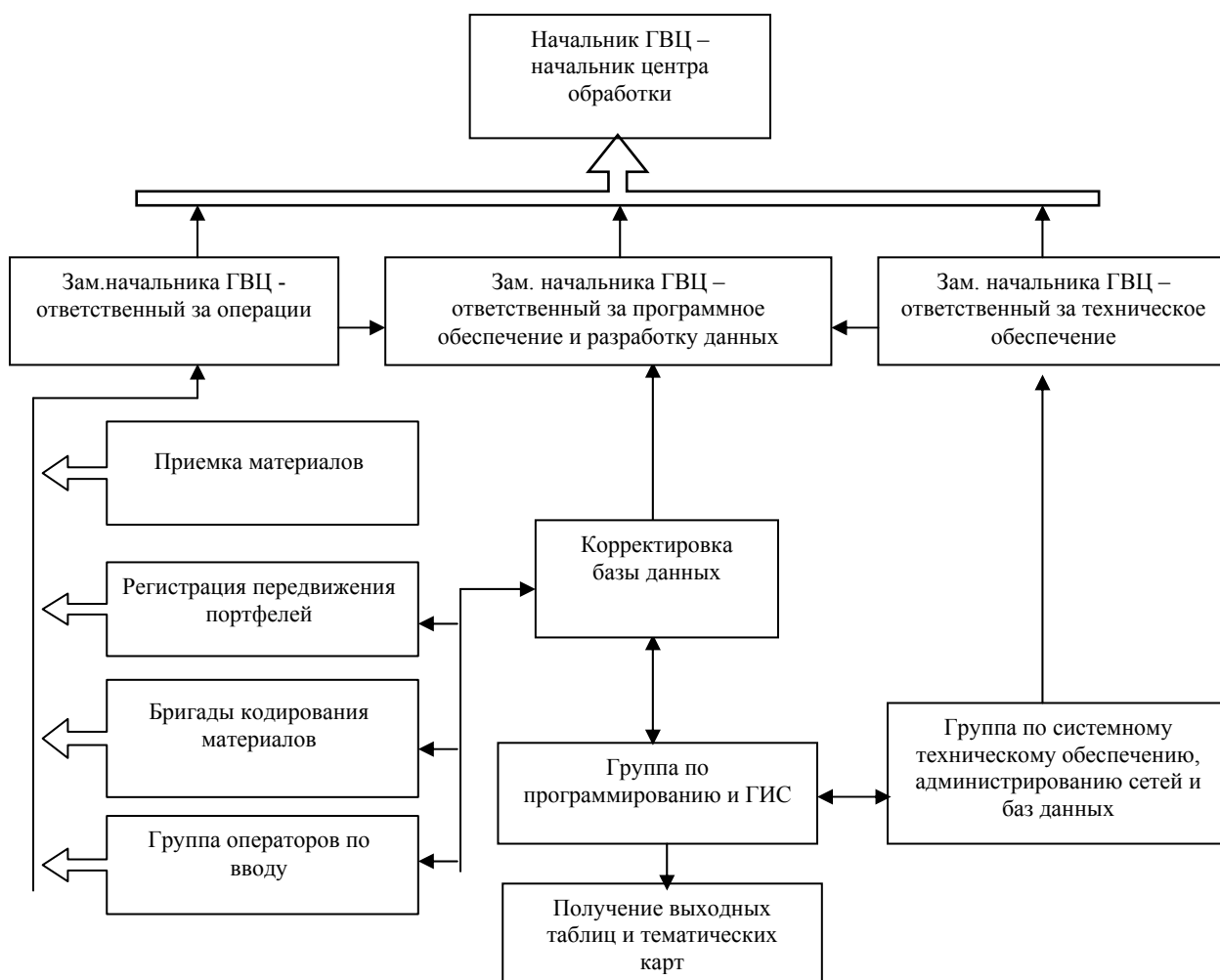


Тема: Технологический цикл обработки материалов Переписи населения и жилищного фонда Кыргызской Республики 2009 года

Обработка материалов переписи - один из наиболее важных этапов проведения переписи, в процессе которого осуществляется преобразование первичных данных в проверенные и агрегированные данные. Весь процесс обработки данных будет проводиться в Главном Вычислительном центре Нацстаткомитета. Этап обработки включает в себя следующие операции:

- кодирование;
- ввод данных;
- корректировка (редактирование);
- получение выходных таблиц
- получение тематических карт

Чтобы добиться положительного результата этого этапа переписи, необходимо внимательно руководить операциями в центре обработки. Для координации и управления всеми операциями этапа обработки в ГВЦ была создана следующая структура подразделений:



Выбор метода ввода

В настоящее время в мировой практике при вводе материалов переписи используются в основном три метода ввода данных:

- ручной ввод с помощью персональных компьютеров
- сканирование
- ввод данных непосредственно респондентами через Интернет.

Из-за отсутствия опыта применения сканеров и web-страничек для ввода данных статистических обследований, а также в условиях жестких временных, финансовых ограничений, дефицита квалифицированных ИТ-специалистов мы посчитали целесообразным использование ручного метода ввода данных, как более дешевый и дающий возможность укрепить техническое состояние статорганов, использование компьютеров в последующих статистических работах и обследованиях.

Подбор и подготовка кадров по обработке (кодировщиков, операторов, экономистов)

Целью успешной работы на этом этапе должен быть найм наиболее квалифицированного персонала из числа кандидатов в пределах установленного времени и предусмотренного бюджета. При этом должны заранее приниматься меры для подбора нужного числа сотрудников соответствующего профиля, необходимого для каждой из операций по обработке материалов переписи.

Исходя из необходимости обеспечения эффективности и экономии средств, кадры подбираются с учетом их компетентности. Частично будет привлечен персонал, участвовавший в предыдущих переписях, для работ по обработке данных – работники, владеющие определенной специальностью.

Учебная программа для кодировщиков и операторов предоставляет возможность практиковаться под наблюдением инструкторов к выполнению операций, которые будут впоследствии осуществляться.

Планируется нанять следующий временный персонал:

Наименование	Количество переписных листов	Норма на 1 чел. в день	Сроки исполнения	Требуемая численность
Приемщики	5300000	6000	40	22
Кодировщики	5300000	215	88	280
Операторы	5300000	347	88	174
корректировщики	4637500	1413	88	37

Для разработки программного обеспечения ввода, контроля, корректировки данных переписи и получения сводных итогов было решено привлечь внешних разработчиков. В октябре этого был проведен тендер на разработку программного обеспечения, по результат которого был заключен договор с частной фирмой. Эта программа должна быть разработана с учетом максимального ускорения работы операторов и получения качественного материала.

Составление расчетов и приобретение компьютерного оборудования и другой техники.

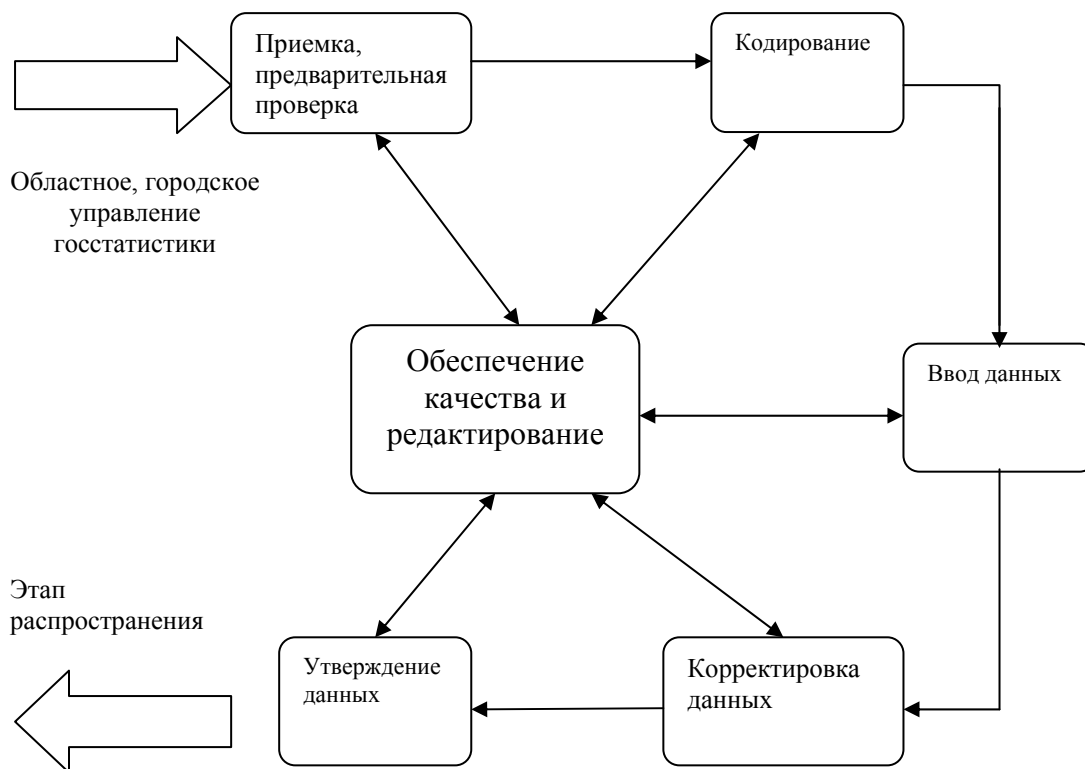
Необходимое количество компьютерного оборудования и оргтехники определяется на основе существующей технической базы, метода ввода информации, предварительных расчетов численности персонала по обработке данных и технологического цикла.

Анализ технической оснащенности показал, что на начало проведения переписи недостаточно компьютерного оборудования и оргтехники, необходимых для проведения такого крупномасштабного мероприятия. На основе анализа были сделаны расчеты.

Цикл обработки данных переписи 2009 года

Хотя обработку данных можно в значительной степени рассматривать как линейный цикл, все операции обычно осуществляются одновременно. Сначала выполняются первые операции, такие как получение и регистрация переписных форм, но вскоре и приступают ко всем остальным операциям, как только в результате первых операций накопится достаточно материалов для обработки. Важно направлять и тщательно координировать движение переписных форм по операциям, с тем, чтобы обеспечить наличие на каждой операции достаточного количества переписных форм для всех обработчиков. Следует создать буферный или резервный запас переписных форм на стыках между операциями (например, на две недели работы), с тем, чтобы избежать простоев из-за отсутствия форм для обработки. Ввод данных не следует начинать, пока не накопится запас переписных форм, достаточный на две недели. Если для обработки такого количества переписных форм на всех операциях до ввода данных требуется одна неделя, то ввод данных не следует начинать ранее третьей недели.

Все операции взаимосвязаны одна с другой через непрерывно выполняемую операцию обеспечения качества. Качество и количество продукции каждой операции цикла обработки оказывают непосредственное влияние на успех или неудачу следующей операции или других последующих операций.



Приемка материалов

Материалы переписи представляются по территориальным массивам строго по графику. Областные (городские) управления госстатистики сдают материал по районам. Каждый портфель должен быть пронумерован. Нумерация портфелей ведется по районам, каждый из них начинается с 1. Необходимо проверить комплектность портфеля, правильность заполнения адресной части. Коды сопроводительного листа, т.е. коды территории заполняются в соответствии со «Справочником территорий». Прием материалов заканчивается оформлением описи массива отдельно на городское и сельское население в 2-х экземплярах. Эта операция предназначена как предварительная стадия получения качественного, достоверного материала переписи.

Кодировка

Основными функциями кодирования является подготовка переписной документации к автоматизированной обработке. Процесс кодирования материалов переписи проводится ручным способом с использованием справочников и инструкций. Процесс кодирования будет проводиться бригадами кодировщиков, каждую из которых возглавлял высококвалифицированный специалист с опытом работы в статистике. Кодировщики будут специализироваться на определенных операциях. Одна группа будет проверять и кодировать социально-демографические показатели, другая - экономические характеристики. После кодировки портфели поступят на ввод.

Ввод

Основные задачи ввода - проверка на полноту, т.е. обеспечение ожидаемого числа опросных листов и исключительность, т.е. обеспечение только одноразового ввода каждого раздела опросного листа, проверка на соответствие допустимому интервалу значений и на непротиворечивость.

Опись массива включает в себя коды областей, районов, кенешей, населенных пунктов, номера портфелей, число населенных пунктов, портфелей и число переписанных лиц в данном массиве. При вводе описи массива регистрируются дата поступления его на ввод. Для учета работы оператору присваивался код и регистрируется дата выдачи портфеля.

После завершения работы оператора портфель с проставленными кодами ошибок передается верификатору (повторный ввод) или корректировщику. По проблеме верификации были рассмотрены 3 варианта:

- контрольная сумма на переписной лист;
- двойной ввод;
- частичные контрольные суммы.

В режиме работы верификатора производится исправление ошибок допущенных оператором и подсчет их количества.

- оператор в службе регистрации получает 1 портфель для ввода;
- в течение смены оператор вводит информацию;
- в конце смены оператор обязан вернуть портфель, независимо от завершенности;
- введенная информация контролируется автоматически и помечается, как ошибочная, в случае если программой контроля обнаружены ошибки.

Корректировка

Основная задача программы контроля - проверка всех физических единиц (список, переписной лист, жилищные условия) на соответствие. Корректировка будет проводиться группой высоко квалифицированных специалистов с компьютера. В режиме работы корректировщика производится исправление ошибок допущенных переписчиком, оператором, кодировщиком и подсчет их количества.

В данном режиме номера ошибочных портфелей высвечиваются на экране компьютера. Выбранный на корректировку портфель был доступен только одному корректировщику. Технология редактирования полностью исключает бумажные протоколы ошибок, необходимость принтеров.

Учет работы оператора

В любое время руководитель может распечатать протокол работы смены, а также каждого оператора в отдельности по количеству введенных портфелей и ошибок допустимых каждым оператором, общее количество портфелей на определенную дату, количество ввода переписных бланков, количество ошибочных портфелей, количество откорректированных портфелей. В процессе работы операторов можно было получить

норму выработки на одного человека, число верных и ошибочных листов, всю информацию по портфелям по данной территории.

Распространение данных, получение выходных таблиц, получение тематических карт

О завершении переписи можно говорить только тогда, когда ее итоги будут разработаны и представлены по полной Программе разработки органам государственного управления и другим категориям пользователей. Поэтому подготовка данных для публикации и распространение являются важным этапом деятельности, которую необходимо рассматривать как продолжение самой переписи. Публикации и распространению результатов переписи должен быть предоставлен такой же приоритет, как и сбору данных с тем, чтобы избежать длительных задержек в предоставлении данных.

Использование карт уже давно стало неотъемлемой частью процесса проведения переписи. Одним из видов новых технологий является применение в переписях географических информационных систем (ГИС). Помимо возможности создания цифровых карт для подготовки и проведения переписи, также появилась возможность создания тематических карт, представляющих результаты переписи, поэтому ГИС в настоящее время играет важную роль в распространении переписи и анализе данных.

Выбранное для распространения результатов переписи техническое и программное обеспечение должно работать с большими объемами данных и подтвердить свою высокую надежность и стабильность при работе с аналогичными по объему и нагрузке приложениями. В этих целях предполагалось приобретение следующего программного обеспечения для создания и ведения базы данных ГИС :

- *MapInfo Professional* – полнофункциональная геоинформационная система.
- *MapBasic* - язык разработчика для создания приложений или внедрения картографических возможностей в соответствующие приложения на базе ГИС MapInfo.
- *MapInfo SpatialWare* - технология управления пространственной информацией в БД SQL Server.
- *MapXtreme 2004 Web Deployment* - программное обеспечение для разработки картографических приложений для Intranet или Internet.
- *MapInfo MapX 5.0* - библиотека разработчика геоинформационных приложений.

В июле этого года НСК сделал запрос в Фонд народонаселения ООН в Кыргызской Республике об оказании технической помощи для создания и ведения базы данных ГИС. Недавно был получен ответ на этот запрос, в котором было предложено заменить пакет MapInfo на эквивалентное программное обеспечение ArcGIS (русская версия).

ГВЦ НСК уже имеет определенный опыт в области применения ГИС-технологий. В рамках работ по внедрению ГИС-технологий в систему статистики на ГВЦ НСК в 2004 году создан сектор ГИС. Разработан Технорабочий проект ГИС, который определяет первоочередные задачи, необходимые условия, долгосрочные перспективы его реализации. Специалисты сектора приняли участие в тренинге по MAPINFO. Для работы был предоставлен лицензионный базовый модуль MAPINFO, позволяющий оцифровывать карты, вести базу данных ГИС и делать тематические карты. Был осуществлен импорт картографических данных, созданных для обработки данных Переписи населения в 1999 году, из пакета Microstation в формат MapInfo и их актуализация в соответствии с произошедшими за этот период изменениями. Был проведен анализ отраслевых баз данных отделов НСК, информационных потоков, потребностей пользователей-статистиков в картографических продуктах. По результатам анализа создается База Данных ГИС, включающая областной, районный уровни и уровень айыл кенешей, а также основные города, населенные пункты и села.

Вместе с тем ГВЦ Нацстаткомитета КР существует ряд проблем в области внедрения ГИС:

- отсутствие лицензионного программного обеспечения ГИС (кроме лицензионного базового модуля MAPINFO);
- недостаточная оснащенность необходимым техническим оборудованием;
- недостаток специалистов, имеющих достаточную квалификацию в области ГИС-технологий.