

Организация геоданных

Статья раскрывает особенности организации геоданных как системы данных. Показаны основные этапы организации геоданных. Раскрываются особенности этапов и их взаимосвязь. Показано, что основой интеграции геоданных являются пространственные данные. Показано, что основой практического использования геоданных являются тематические данные. Показано, что геоданные являются новым информационным ресурсом.

Ключевые слова: данные, геоданные, изучение окружающего мира, система данных, информационные ресурсы.

S. A. Kudzh

Organization geodata

The article reveals the characteristics of the organization as a system of geo-data. The basic steps in organizing geodatabase. This article describes the steps of forming a geodatabase and their relationship. It is proved that the basis of the integration of geo-spatial data are. Shown that the basis of the practical use of geodata are thematic data. Article argues that geodata are new information resource.

Keywords: data, geodata, the study of the world, the system data, information resources.

Геоданными называют обобщенные данные о Земле, которые включают как основу пространственную информацию [1]. На эту основу нанизывают разнообразные данные для последующего пространственного, экономического, регионального и других видов анализа. Геоданные являются не просто данными, а представляют собой систему данных и информационный ресурс [2]. Одной из особенностей геоданных является то, что они отражают реально существующие пространственные отношения [3].

Множество геоданных собирается с помощью разных технологий и систем [4]. Данные отражают различные характеристики и свойства. Они могут иметь различные размерности, разное количество значащих цифр, разное число разрядов, разную точность и т.д. Собранные данные могут храниться в виде наборов или файлов. Кроме того, при сборе данные могут организовывать связанные совокупности, называемые моделями [5]. Для того чтобы разнородные данные и модели можно было обрабатывать в одной системе они должны быть упорядочены и сведены к единой информационной модели, в которой они будут дополнять друг друга.

Организацией данных (рис.1) называется процедура сведения разнородных данных и моделей в единую непротиворечивую информационную модель, которую в дальнейшем можно будет эффективно применять в различных

технологиях анализа и управления. Эту особую информационную модель называют информационной основой.

Результатом организации данных является создание только такой информационной модели, которая позволяет организовать хранение в базе данных [6]. Следовательно, организация геоданных дает возможности создания для БД и возможность их автоматизированной обработки.

Геоданные для их использования должны быть классифицированы, унифицированы, интегрированы и стратифицированы. Последовательность этих процедур показана на рис. 1.

Первым этапом является сбор информации. Он формирует так называемые первичные данные. Исходная первичная информация включает множество параметров, многие из которых дублируют друг друга. Уменьшение числа данных о реальных объектах достигается применением разных моделей, сохраняющих основные свойства объектов исследования и не содержащих второстепенных свойств.

Одной из особенностей сбора данных в геоинформатике является то, что исходные данные могут иметь не только разные размерности, но и измеряться в разных шкалах измерений. Организация геоданных направлена на объединение данных разных размерностей и шкал измерений в единую систему данных для их хранения и последующей обработки. Именно это создает возможность комплексного анализа данных [7], при



Рис.1. Общая схема организации геоданных

работе с разнородными исходными данными, измеренными в разных шкалах измерений.

Поэтому следующим этапом является классификация собранной информации, которая служит основой дальнейших действий. Классификация данных позволяет соотносит различные модели и их характеристики к разным классам, подклассам и типам, что дает возможность систематизировать исходные наборы данных и использовать свойства классов при анализе конкретных данных. Как дополнительный этап классификации геоданных в геоинформатике присутствует процедура локализации данных [8].

После того, как данные классифицированы, осуществляется их унификация. Разнообразие технологий и методов сбора данных порождает разнообразие типов данных, которые впоследствии необходимо обрабатывать. Обрабатывать множество различных данных неудобно и неэффективно. Для упрощения процесса обработки, хранения и обмена разнородные данные приводят к единому структурному виду, который используется при последующей обработке информации. Такие данные называют унифицированными.

Процедура сведения разнородных видов и структур данных к единому виду и структуре называется унификацией. В ходе унификации данных осуществляется построение единой формы данных.

После этих процедур возможно построение интегрированной модели. Унификация не создает систему данных, преобразует исходную совокупность разнородных и несогласованных данных в другую, но уже более согласованную и менее разнородную.

Для обработки по единой технологической системе и в единой информационной среде модели должны быть объединены на основе правила или метода, отвечающего требованиям оптимального хранения и обработки. Таким

объединяющим методом является интеграция данных. интеграция данных и создает систему данных вместо совокупности данных

Необходимо отметить, что геоданные образуют естественную информационную систему данных [9]. Это обусловлено тем, что они отображают реальные объекты и явления земной поверхности, которые расположены не произвольно, а организовано и имеют объективные связи друг с другом. Можно говорить, что информация об объектах и явлениях земной поверхности образует некую систему. Отдельные модели или данные являются элементами такой системы.

Интеграцией называют восстановление и (или) повышение качественного уровня взаимосвязей между элементами системы, а также процесс создания из нескольких разнородных систем единой системы, с целью исключения (до технически необходимого минимума) функциональной и структурной избыточности и повышения общей эффективности функционирования [10].

Таким образом, интеграция данных приводит к установлению дополнительных связей между данными и эти связи можно назвать системными. Можно сказать, что именно интеграция данных приводит к появлению геоданных как системы. Можно также сказать, что интеграция данных создает интегрированную модель геоданных.

Интегрированная модель не является просто суммой информационных частей ее образующих. Она, как правило, имеет меньший объем физической памяти при сохранении информационной емкости по сравнению с информационными моделями, ее составляющими, хотя включает данные о связях и дополнительную служебную информацию. Кроме того, она включает дополнительные связи между исходными данными, что создает синергетический эффект. Как следствие появляется возможность решения больше-

го количества задач, в частности комплексного анализа данных и коррелятивного анализа [11, 12].

В реальности многие модели можно отнести к интегрированным. Поэтому говорят о степени интеграции, однако другим важным параметром является критерий или аспект интеграции. Он служит основой объединения данных в интегрированную модель.

Важным свойством интеграции является то, что интеграция это не просто объединение данных, а приобретение этой моделью дополнительных свойств. В результате интеграции данных создается модель, обладающая дополнительными свойствами или, говоря языком синергетики, имеющая синергетический эффект.

Интегрированная модель является развитием информационной модели. Она более сложная, но по этой причине не только описывает информационные свойства объекта, но позволяет проводить эффективную обработку данных, относящихся к исследуемому объекту. В этом преимущество интегрированной модели.

Аспект интеграции связан с выбором устойчивого критерия интеграции. В геоинформатике имеется особенность аспекта интеграции данных. Она заключается в том, что геоданные рассматривают с учетом трех аспектов: пространственного, временного и тематического. Это означает, что данные, собираемые и хранимые в базе геоданных (БГД), группируют по трем характеристикам: “место”, “время”, “тема”.

Данные, которые выбирают для интеграции, должны быть наиболее устойчивыми или наименее изменяющимися. Временные данные по определению изменяющиеся и поэтому не могут служить основой интеграции. Тематические данные также изменчивы, они могут меняться могут исчезать или появляться в новых видах, поэтому и они не могут служить основой интеграции.

Пространственные данные — наиболее устойчивые и наименее меняющиеся, поэтому в этой группе следует искать основу для интеграции. Среди пространственных данных наиболее устойчивыми (наименее изменчивыми) являются координаты. Именно они являются основой для объединения различных данных, т.е. основой для интеграции.

Характеристика “место” является наиболее устойчивой в системе координат земной поверхности, в то время как характеристики “время” и “тема” являются изменчивыми от объекта к объекту. Глобальная устойчивость характеристики “место” и послужила основой интеграции других видов информации на этой основе.

Таким образом, если локализация создает совокупность данных с вертикальными связями, то интеграция создает систему унифицированных данных с вертикальными и горизонтальными связями. Именно системность организа-

ции данных на основе интеграции обеспечивает эффективность анализа и обработки геоданных как в геоинформатике, так и в других научных направлениях.

В результате интеграции получается некая система данных напоминающая таблицу или “универсальное отношение” из теории реляционных баз данных. Работать с такой одной таблицей неудобно и как следует из теории баз данных, ее разбивают, используя процедуры нормализации.

Другими словами, в полученной системе геоданных целесообразно задать некую структуру для удобства анализа и обработки. Для структуризации системы геоданных применяю процесс называемый стратификацией. Стратификация означает разбиение совокупности или системы на части, называемые стратами или слоями.

Стратификация координатных данных основана на важной функции координатных моделей отображать пространственные свойства объектов. Пространственные объекты характерны тем, что имеют графическую форму представления.

На рис. 2. показан процесс стратификации. Верхний уровень или слой относится к континенту. Следующий слой соотносит геоданные со страной. Следующий слой определяет регион. Под регионом находится объект. Это может быть 6 город, поселок, городской район, промышленное или транспортное предприятие.

Далее слои группируются в соответствии с задаваемыми темами, которые соответствуют объектам. Группировка может быть по некой теме, например “транспорт” или “подземные коммуникации”.

Самый нижний слой называют элементарным [13]. Он разбивает геоданные на три пространственных типа. Это данные ареальные — А, линейные — Л, точечные Т. Далее слои группируются в соответствии с задаваемыми темами, которые соответствуют объектам.

Таким образом, стратификация это не просто структуризация геоданных, а создание инструмента анализа и обобщения данных на разных территориальных или административно-территориальных уровнях.

Кроме того, стратификация превращает геоданные в уникальный информационный ресурс. В целом геоданные можно рассматривать как систему данных. Но на нижнем уровне стратификации геоданные предстают в виде информационных единиц [14]. Это дает возможность организации геоинформационного моделирования [15] с уровня информационных единиц на глобальный уровень [16].

Особенностью геоданных является наличие динамической связи между графическими данными и атрибутивными данными. Изменение атрибутивных данных влечет автоматическую замену графической информации. Это создает

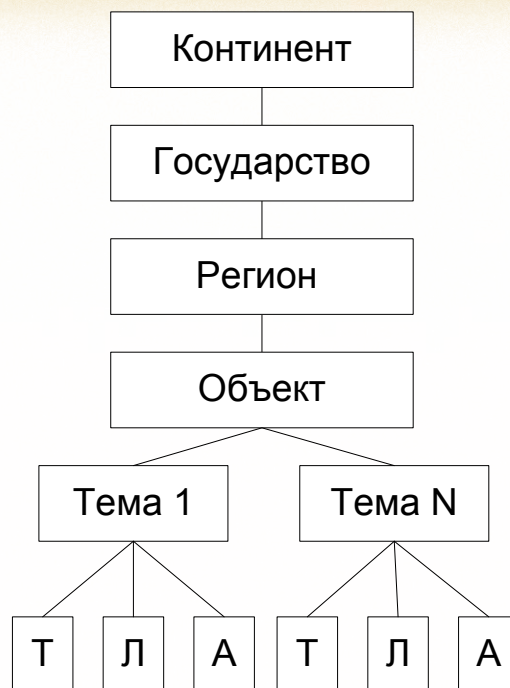


Рис.2. Стратификация геоданных

хорошую основу для пространственного анализа и управления.

Геоданные организуют с учетом семиотического подхода, а именно в виде семантической, синтаксической и прагматической частей.

Семантическая часть содержит информацию об объектах и способ ее кодирования. Синтаксическая часть включает правила построения моделей объектов и способ их отнесения к классу известных моделей. Прагматическая часть определяет ценность информации или дает возможность ее оценить. При отсутствии любой из этих трех частей информационная модель геоданных не пригодна для использования.

Таким образом геоданные являются одним из многих универсальных средств анализа пространственных объектов и явления и инструментом познания окружающего мира. Они применяются не только в геоинформатике, но и в других научных направлениях [17], включая искусственный интеллект [18].

Проблема организации геоданных сводится к решению ряда проблем. Однако организация геоданных приводит к созданию интегрированной системы данных, включающей систему моделей и систему информационных единиц. Это определяет геоданные как уникальный информационный ресурс который применяют в науке образовании и на производстве.

ЛИТЕРАТУРА

1. Цветков В.Я. Модель геоданных для управления транспортом // Успехи современного естествознания, 2009. – №4. – С. 50-51.
2. Соловьев И.В., Цветков В.Я. О содержании и взаимосвязях категорий «информация», «информационные ресурсы», «знания» // Дистанционное и виртуальное обучение, 2011. – №6 (48) – С.11-21.
3. Цветков В.Я. Пространственные отношения в геоинформатике // Международный научно-технический и производственный журнал «Науки о Земле», 2012. – №1. – С.59-61.
4. Кудж С.А. Добыча геоданных // Науки о Земле, 2013. – № 2-3. – С.82-84.
5. Поляков А.А., Цветков В.Я. Прикладная информатика: Учебно-методическое пособие для студентов, обучающихся по специальности «прикладная информатика» (по областям) и другим междисциплинарным специальностям: В 2-х частях / Под общ. ред. А.Н. Тихонова. – М.: МАКС Пресс, 2008.
6. Майоров А.А., Соловьев И.В., Кудж С.А. О новом подходе к доступу и хранению электронных аэрокосмических снимков и планов // Известия высших учебных заведений. Геодезия и аэрофотосъемка, 2011. – № 6. – С. 80-84.
7. S. A. Kudzh Geoinformation Analysis // European Researcher, 2013, Vol.(60), № 10-1 , p.2358-2365.
8. Цветков В.Я. Информатизация, инновационные процессы и геоинформационные технологии // Геодезия и аэрофотосъемка, 2006. – №4. – С.112-118.
9. Соловьев И.В. Об информационном объекте и субъекте Дистанционное и виртуальное обучение, 2012. – №5. – С. 80-84.
10. Соловьев И.В. Применение модели информационной ситуации в геоинформатике Науки о Земле, 2012. – №1. – С. 54-58.
11. Кудж С.А. Коррелятивный анализ как метод познания // Перспективы науки и образования, 2013. – №5. – С.9-13.
12. V. Ya. Tsvetkov. Framework of Correlative Analysis // European Researcher, 2012, Vol.(23), № 6-1, p.839-844.
13. S. A. Kudzh, I. V. Solovjev, V. Y. Tsvetkov System Elements Heterogeneity // European Researcher, 2013, Vol.(60), № 10-1 , p.2366-2373.
14. Цветков В. Я. Информационные единицы сообщений // Фундаментальные исследования, 2007. – №12. – С.123-124.
15. Цветков В.Я. Геоинформационное моделирование // Информационные технологии, 1999. – №3. – С.23-27.
16. V. Ya. Tsvetkov. Global Monitoring // European Researcher, 2012, Vol.(33), № 11-1, p.1843- 1851.
17. Савиных В.П., Цветков В.Я. Развитие методов искусственного интеллекта в геоинформатике // Транспорт Российской

Федерации, 2010. – № 5. – С.41-43.

18. Соловьев И.В., Кудж С.А., Дедегкаев З.Н. Об использовании универсального ключа хранения и поиска электронных аэрокосмических снимков и планов // Инженерные изыскания, 2010. – №9. – С.62-65.

REFERENCES

1. Tsvetkov V.Ia. Geodatabase model for mobility management. *Uspekhi sovremennogo estestvoznaniia - Successes of modern natural sciences*, 2009, no.4, pp.50-51 (in Russian).
2. Solov'ev I.V., Tsvetkov V.Ia. About content and interrelationship of categories «information», «information resources», «knowledge». *Distantionnoe i virtual'noe obuchenie - Distance and virtual learning*, 2011, no.6(48), pp.11-21 (in Russian).
3. Tsvetkov V.Ia. Spatial relations in the geoinformatics. *Mezhdunarodnyi nauchno-tekhnicheskii i proizvodstvennyi zhurnal «Nauki o Zemle» - International scientific-technical and production journal «Earth Sciences»*, 2012, no.1, pp.59-61 (in Russian).
4. Kudzh. S.A. Mining geodata. *Nauki o Zemle - Earth Sciences*, 2013, no.2-3, pp.82-84 (in Russian).
5. Poliakov A.A., Tsvetkov V.Ia. *Prikladnaia informatika: Uchebno-metodicheskoe posobie dlia studentov, obuchaiushchikhsia po spetsial'nosti «prikladnaia informatika» (po oblastiam) i drugim mezhdistsiplinarnym spetsial'nostiam: V 2-kh chastiakh / Pod obshch. red. A.N. Tikhonova* [Applied Informatics: textbook for the students of a speciality «applied computer science» (in some areas) and other interdisciplinary specialties: In 2 parts / edited by A.N. Tikhonov]. Moscow, MAKS Press, 2008.
6. Maiorov A.A., Solov'ev I.V., Kudzh S.A. On a new approach to access and storage of electronic aerospace images and plans. *Izvestiia vysshikh uchebnykh zavedenii. Geodeziia i aerofotos"emka - News of of higher schools. Geodesy and air photography*, 2011., no.6, pp.80-84 (in Russian).
7. S. A. Kudzh. Geoinformation Analysis. *European Researcher*, 2013, Vol.(60), no.10-1, pp.2358-2365.
8. Tsvetkov V.Ia. Information technology, innovative processes and geoinformation technologies. *Geodeziia i aerofotos"emka - Geodesy and air photography*, 2006, no.4, pp.112-118 (in Russian).
9. Solov'ev I.V. information about the object and subject. *Distantionnoe i virtual'noe obuchenie - Distance and virtual learning*, 2012, no.5, pp.80-84.
10. Solov'ev I.V. Application of a model of information situation in Geomatics. *Nauki o Zemle - Earth Sciences*, 2012, no.1, pp.54-58 (in Russian).
11. Kudzh S.A. Correlative analysis as a method of cognition. *Perspektivy nauki i obrazovaniia - Perspectives of science and education*, 2013, no.5, pp.9-13 (in Russian).
12. V. Ya. Tsvetkov. Framework of Correlative Analysis. *European Researcher*, 2012, Vol.(23), no.6-1, pp.839-844.
13. S. A. Kudzh, I. V. Solovjev, V. Y. Tsvetkov System Elements Heterogeneity. *European Researcher*, 2013, Vol.(60), no.10-1, pp.2366-2373.
14. Tsvetkov V. Ia. Information units message. *Fundamental'nye issledovaniia - Fundamental research*, 2007, no.12, pp.123-124 (in Russian).
15. Tsvetkov V.Ia. Geoinformation modeling. *Informatsionnye tekhnologii - Information technology*, 1999, no.3, pp.23-27 (in Russian).
16. V. Ya. Tsvetkov. Global Monitoring. *European Researcher*, 2012, Vol.(33), no.11-1, pp.1843-1851.
17. Savinykh V.P., Tsvetkov V.Ia. Development of artificial intelligence methods in geoinformatics. *Transport Rossiiskoi Federatsii - Transport of the Russian Federation*, 2010, no.5, pp.41-43 (in Russian).
18. Solov'ev I.V., Kudzh S.A., Dedegkaev Z.N. On the use of universal key of the storage and retrieval of electronic aerospace images and plans. *Inzhenernye izyskaniia - Engineering surveys*, 2010, no.9, pp.62-65 (in Russian).

Информация об авторе

Кудж Станислав Алексеевич
(Россия, Москва)

Профессор, доктор технических наук, ректор.
Московский государственный технический
университет радиотехники, электроники и
автоматики.

E-mail: mirearec1@yandex.ru

Information about the author

Kudzh Stanislav Alekseevich
(Russia, Moscow)

Professor, Doctor of Technical Sciences. Rector.
Moscow State Technical University
of Radio Engineering,
Electronics and Automation.

E-mail: mirearec1@yandex.ru