

Применение новых технологий при составлении географических списков для переписи населения

Выбор тем в международных переписях населения¹

Выпущено в декабре 2015 года

ВВЕДЕНИЕ

Данный документ серии «Выбор тем в международных переписях населения» содержит сведения для национальных статистических служб (NSO, в соответствии с английским акронимом) о появившихся за последние десять лет технологиях составления географических списков для переписи населения. Географические списки имеют важное значение для хода последующей переписи. В рамках составления списков сотрудники переписи идентифицируют жилые единицы и составляют список домохозяйств на территории переписи, то есть, как правило, — всей страны. При традиционном выполнении такого подсчета с помощью бумаги и карандаша сотрудники рисовали пиктографические (эскизные) карты местности, чтобы затем пользоваться ими в рамках полномасштабной переписи населения и жилищного фонда. Ко времени проведения раунда переписи 2000 года широко доступным стало программное обеспечение географических информационных систем (GIS, в соответствии с английским акронимом) для настольных компьютеров с удобным графическим интерфейсом пользователя. Однако лишь ко времени проведения раунда переписи населения 2010 года (2005–2014 гг.) многие службы NSO перешли от бумажных карт к цифровым. Такой переход требует оцифровки границ территории учета, зафиксированных на эскизных картах. При этом для правильной оцифровки на эскизных картах должны быть также изображены физические границы. Для восстановления границ территории переписи используются физические объекты местности в цифровом формате, соответствующие изображенным на эскизных картах. В ходе раунда переписи 2010 года во многих службах NSO также стали изучать и внедрять решения на базе корпоративных GIS, средств анализа спутниковых снимков и карманных устройств с поддержкой глобальной навигационной системы (GPS, в соответствии с английским

Географические определения, встречающиеся в обзоре

В каждой NSO могут быть свои толкования терминов, используемых при описании работ по составлению списков. Приведенные определения призваны обеспечить ясность для целей данного обзора, но не заменить определения конкретной NSO или указать на потребность полного единообразия. Следует также отметить существование отличий между жилищами с учетом имеющихся в них активов и услуг. Дополнительные сведения о данных определениях см. на страницах 192–197 документа *Principles and Recommendations for Population and Housing Censuses* (UNSD, 2015).

Точки застройки — полный перечень точек местоположения строений на территории переписи (обычно соответствует территории всей страны), которые могут содержать или не содержать одно или несколько жилых помещений.

Жилое помещение — любое населенное строение, которое в период проведения переписи не используется целиком в нежилых целях.

Жилая единица — комната или группа комнат (место жительства) в составе капитального строения, сооруженного специально для размещения одного домохозяйства, и имеющего выход на улицу или пространство общего пользования.

Продолжение на следующей странице

акронимом). Однако большие затраты и высокий уровень сложности решений стали серьезными препятствиями на пути их широкого внедрения. В ходе раунда переписи 2020 года службы NSO будут применять такие технологии для осуществления переписи намного шире.

¹ Настоящая техническая записка является одной из серии «Избранные темы международных переписей населения», в которой рассматриваются вопросы, представляющие интерес для международного статистического сообщества. Бюро переписи населения США помогает странам совершенствовать национальные системы статистики, содействуя устойчивому расширению статистических компетенций.

Начало на предыдущей странице

Нетипичная жилая единица — комната или группа комнат (место жительства), которая ввиду своих конструктивных особенностей 1) не была рассчитана на длительное использование, 2) была рассчитана на передвижение, 3) была сооружена не для целей проживания людей или 4) не имеет услуг, достаточных, чтобы считать помещения жилыми, однако они используются в качестве таковых.

Жилищные единицы — все типичные и нетипичные жилые единицы, то есть все отдельные и независимые места жительства, занятые в период проведения переписи.

Домохозяйство — группа людей, которые совместно принимают решения, а также вместе осуществляют планирование снабжением пищей и проводят соответствующую подготовку.

Территория переписи — операционная единица для сбора данных во время переписи населения, соответствующая самому нижнему уровню географической иерархии административных и статистических единиц.

АНАЛИЗ СПУТНИКОВЫХ СНИМКОВ

Эскизные карты могут содержать точки застройки, однако их оцифровка обычно не позволяет обеспечить приемлемую точность определения местоположения. Для точной оцифровки сотрудники NSO могут с помощью спутниковых снимков фиксировать точки застройки на настольном компьютере, а для проверки на месте можно воспользоваться карманными устройствами с GPS. Совместное применение этих двух технологий позволяет уменьшить объем ресурсов, необходимый для составления списка зданий. Такая методика не только дает возможность быстрой фиксации точек застройки, но и служит обходным решением при отсутствии системы адресов. Однако чтобы отличать жилые помещения от нежилых и регистрировать индивидуальные жилищные единицы, по-прежнему необходима работа на местах.

Сотрудники NSO могут регистрировать точки застройки на изображениях спутниковой или аэрофотосъемки с применением методов интерпретации снимков и широко доступного программного обеспечения GIS для настольных компьютеров. При этом возможно использование двух методов: 1) фиксируются все возможные строения или 2) сотрудники выполняют визуальную интерпретацию изображений, исключая нежилые объекты. Например, объект с сильной отражающей способностью, имеющий четкие линейные контуры, может быть зданием. Однако если он не отбрасывает тени, скорее всего, это сухая уплотненная почва. Такая процедура интерпретации чревата ошибками, ее результативность зависит от опыта и обучения специалистов по оцифровке (см. рисунок 1). Поэтому можно пользоваться иным методом — в рамках оцифровки фиксируются все возможные объекты, затем после проверки на местности формируется окончательный набор данных.

Рисунок 1.

Пример интерпретации фотоснимка



Часть изображенных объектов похожи на здания, но многие из них не являются. При интерпретации снимков приходится полагаться на визуальную идентификацию и интуицию.

Источник: U.S. Geological Survey

Анализ снимков также можно автоматизировать с помощью средств, доступных в составе программного обеспечения GIS. Функция классификации изображений, например, анализирует спутниковые снимки, классифицируя их на основе общих спектральных или пространственных характеристик. Данный метод полезен для классификации точек застройки и идентификации новонаселенных областей, появившихся со времени предыдущей переписи населения. Однако для автоматической классификации снимки должны быть сделаны с помощью многоспектральной съемки. При такой съемке фиксируется несколько участков электромагнитного спектра, что позволяет анализировать снимки для классификации. Карты из онлайн-сервисов наподобие Google, Microsoft или Esri классифицировать невозможно, так как эти изображения не содержат спектральных данных. Если NSO пользуется онлайн-изображениями, единственный возможный вариант — ручная классификация. Классификация снимков позволяет формировать некоторые наборы данных, например, о дорожных сетях и новонаселенных территориях, но проверить сведения о жилых помещениях можно только путем работы на местах.

Методы классификации снимков можно разделить на две широкие категории — пиксельные и объектные. Пиксельная классификация подразумевает присвоение каждому пикселу (минимальной единице данных изображения) категории с учетом характеристик отражения света. Объектная классификация основана на извлечении объектов — элементов ландшафта (зданий, дорог, полей, оград) с учетом их формы и характеристик отражения света. Средства объектной классификации обычно не включаются в стандартное программное обеспечение GIS — их необходимо приобретать или разрабатывать

Рисунок 2.

Методы классификации снимков и необходимые навыки персонала

	Почвенно-растительный покров Методы доступны опытным пользователям GIS. При изменении программного обеспечения могут возникнуть дополнительные потребности в обучении.	Земледелие и землепользование Руководитель проекта должен пройти специальное обучение и иметь опыт исследований с использованием дистанционного зондирования.
Пиксельная классификация Возможна с использованием стандартного программного обеспечения GIS.	Обнаружение изменений — новых поселений и городских кластеров.	Обнаружение изменений территорий в пределах существующих поселений и прилегающих к ним; приблизительное указание количества жилищных единиц.
Объектная классификация Требует специализированного программного обеспечения.	Обнаружение изменений — новых зданий и физических объектов (дорог, оград, линий электропередачи).	Возможность построения карт для каждого здания с приблизительным указанием числа жилых помещений и жилищных единиц.

Источник: U.S. Census Bureau

отдельно. Персоналу NSO также потребуется дополнительное обучение для использования программного обеспечения классификации объектов.

Существует также отличие между классификацией почвенно-растительного покрова и классификацией землепользования. Классификация для определения почвенно-растительного покрова — наработанный метод геопространственного анализа, реализованный в большинстве пакетов GIS и программных систем для анализа снимков. Профессиональных пользователей GIS обычно можно научить эффективному выполнению анализа почвенно-растительного покрова. Классификация землепользования — более сложный и менее устоявшийся процесс, требующий дополнительных геопространственных слоев, таких как карты зонирования и кадастровые карты, которые в отдельных странах могут отсутствовать. Исследования землепользования требуют специализированного обучения — в NSO прошедших такое обучение сотрудников обычно нет (рисунок 2).

При создании геопространственных данных и работе с ними можно пользоваться современными технологиями классификации изображений. Однако соответствующее программное обеспечение вряд ли отменит в ближайшем будущем необходимость в занимающей много времени ручной оцифровке, которая выполняется при использовании снимков для составления списка адресов.

СБОР ДАННЫХ GIS НА МЕСТАХ С ПОМОЩЬЮ КАРМАННЫХ УСТРОЙСТВ

Карманные устройства используются для составления географических списков еще с 1990-х² годов, когда стала широко доступной портативная техника с функциями GPS. Применение таких устройств способствует повышению точности цифровых данных GIS.

Современные специализированные устройства с GPS позволяют регистрировать географические данные и дополнительные атрибуты. Однако они не всегда отвечают сложным требованиям, предъявляемым к подобной технике при составлении географических списков для переписи населения, и могут, соответственно, нуждаться в значительной доработке. Кроме того, с помощью карманных устройств с GPS можно фиксировать лишь доступные объекты. Фиксация протяженных объектов, таких как реки и дороги, может быть трудной, если местный персонал не имеет возможности легко вдоль них перемещаться. Регистрация границ с помощью карманных устройств с GPS сопряжена с дополнительными сложностями, если на местности отсутствуют четкие ориентировочные знаки.

Для одновременной фиксации пространственных и демографических данных на карманном устройстве должны быть специализированные приложения.

С появлением планшетов и смартфонов, поддерживающих технологию GPS, стала доступной возможность одновременно собирать географические данные и сложные атрибуты, а также создавать или корректировать географические объекты на местах. Использование карманных устройств для составления списка адресов и корректировки границ предполагает применение двух связанных, но самостоятельных технологий: 1) анкеты для составления списка адресов в варианте для личного опроса с применением компьютера (CAPI, в соответствии с английским акронимом) и 2) пользовательского интерфейса и ядра обработки операций с пространственными данными.

Специализированные мобильные приложения для составления географических списков позволяют выполнять операции со всеми видами географических примитивов (точками, линиями и многоугольниками) на экране и снабжены базовой картой (на основе спутниковых снимков или опорной картой), которая выполняет роль справочного руководства для сотрудников на местах. Такие приложения

² Первый коммерчески доступный приемник сигналов GPS, Magellan Nav 1000, был выпущен в мае 1989 года.

Рисунок 3.

Маршрут обмена данными между техническими средствами для составления географических списков



Источник: U.S. Census Bureau

позволяют собирать данные для разбиения наборов данных о точках застройки на индивидуальные жилые помещения, фиксировать количество жилищных единиц и собирать данные на уровне домохозяйства.

С помощью карманных устройств можно относительно легко собирать точечные данные и выполнять операции с ними. Однако работы по корректировке границ на местах требуют значительного усложнения самой программы составления списков и процесса реинтеграции собранных на местах данных в географическую иерархию. Реинтеграция линейных данных или многоугольников, которые корректировались на местах, требует сложного рабочего процесса с мерами по сохранению топологии и иерархии границ с вложениями. Кроме того, если такие данные одновременно редактируются несколькими пользователями, возможно, потребуются покупка решения GIS корпоративного класса (см. следующий раздел) для эффективного управления потоком данных. На рисунке 3 оранжевым выделены элементы, необходимые, когда приложение для составления списков включает средства корректировки статистических границ на местах. Более простое решение, которое можно рассмотреть в NSO — использование актуальных снимков высокого разрешения для большей части объема работ по корректировке границ территории переписи.

К числу важных факторов, которые надо учесть при внедрении карманных устройств для переписи и опросов, относятся характеристики точности определения местонахождения устройства и операционной системы. Планшеты и смартфоны в первую очередь рассчитаны на

использование потребителями в качестве коммуникационных устройств, и точность определения местоположения у них может быть низкой. Национальной статистической службе следует требовать от поставщиков устройств предоставления характеристик их точности и выполнять местное тестирование в различных условиях (в городах, пригородах и сельской местности). В NSO также необходимо соотнести потребности службы с возможностями программного обеспечения, доступного для различных операционных систем карманных устройств, таких как Android, iOS и Windows.

ПАКЕТЫ GIS И СИСТЕМЫ УПРАВЛЕНИЯ ПРОСТРАНСТВЕННЫМИ ДАННЫМИ КОРПОРАТИВНОГО КЛАССА

GIS корпоративного класса для составления географических списков рекомендуется внедрить на ранней стадии планирования программы переписи населения. Здесь «корпоративного класса» означает систему, рассчитанную на применение в офисе или организации. GIS для предприятий обеспечивает легкую маршрутизацию данных между различными отделами и индивидуальными пользователями без компрометации качества и безопасности данных.

Корпоративная база пространственных

GIS корпоративного класса обеспечивает легкость маршрутизации данных между различными отделами и индивидуальными пользователями без принесения в жертву качества и защиты данных.

Рисунок 4.

Основные отличия свободного программного обеспечения с открытым кодом (FOSS) и коммерческого программного обеспечения

	Свободное программное обеспечение с открытым кодом	Коммерческое программное обеспечение
Оплата лицензий	Отсутствует.	Возможны авансовые платежи и/или ежегодная оплата за сопровождение.
Исходный код	Полный доступ и, соответственно, более широкие возможности модификации для разработчиков программного обеспечения. Не все свободное программное обеспечение имеет открытый код.	Не является открытым для общественности и защищено авторским правом.
Легкость использования	Может иметь только интерфейс командной строки и требовать знаний в области программирования: необходима более высокая квалификация пользователя.	Обычно имеет удобный графический интерфейс, доступный менее опытным пользователям.
Техническая поддержка	Только со стороны Интернет-сообщества пользователей; профессиональные услуги поддержки могут предоставляться на платной основе частными структурами.	Обычно предоставляется непосредственно издателем программного обеспечения.

Источник: U.S. Census Bureau.

данных (геоданных) позволяет хранить пространственные данные и работать с ними. Обычно она администрируется с помощью системы управления реляционными базами данных (RDBMS, в соответствии с английским акронимом). Такая база может быть доступной конкретной рабочей группе (например, специалистам по работе с GIS или в пределах всей NSO — возможно, через защищенный Интернет-портал, в зависимости от потребностей в рамках работ по составлению списков.

Службе NSO необходимо сделать выбор между проприетарным коммерческим программным обеспечением и свободным программным обеспечением с открытым кодом (FOSS, в соответствии с английским акронимом) для использования в основе корпоративной GIS для составления географических списков. Как показано на рисунке 4, у обоих вариантов есть свои преимущества и недостатки. Службе NSO следует обсудить выбор программного обеспечения с сотрудниками (если разработка осуществляется собственными силами) или поставщиком

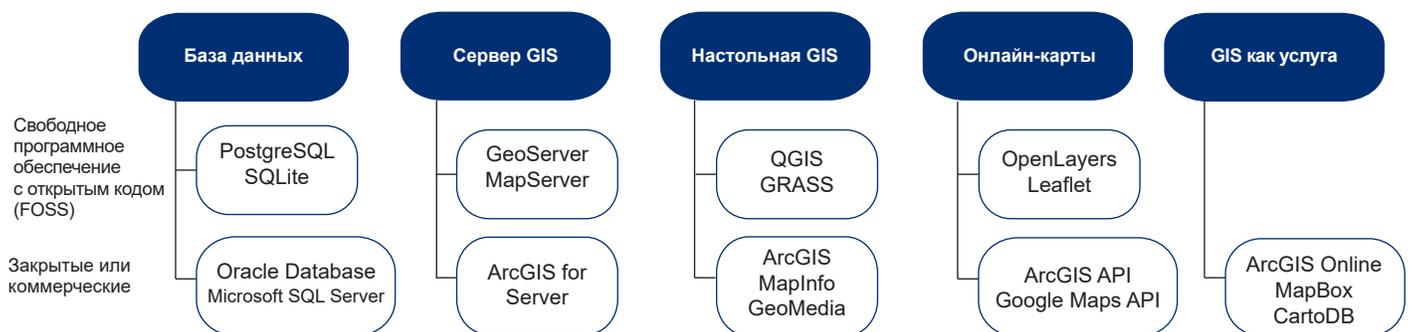
(в случае аутсорсинга) на раннем этапе планирования переписи. Независимо от того, выбрано программное обеспечение с открытым кодом или коммерческое, службе NSO необходимо позаботиться о том, чтобы созданная корпоративная GIS была доступной для использования как на протяжении всего цикла опроса или переписи, так и после его завершения.

Открытое и коммерческое программное обеспечение можно использовать совместно для различных компонентов корпоративной GIS. Многие из таких компонентов являются совместимыми, то есть NSO может воспользоваться решением с открытым кодом для одного из компонентов и коммерческим для другого, в зависимости от требований рабочих процессов. На рисунке 5 перечислены примеры популярного открытого и коммерческого программного обеспечения GIS.

Такие решения обычно размещаются службой NSO на локальных серверах или в облачных сервисах, а их

Рисунок 5.

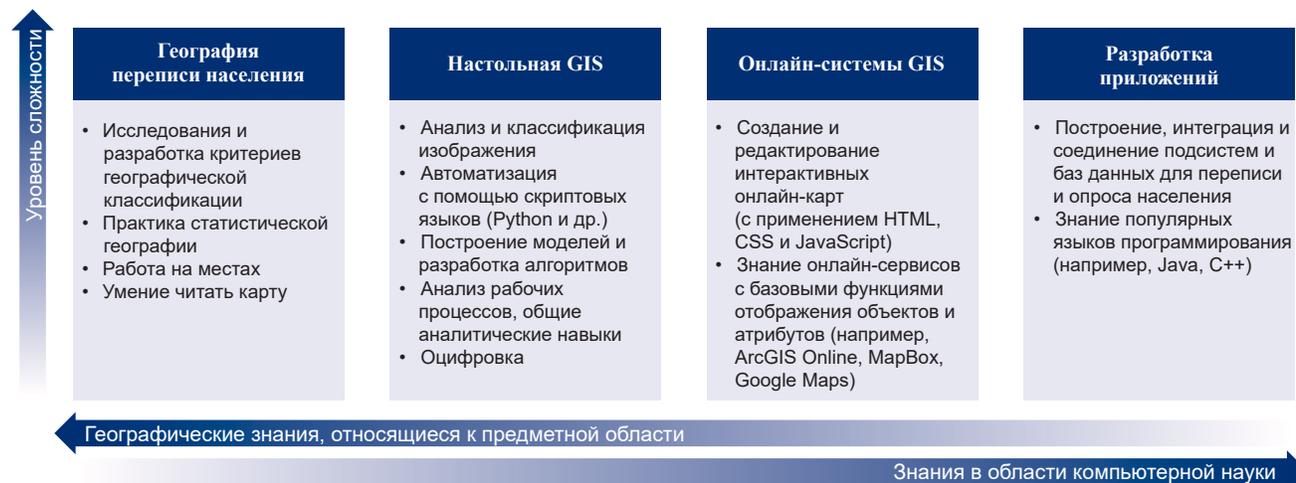
Примеры компонентов программного обеспечения GIS корпоративного класса



Источник: U.S. Census Bureau.

Рисунок 6.

Области знаний, необходимые для успешного внедрения новой технологии GIS



Источник: U.S. Census Bureau.

управлением занимается сама служба либо подрядчик. На рисунке 5 также перечислены альтернативные решения нового типа, относящиеся к категории «GIS как услуга» (GIS-as-a-Service). Такие решения сочетают функции баз данных, серверов и онлайн-карт и могут обладать различными наборами возможностей. Решение GIS-as-a-Service потенциально позволяет уменьшить число специалистов по работе с геопространственными данными, сократив соответствующие расходы. Однако системы GIS-as-a-Service предоставляют меньше возможностей контроля и произвольной настройки по сравнению с локальным серверным решением, в связи с чем их лучше использовать не на стадии проведения переписи, а на этапе обнародования результатов.

ОЦЕНКА И РАСШИРЕНИЕ НАВЫКОВ ПЕРСОНАЛА ПО РАБОТЕ С GIS

Новые геопространственные технологии, представленные в данном обзоре, требуют переоценки технических навыков, которыми владеют сотрудники по работе с GIS службы NSO.

Традиционно наборы навыков географов и картографов перекрывались: первые отвечали за разработку классификации географических районов и анализ пространственных демографических тенденций, а вторые создавали карты, используемые на местах и в обнародуемых материалах. Рост применения настольных GIS привел к объединению двух этих специальностей, а затраты времени и труда на создание карт уменьшились.

Однако набор навыков, необходимых для работы с современными геопространственными технологиями, изменился (рисунок 6). Сотрудники должны уметь автоматизировать работу с GIS с помощью скриптовых языков, проектировать базы данных для хранения геопространственных данных и создавать интерактивные онлайн-карты. Нанять одного специалиста, обладающего всеми необходимыми навыками использования новых технологий, непросто, однако не всем

Для обеспечения продуктивной работы руководителям по GIS необходимо разделить деятельность на несколько направлений, требующих разных навыков, поощряя при этом обмен информацией, сотрудничество и взаимодействие между разными направлениями.

сотрудникам по работе с GIS статистической службы нужно становиться экспертами во всех перечисленных областях. Поэтому распределение обязанностей персонала по работе с GIS требует тщательного планирования с учетом круга навыков, необходимого для достижения каждой конкретной цели. Для обеспечения продуктивной работы руководителям по GIS необходимо разделить деятельность на несколько направлений, требующих разных навыков, поощряя при этом обмен информацией, сотрудничество и взаимодействие между разными направлениями. Руководителям также необходимо побуждать сотрудников по работе с GIS к самостоятельному решению проблем и обучению, учитывая стремительное развитие и высокий уровень технической сложности современных GIS.

АУТСОРСИНГ

После оценки способностей сотрудников NSO по работе с GIS, возможно, выяснится нецелесообразность развития собственных навыков, требуемых для создания, интеграции и развертывания полноценной системы построения географических списков. В этом случае NSO может заключить договор аутсорсинга разработки такой системы с частным подрядчиком. Главной целью аутсорсинга должно быть временное получение доступа к навыкам, отсутствующим в самой NSO, или увеличение численности сотрудников

Ответственность за итоговый успех или провал процесса составления списков должна лежать на NSO, а не на подрядчике.

с определенным набором навыков. В случае аутсорсинга следует придерживаться следующих рекомендаций:

1. Не уступайте полностью контроль над проектированием и разработкой системы подрядчику. Ответственность за итоговый успех или провал процесса составления списков должна лежать на NSO, а не на подрядчике. Соответственно, в NSO должно быть четкое представление о внедряемом техническом решении и соответствующих рисках.
2. Используйте практические знания сотрудников по работе с GIS, имеющих опыт корректировки статистических границ и составления списка адресов.
3. Документируйте рабочий процесс обновления карт переписи до начала применения новых геопространственных технологий и используйте эту документацию при разработке системы составления списков адресов вместе с подрядчиком.
4. Не допускайте того, чтобы технологии определяли архитектуру системы составления списков.

При выборе подрядчика в NSO должны учитывать потребности дальнейшего сопровождения и расширения системы, а также навыки персонала. Поставщики нередко специализируются либо на коммерческих системах, либо на системах с открытым кодом.

ДРУГИЕ ВАЖНЫЕ ФАКТОРЫ

Внедрение новых технологий связано с риском. Сотрудники могут не приветствовать изменение привычных процедур работы и сопротивляться внедрению новых технологий. Кроме того, новые технологии требуют инвестиций в защиту данных и обучение персонала в целях предотвращения потери конфиденциальных личных сведений.

Перед внедрением новые технологии необходимо подвергнуть тщательному тестированию. На тестирование нужно отвести время в графике проекта с расчетом на то, чтобы успеть внести доработки перед внедрением. На случай, если тестирование покажет, что программное решение еще не будет соответствовать необходимым требованиям ко времени начала работ, нужно предусмотреть альтернативный план, который позволит успешно завершить работы.

Сотрудники по работе с GIS, их руководители и высшее руководство играют критически важную роль в успехе внедрения новых технологий в NSO и несут общую ответственность за такой успех.

ЛИТЕРАТУРА

Jensen, John R., *Remote Sensing of the Environment: An Earth Resource Perspective*, Second Edition, Prentice Hall, Upper Saddle River, NJ, 2006.

_____, *Introductory Digital Image Processing*, Third Edition, Prentice Hall, Upper Saddle River, NJ, 2005.

Longley, Paul A., Michael F. Goodchild, David J. Maguire, and David W. Rhind, *Geographic Information Systems and Science*, Third Edition, Wiley, Hoboken, NJ, 2011.

United Nations Statistics Division, *Principles and Recommendations for Population and Housing Censuses*, Revision 3, United Nations Publications, New York, 2015.

_____, *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations Publications, New York, 2009.