

Использование метамодели в качестве основы для описания пространственных данных

Т.В. Шушкевич

Южный федеральный университет, Таганрог

Аннотация: В статье рассматриваются вопросы представления пространственных данных в рамках метамодели и оценки качества данных. Приводятся дескрипторы описания для частного набора пространственных данных.

Ключевые слова: метамодель, качество пространственных данных, набор дескрипторов оценки, оценка качества.

Цифровые данные, полученные при решении той или иной задачи, зачастую могут использоваться повторно с той же или близкой к первоначальной цели. Например, они могут служить доказательной базой для принятия управленческих решений. Целесообразно такие данные накапливать и объединять в фонды, используя различные базы данных и управляющие системы к ним. При формировании фонда информационных данных на любом уровне всегда поднимается вопрос качества данных, составляющих основу фонда. Особенно актуален вопрос качества данных для Федеральных информационных фондов, служащих основой для принятия каких-либо юридических, правовых, экономических и т.п. решений.

Методы оценки качества данных в определенной степени всегда субъективны, поскольку различные потребители информации о качестве данных могут использовать набор данных или выборку из набора с разными целями, в разных условиях. Выборки могут подвергаться различной обработке и быть соответственно репрезентативными или не обладать этим свойством при решении конкретной задачи. Для согласования методов оценки качества данных область международного и национального технического регулирования предлагает стандарты, описывающие один из подходов к оценке качества данных.

Согласно принятым международному (ISO 19157:2013 Geographic information - Data quality) и национальному (ГОСТ Р 57773-2017 Пространственные данные. Качество данных) стандартам, качество пространственных данных может быть оценено с помощью стандартизованного набора мер качества данных. Стандарты рекомендуют использовать шесть элементов качества данных и предлагают формальные модели описания на языке UML.

Следует отметить, что всякая оценка качества данных зависит от цели, с которой она проводится, и от системы, в которую эти данные внесены [1, 2]. Вне системы использования оценить качество не представляется возможным, поэтому имеет смысл представить оценку качества данных как подсистему системы использования совокупности пространственных данных, в том числе данных, используемых в геоинформационных системах.

Как правило, говоря о качестве пространственных данных, подразумевают точность позиционирования объектов на картах или планах или точность определения координат объектов или характерных точек [3, 4], однако такая трактовка не позволяет учитывать цели использования пространственных данных и их наборов и является только частной задачей определения уровня качества. Такие аспекты, как полнота набора, удовлетворяющая требованиям потребителя данных, или возможное изменение интересующих потребителя характеристик во времени обычно в оценку качества данных не включаются.

В работах [5 – 8] изложен подход, позволяющий представить геоинформационную систему в виде метамодели в рамках метасистемного подхода (System of system, SoS) [9, 10]:

$$M = \{M_O, M_E, M_{OE}, M_D, M_{ГИС}, Q, M_{ГПР}, M_U, M_H, A\},$$

где: M_O – идентифицирующая модель системы (объекта), M_E – модель окружающей среды, M_{OE} – модель взаимодействия системы и среды, M_D –

модель поведения системы, реализуемая при имитационном моделировании возмущающих и управляющих воздействий; $M_{\text{ГИС}}$ – модель геоинформационной системы (ГИС); Q – возмущающие/управляющие воздействия, $M_{\text{ПР}}$ – модель принятия решений, M_U – модель управляющей системы, M_H – модель «наблюдателя» (инженера, эксперта, оператора), A – правила объединений моделей и выбора процессов изменения объекта.

Модель можно дополнить кортежем системы оценки, построенным на мерах указанных стандартов:

$$M_M = \langle DQ_C, DQ_LC, DQ_TA, DQ_TQ, DQ_PA, DQ_UE \rangle$$

где: DQ_C – подсистема мер, характеризующих полноту набора данных; DQ_LC – подсистема мер, характеризующих логическую согласованность; DQ_TA – подсистема мер, характеризующих тематическую точность; DQ_TQ – подсистема мер, характеризующих временное качество; DQ_PA – подсистема мер, характеризующих позиционную точность; DQ_UE – дескриптор возможности корректного использования данных.

Каждая из подсистем, в свою очередь, является кортежем. Набор мер, характеризующих полноту набора данных, включает меру избыточности данных (DQ_CC) и меру отсутствия элементов, наборов или структуры данных (DQ_CO). Для логической согласованности составляющими кортежа выступают меры концептуальной (DQ_CCon), доменной (DQ_DCon), структурной (DQ_FCon) и топологической (DQ_TCon) согласованности. Пример использования дескрипторов оценки для первых двух подсистем показан в таблице 1. Такой набор исходных данных характеризуют охранную зону распределительного подземного и надземного газопровода низкого давления. И хотя охранные зоны с 1.01.2016 не относятся к объектам землеустройства, сведения о них вносятся в Росреестр, следовательно, данные должны иметь определенный уровень качества, который можно оценить, например, по критериям, предлагаемым ISO 19157:2013 и ГОСТ Р 57773-2017.

В таблице 1 приведена часть стандартного набора данных, используемого для формирования текстового и графического описания охранной зоны. Цифровые обозначения дескрипторов соответствуют ГОСТ Р 57773-2017. Курсивом в таблице выделены классы и подклассы данных (классы – верхний уровень перечислений, подклассы – второй уровень перечислений). Для классов и подклассов набор мер может быть шире, чем для отдельных элементов, поскольку рассматривается совокупность частных критериев оценки и их взаимодействие или взаимосвязь.

Из приведенных наборов данных можно извлечь поднабор в соответствии с целью оценки – например, для подтверждения соответствия предметной области или для контроля правильности определения характерных точек границ объектов. Из подклассов данных могут формироваться новые наборы и классы данных, например, для графического отображения объекта на публичной карте можно извлечь сведения из набора о местоположении. Отдельные характеристики нескольких наборов данных (нескольких зон) могут использоваться для проверки согласованности расположения зон и границ объектов. Выбор поднабора может осуществлять непосредственно пользователь или же он может быть фиксирован и формироваться при выполнении определенного стандартизованного пользовательского запроса в системе, хранящей данные. Структура подкласса 3.3 в приведенной таблице полностью идентична структуре подкласса 3.2, включая перечень характеристик и дескрипторы оценки для класса и подкласса, поэтому для исключения дублирования в таблице не приводится.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-01-00412 А «Разработка моделей и алгоритмов когнитивного и нейроэволюционного моделирования для поддержки принятия решения в интеллектуальных информационно-управляющих системах».

Таблица № 1

Использование дескрипторов полноты и логической согласованности

Класс/элемент данных	Полнота набора DQ_C		Логическая согласованность данных DQ_LC			
	DQ_CC	DQ_CO	DQ_CCon	DQ_DCCon	DQ_FCon	DQ_TCon
<i>1. Сведения об объекте</i>	1 - 4	5 - 7		14 - 18	19, 20	
1.1. Местоположение объекта	1, 2, 3, 4	5, 6, 7		14 - 18	19, 20	
1.2. Площадь объекта, м ²		5, 6, 7			19, 20	
1.3. Погрешность определения площади, м ²		5, 6, 7		14 - 18		
1.4. Иные характеристики объекта		5, 6, 7				
<i>2. Сведения о местоположении границ объекта</i>	1 - 4	5 - 7	8 - 10, 12, 13	14 - 18	19, 20	
2.1. Используемая система координат	1 - 4	5 - 7	8, 9	14, 15	19	
<i>2.2. Сведения о характерных точках границ объекта:</i>	1 - 4	5 - 7	8, 9, 12, 13	14 - 18	19, 20	
2.2.1. Обозначение характерных точек границы	4	5, 6, 7				
2.2.2. Координаты характерных точек границ образуемых земельных участков X, Y; м	4	5, 6, 7	8 - 10, 12, 13	14 - 18	19, 20	
2.2.3. Метод определения координат	1 - 4	5, 6, 7	8 - 10	14 - 18	19, 20	
2.2.4. Средняя квадратическая погрешность положения характерной точки, м		5, 6, 7		14 - 18		
2.2.5. Описание закрепления точки		5, 6, 7				

Окончание таблицы 1

3. Сведения о местоположении измененных (уточненных) границ объекта	1 - 4	5 - 7	8 - 10, 12, 13	14 - 18	19, 20	
3.1. Система координат	1 - 4	5 - 7	8, 9	14, 15	19	
3.2. Сведения о характерных точках границ объекта:	1, 2, 3, 4	5, 6, 7	8, 9, 12, 13	14 - 18	19, 20	
3.2.1. Обозначение характерных точек границы	4	5, 6, 7				
3.2.2. Существующие координаты X, Y; м	4	5, 6, 7	8-10, 12, 13	14 - 18	19, 20	
3.2.3. Измененные (уточненные) координаты X, Y; м	4	5, 6, 7	8 - 10, 12, 13	14 - 18	19, 20	
3.2.4. Метод определения координат	1 - 4	5, 6, 7	8 - 10	14 - 18	19, 20	
3.2.5. Средняя квадратическая погрешность положения характерной точки, м		5, 6, 7		14 - 18		
3.2.6. Описание закрепления точки		5, 6, 7				
3.3. Сведения о характерных точках части (частей) границы объекта	1, 2, 3, 4	5, 6, 7	8, 9, 12, 13	14 - 18	19, 20	
3.4. Сведения о частях границ объекта, совпадающих с местоположением внешних границ природных объектов и (или) объектов искусственного происхождения	4	5 - 7				21- 26
3.4.1. Обозначение части границ от точки	4	5, 6, 7				
3.4.2. Обозначение части границ до точки	4	5, 6, 7				
3.4.3. Существующее описание прохождения части границ		5, 6, 7				
3.4.4. Измененное (уточненное) описание прохождения части границ		5, 6, 7				

Литература

1. Tsvetkov V.Ya., Shaitura S.V., Minitaeva A.M., Feoktistova V.M., Kozhaev Yu.P., Belyu L.P. Metamodelling in the information field // Amazonia Investiga. 2020. Т. 9. № 25. С. 395-402.
2. Дышленко С.Г. Развитие геоинформационного пространства // ИТНОУ: Информационные технологии в науке, образовании и управлении. 2019. № 1 (11). С. 46-52.
3. Хатоум Т. С. Критерии оценки точности геопространственных данных//ИнтерэкспоГео-Сибирь. Новосибирск, СГГА, 2007. С. 284-289.
4. Воронин А.В. Целевая функция точности определения геоданных//Вестник Воронежского государственного университета. Системный анализ и информационные технологии. 2019. № 1. С. 44-50.
5. Звягин Л.С. Метасистемный подход в экономике и управлении // Вопросы экономики и управления. 2016. №4. С. 6-11.
6. Ginis L.A., Gorelova G.V., Kolodenkova A.E. Cognitive and simulation modeling of development of regional economy system. International Journal of Economics and Financial Issues, vol. 6, no. 5S, 2016, pp. 97-103.
7. Гинис Л.А., Гордиенко Л.В., Левонюк С.В. Разработка концептуальной проблемно-ориентированной метамоделю образного представления сложной системы на основе геоинформационной системы// Инженерный вестник Дона, 2017, №1 URL: ivdon.ru/magazine/archive/n1y2017/4065/.
8. Гинис Л.А., Вовк С.П. Определение четко доминирующих тактик для выработки альтернативных управляющих решений в условиях полной неопределенности// Инженерный вестник Дона, 2014, №2 URL: ivdon.ru/magazine/archive/n2y2014/2327/.
9. James E. Campbell, Dennis E. Longsine, Donald Shirah, and Dennis J. Anderson, System of Systems Modeling and Analysis, Springfield, VA, 2005, 135 p.

10. Цветков В.Я., Булгаков С.В., Титов Е.К., Рогов И.Е. Метамоделирование в геоинформатике// Информация и космос. 2020. № 1. С. 112-119.

References

1. Tsvetkov V.Ya., Shaitura S.V., Minitaeva A.M., Feoktistova V.M., Kozhaev Yu.P., Belyu L.P. Amazonia Investiga. 2020. P. 9. № 25, pp. 395-402.
2. Dyshlenko S.G. ITNOU: Informacionnye tehnologii v nauke, obrazovanii i upravlenii. 2019, № 1 (11), pp. 46-52.
3. Khatoum T.S. InterekspoGeo-Sibir'. Novosibirsk, SGGGA, 2007, pp. 284-289.
4. Voronin A.V. Vestnik Voronezhskogo gosudarstvennogo universiteta. Sistemnyy analiz i informatsionnye tekhnologii. 2019. № 1. pp. 44-50.
5. Zvyagin L.S. Voprosy jekonomiki i upravlenija [Economic and management issues]. 2016. №4. pp. 6-11.
6. Ginis L.A., Gorelova G.V., Kolodenkova A.E. International Journal of Economics and Financial Issues, vol. 6, no. 5S, 2016, pp. 97-103.
7. Ginis L.A., Gordienko L.V., Levonyuk S.V. Inzhenernyj vestnik Dona (Rus), 2017, №1 URL: ivdon.ru/magazine/archive/n1y2017/4065/.
8. Ginis L.A., Vovk S.P. Inzhenernyj vestnik Dona (Rus), 2014, №2 URL: ivdon.ru/magazine/archive/n2y2014/2327/.
9. James E. Campbell, Dennis E. Longsine, Donald Shirah, and Dennis J. Anderson, System of Systems Modeling and Analysis, Springfield, VA, 2005, 135 p.
10. Tsvetkov V.Ya., Bulgakov S.V., Titov E.K., Rogov I.E., Informatsiya i kosmos. 2020. № 1, pp. 112-119.