

Глава IV

Офисная обработка — хранение информации и ведение записей: некоторые соображения относительно полей данных¹

Хелен Керфут (Канада)

Одной из основных обязанностей национального официального органа по географическим названиям является его деятельность в качестве координатора этого процесса. Совету необходима вспомогательная служба для обработки всей информации о названиях, ведения подробных записей о текущей деятельности и обеспечения доступности данных для широкого распространения в государственных органах и среди населения в целом.

Ранее записи в основном хранились в огромных гроссбухах, на карточках и в разного рода рукописных документах. В настоящее время гораздо больше внимания уделяется электронной обработке текстовых данных, цифровым базам данных и информации, доступной для населения на веб-сайтах.

Изменился не столько тип информации, необходимой для записи географических названий, сколько методы обработки, хранения и распространения информации, которые видоизменяются по мере развития технологий. Тем не менее такие основные качества при ведении записей, как точность и тщательность, важны не меньше, чем прежде. Полезность записей и базы данных напрямую зависит от качества введенной информации. Всегда лучше вносить записи медленно, но правильно с первого раза. Последующее исправление записей часто оказывается сложной и трудоемкой задачей, поскольку ошибки могут остаться незамеченными. Распространение ошибок способно привести к созданию ненадежного набора данных сомнительной ценности.

Хранение информации и ведение записей можно рассматривать как часть «офисной обработки» географических названий, которая в широком плане состоит из следующих этапов:

- a) сбор топонимической информации;
- b) утверждение топонимов;
- c) хранение информации и ведение записей;
- d) распространение информации.

Как правило, в настоящее время записи ведутся в определенном формате базы данных в компьютерной операционной среде. Для временного хранения данных о географических названиях может подойти электронная таблица (типа той, что применяется для

¹ На основе отрывка из Helen Kerfoot, «Office processing of geographical names», *Lecture Notes*, vol. 1, Enschede, Netherlands/Frankfurt am Main, Germany, 10–24 August 2002, pp. 83–85, для учебного курса по топонимике, организованного Отделом голландско- и германоязычных стран/Группой экспертов Организации Объединенных Наций по географическим названиям, отредактировано в Утрехтском университете, Утрехт, Нидерланды.

бухгалтерских и финансовых целей), но использовать ее для постоянной базы текстовых топонимических данных не рекомендуется. Могут применяться более сложные системы, например, с базой данных, которая ведется в Интернете с непосредственным официально разрешенным региональным вводом данных.

Цифровые базы данных можно создавать на основе существующих рукописных данных на карточках, списков различных типов, путем оцифровывания карт в одном или более масштабах или путем введения вновь собранных данных на местах. В процессе введения записей с карточек потребуются принятие многочисленных решений, так как записи вряд ли велись так строго, как это необходимо для заполнения полей цифровой базы данных. Даже если на промежуточном этапе записи по-прежнему придется вести на карточках, нужно будет принять ряд решений относительно того, какие поля следует собирать и хранить для каждого названия.

Поля данных

Поля данных для каждого географического названия будут значительно отличаться в разных странах в соответствии с потребностями. Тем не менее существует несколько ключевых необходимых полей данных. Приведенный ниже список можно, конечно, видоизменить с учетом индивидуальных потребностей, но он базируется в основном на содержании баз данных, информация о которых поступила в Группу экспертов Организации Объединенных Наций по географическим названиям.

1. Топоним

Стандартизированные названия следует вводить с правильным употреблением заглавных букв, дефисами, диакритиками и т. д. Аббревиатуры следует вводить, только если они являются стандартизированной частью названия (например, St. John's).

- a) Естественный порядок слов, как дается на карте или в тексте (например, Lake Phillip);
- b) обратный порядок слов, когда это требуется для алфавитного справочного перечня (например, Phillip, Lake). Названия населенных пунктов (например, Harbour Grace) не вносятся в обратном порядке.

2. Тип объекта

Например: река, гора, населенный пункт и т. д. (или более подробная разбивка).

Тип объекта очень важен, особенно там, где существует неоднозначность. Например, Baker Lake и Mount Pleasant — это физико-географические объекты или населенные пункты?

3. Административные единицы, в которых расположен данный объект

- a) Первый уровень административной единицы ниже уровня страны как таковой (например, провинция, штат, район и т. д.);
- b) второй уровень (и третий уровень также может оказаться полезным).

Если такой объект, как река, пересекает административные границы, в записи названия необходимо указать более одной административной единицы.

4. Географические координаты

- a) Широта (градусы, минуты, секунды: две цифры для каждого);

- b) долгота (градусы, минуты, секунды: три цифры для градусов, две цифры для минут и две цифры для секунд).

Даже если используются другие системы координат, математический алгоритм должен позволять переводить их в системы, которые можно было бы объединить с данными из других стран. Кроме того, если существует возможность путаницы в названиях в пределах одной страны, скорее всего, потребуется указывать расположение к северу или югу от экватора и к востоку или западу от нулевого меридиана.

Для использования данных за пределами страны необходимо будет соответствующим образом добавлять указатель N (север), S (юг), E (восток) или W (запад). В следующих примерах указаны стандарты для применения в географической информационной системе (ГИС) и при обмене данными:

Широта, к северу от 0 градусов	напр.	30
Широта, к югу от 0 градусов	напр.	-30
Долгота, к востоку от 0 градусов	напр.	120
Долгота, к западу от 0 градусов	напр.	-120

Координаты должны считываться с максимально возможной точностью с крупномасштабных карт или, возможно, из Глобальной навигационной спутниковой системы (GPS).

Даже если пределы объекта определяются на цифровой графической системе, полезно иметь один комплект справочных координат. Для оконтуренных объектов (например, озер) они выбираются по центру; для текущих водных объектов используется устье (а координаты истока можно записать в качестве вторичного значения); для городских районов записываемым значением обычно является центр города. Выбор координат для некоторых объектов, таких как ледники, каналы, множественные объекты (например, группа островов, двуглавые вершины), требует специального рассмотрения.

5. Листы карт

В зависимости от координат объект может быть представлен на одном листе карты. Он может также располагаться на нескольких листах карты (например, Mackenzie River). Отдельные поля могут использоваться для различных масштабов (например, 1:50 000; 1:250 000) в зависимости от охвата карты.

6. Варианты названий

Это поле предназначено для других названий, для которых следует сделать перекрестные с основным топонимом ссылки, например исторические названия, другие языковые формы, другие варианты написания.

7. Статус

Как минимум, может быть указано, имеет ли название утвержденный статус, имело ли оно утвержденный статус в прошлом, либо оно не утверждено. Это поле можно расширить, чтобы указать многие другие уровни информации (например, изменение названия, одно из многочисленных официальных названий, объявленное советом недействительным). Это же поле или подобное можно было бы использовать также для того, чтобы показать, какого этапа обработки достигло это название (например, готово для передачи в совет, ожидается подпись управляющего департаментом).

8. Дата утверждения

Год, месяц, день, когда данное название приобрело статус утвержденного. (Если дата рекомендации совета отличается от даты фактического утверждения, эту информацию можно добавить в другом поле).

9. Идентификатор записи

В цифровой системе каждая запись приобретает собственную уникальную идентификацию — скорее всего числовое поле, хотя буквенные или буквенно-числовые поля тоже возможны. В строго топонимической базе данных уникальный идентификатор придается каждой записи названия. Поскольку названия могут дублироваться, само поле названия или его сочетание с другим полем не подходит в качестве идентификатора. (В системе ГИС объекты будут иметь уникальные идентификаторы, при этом записи названия станут атрибутами записей объектов. Можно также использовать и название, и идентификаторы объекта, когда системы взаимосвязаны).

Можно добавлять другие поля, чтобы показать, например, геодезическую характеристику используемых для карты объектов, язык, на котором дается название, его значение и происхождение, правовой статус, источники данных и т. д. Можно просмотреть более старые записи на карточках, сканировать и добавить их как поля информации для записи названия. Если за утверждение названия отвечает не один совет, возможно, понадобится поле с указанием одобрявшего название органа. В некоторые базы данных были включены описания расположения (например, «к северу от Пембины» или «течет на юг в Голубую реку»).

Некоторые топонимические базы данных могут включать другую информацию, такую как данные о населении и высоту над уровнем моря. Однако следует помнить о том, что клиентам нужна современная информация, а для того чтобы обеспечивать точность информации, за которую ты не несешь прямой ответственности, нужно больше средств. В современной цифровой среде, вероятно, лучше всего поддерживать связь с другими базами данных, ответственными за такую информацию.

В любой цифровой базе данных создаются другие поля, чтобы сделать возможной идентификацию оператора, который вводил записи или видоизменял их, а когда это уже сделано, объяснения любых кодов, использованных в различных полях данных, и т. д.

С точки зрения обмена информацией важно, чтобы база данных для того или иного языка (языков) была создана в соответствии со стандартами Международной организации по стандартизации (ИСО). Даже в этом случае могут встретиться символы, которые изобразить невозможно. Например, в атапасских языках существуют «трудные для построения» символы, которые в настоящее время невозможно ввести в национальную базу данных, следующую каким-либо национальным или международным стандартам. До тех пор, пока не появятся такие стандарты, для ввода подобных названий будут использоваться символы-заменители.

Что касается офисной обработки, сотрудники, занимающиеся вводом данных, должны иметь в своем распоряжении «Руководство по введению записей». В нем будет указано, как форматировать данные для ввода в различные поля или их модификации. Чем качественнее инструкции и чем лучше сотрудники знакомы с базой данных, тем точнее будут сделанные записи, а ведь самую важную роль играют именно записи, которые будут доступны и будут использоваться неоднократно.

Помимо буквенно-числовых данных о топонимах важно, чтобы размеры названных объектов (то есть применение названия к ландшафту) были представлены в пространственной (графической) форме. Это может быть сделано на справочных экземплярах отпечатанных на бумаге карт или — в современном мире, где все большую роль играет цифровая информация, — границы можно было бы включать в цифровые файлы карт.