГИС "Base-flot".

63. Нужно ответить на несколько вопросов (см. ниже):

- Как формировать участки на основе этого уровня?
- Как предоставить правильную информацию счетчикам?
- Можно ли иначе организовать сбор?

64. Само собой разумеющиеся вопросы:

- Какими должны быть технические характеристики этого дополнительного уровня?
- Какими будут инвестиционные и эксплуатационные расходы и процедуры?

III.1.2 Как формировать участки?

65. Предположим, что имеются лишь небольшие строения с небольшим числом жилищных единиц. Сформировать участок можно легко, объединяя смежные строения до тех пор, пока число жилищных единиц не достигнет определенной величины (например, 10). Этот процесс был апробирован в 1997 году в одном из заморских

департаментов Франции (Реюньон), где, за некоторыми исключениями, расселение действительно имеет такой характер. Этот эксперимент показал, что процесс разграничения территории и связанный с ним процесс формирования выборки могут быть автоматизированы благодаря использованию топографической информации, включенной в ГИС.

66. В отношении более крупных строений предполагается получить достаточный объем информации для их разделения на более мелкие части – подгруппы квартир – вероятно по этажам. Кроме того, эту категорию строений можно рассматривать отдельно, при этом конкретные застроенные районы, охватывающие лишь часть крупных строений, могут составлять один или несколько уровней.

III.1.3 Какую информацию предоставлять счетчикам?

67. Выше уже были описаны карты счетчиков для переписи 1999 года и охарактеризована внутренняя дискуссия по поводу того, насколько детализированной должна быть информация, чтобы они могли хорошо работать. Новый подход должен дать новый толчок этой дискуссии, так как различимость участков на местности, безусловно, меньше, чем различимость кварталов. Предварительный вывод состоит в том, что нам необходимы карты с указанием местоположения строений и что в некоторых случаях могут быть полезны комментарии 18/ или даже схемы, подготовленные после первого обследования.

III.1.4 Организация сбора в городах: зоны контролируемого качества

68. Характерной чертой нового процесса по сравнению с традиционным 19/ является наличие времени. В то же время на национальном и всех местных уровнях существует значительная потребность в оценке качества, и эта потребность, вероятно, еще больше возрастет, тем более, что местные органы самоуправления берут на себя все больше политических функций и все активнее стремятся к получению данных, утрачивая при этом возможности вмешательства в процесс сбора данных при проведении переписи нового типа.

18/ Например, "чтобы выйти к этому строению, необходимо, пройдя через подъезд здания № xx, подойти к правому краю внутреннего двора".

19/ Период сбора будет, вероятно, продлен с одного месяца (продолжительность переписи классического типа) до восьми или девяти месяцев учебного года, при этом будет опускаться лишь период каникул. Это позволит НИСЭИ осуществлять сбор данных без помощи местных органов самоуправления.

69. Существует широкий круг критериев оценки качества результатов, полученных на местах: исчерпываемость или – в данном случае – увеличение выборки; методы корректировки; исходное качество внешних источников; социально-демографическая гетерогенность населения; трудности сбора данных в зоне, измеряемые долей не ответивших лиц... Уделение достаточного времени перечисленным аспектам и использование некоторых инструментов, например корректировка объема выборки, обучение счетчиков, подготовка процесса сбора на месте с использованием особых отношений с местными средствами массовой информации, процедуры условного исчисления... могут позволить НИСЭИ отреагировать на недостатки в качестве. Это еще одно из значительных преимуществ непрерывной переписи.

70. Один из предполагаемых вариантов состоит сейчас в формировании "зон контролируемого качества", где НИСЭИ должен будет поддерживать его на высшем уровне. Если будет выбран этот вариант, то ГИС вновь станет основным инструментом анализа и демонстрации результатов.

III.1.5 Технические характеристики дополнительного уровня ГИС

71. "Base-îlot" будет включать в себя новый уровень – RIL. Этот уровень будет охватывать точки расположения строений. Мы не собираемся отмечать границы участков или показывать форму строений, как это делается на кадастровых картах, а будем использовать лишь точки. Эти точки будут наноситься приблизительно – по крайней мере, в ближайшие годы. Мы не можем пойти на точное (например, отклонение в 2 м) указание местоположения, до это и не существенно при проведении обследований или переписи (см. выше эксперимент 1993 года). Важными же элементами мы считаем полноту, порядок следования объектов на участке улицы и указание того, на какой стороне участка он фактически расположен. Позднее мы думаем указывать местоположение с большей точностью в тех случаях, когда это будет представляться полезным и поможет избежать трудностей при сборе, например за счет использования кадастровых карт или ортоснимков в качестве фона при оцифровке или GPS.

72. На уровне RIL привязка к точкам будет делаться для очень небольшой группы данных: о существующих жилищных единицах и оценке их числа на основе последнего обследования или переписи и, вероятно, существующих единицах, предназначенных для других целей (магазин, ...), в случае комментариев, могущих помочь счетчикам в поиске нужной единицы, и идентификационных номеров для выхода на статистические файлы, используемые при проведении переписи.

III.1.6 Инвестиционные и эксплуатационные расходы и процедуры

73. В настоящее время производится адаптация программных средств, разработанных для управления системой "Base-îlot". Их можно будет использовать в сентябре 1999 года, что позволит своевременно интегрировать данные переписи 1999 года (около 4,5 млн. строений). В 2000 году эта интеграция будет произведена по большинству городов с использованием программного обеспечения MapInfo для

автоматического геокодирования на основе почтовых адресов, полученных во время переписи. Сейчас проводятся испытания для оценки того, на какое число неавтоматически кодируемых адресов необходимо рассчитывать. По нынешним оценкам, обработка этих адресов потребует 30 000 рабочих часов.

74. Поддержание такой географической базы данных предполагает постоянное взаимодействие с местными органами самоуправления, определяющими и применяющими градостроительные правила. НИСЭИ уже поддерживает такие связи, но не везде на постоянной основе: иногда они ограничиваются необходимым для переписи обновлением системы "Base-îlot" и разграничением кварталов. Поскольку постоянная работа становится обязательной, нам необходимо налаживать официальные партнерские связи с местными органами самоуправления.

75. Принимаются во внимание и другие источники. Например, хорошо осведомлены о любых геометрических изменениях в уличной сети компании, занимающиеся водоснабжением (или другие поставщики коммунальных услуг), которые получают актуализированную информацию о строительстве и сносе сооружений, имеющую важное значение для поддержания RIL.

76. Третьим источником информации может быть также полевая работа при проведении мероприятий переписи "нового типа", но, исходя из соображений, связанных с выборочными обследованиями, мы предпочитаем рассматривать ГИС (включая RIL) не в качестве конечного продукта, а как ресурс, используемый в процессе работы.

77. Наши нынешние партнерские отношения с НГИ следует поставить на новую основу. Вероятно, их сфера не будет расширена за счет RIL, а останется ограниченной нынешней системой "Base-îlot". Новая основа этих партнерских связей, вероятно, будет иной: благодаря поддержанию постоянных отношений с первичными источниками мы располагаем потенциалом для того, чтобы обновлять систему "Base-îlot" самостоятельно (геометрия улиц, названия улиц, почтовые адреса и кварталы); НГИ, располагающий более квалифицированными специалистами для точной картографической работы, подключится к работе по повышению геометрического качества базы данных с некоторой задержкой.

78. Согласно оценкам, годовые эксплуатационные расходы будут соответствовать 80 000 рабочих часов, половина из которых потребуется для RIL. Ежегодное число изменений в RIL (создание и ликвидация строений) должно быть менее 200 000.

III.2 Переход к другим целям

79. Помимо жилых зданий существуют также промышленные и торговые структуры, размещаемые в специальных зданиях или вместе с жилищными единицами, а также различные объекты общего пользования (игровые площадки, больницы, школы...). При проведении опросов в рамках переписи нового типа счетчики работали бы лучше, если бы им было известно расположение подлежащих обследованию сооружений относительно других

сооружений любого типа. Помимо этого первого довода в пользу расширения сферы охвата RIL за счет других структур, существуют и конкретные причины, характерные для каждой категории.

80. Ведение реестра предприятий предполагает получение точных адресов, которые должны характеризовать местоположение существующих строений; сейчас такой контроль тщательно не ведется (проверяется лишь существование улиц), но наличие RIL делает его возможным. Распространение данных из Реестра предприятий 20/ предполагает определение точного местонахождения единиц. В настоящее время это делается на последующем этапе на уровне кварталов, но было бы экономичнее производить эту операцию непосредственно при кодировании новой единицы.

81. Существует спрос и на данные о точном местонахождении объектов общего пользования. Первое обследование было проведено в 1993 году, но формирование RIL дает нам возможность обновить методику и получить актуализированные данные, необходимые для придания законченности набору переписных данных, характеризующих совокупность. Эти единицы опять-таки не являются мобильными, и эксплуатационные расходы не должны быть большими.

82. В силу всех этих причин НИСЭИ принял решение о скорейшем внедрении RIL после завершения стадии сбора переписи 1999 года даже при невозможности принятия к тому времени решения по будущей переписи.

20/ Важный продукт НИСЭИ, поскольку это не "статистический", а административный реестр предприятий и все его содержание не подпадает под действие статистического законодательства.

COEFFIC, Nicole, L'enquête de mesure de l'exhaustivité du recensement de 1990, Recensement de la population 1990 - Innovations méthodologiques, Novembre 1995, INSEE-Méthodes n° 52-53

DEVILLE, Jean Claude, L'échantillonnage équilibré de districts et l'exploitation au 1/20 du recensement, Recensement de La population 1990 - Innovations méthodologiques, Novembre 1995, INSEE-Méthodes n° 52-53

DEVILLE, Jean Claude, JACOD, Michel, (1996) Replacing the traditional French Census by a large scale continuous population survey, Annual Research Conference, March 1996, Washington, Bureau of the Census

FANQUILLET, Jean Christophe, XXX, Recensement de la population 1990 - Innovations méthodologiques, INSEE-Méthodes n° XXX (to be published)
~~ISNARD, Michel, 1999, post 1990 Census Research Project in INSEE, рабочий документ, с которым можно познакомиться, обратившись к автору по адресу "michel.isnard@insee.fr."~~
