

МОДЕЛИ И МЕТОДЫ СТРУКТУРИРОВАНИЯ И АНАЛИЗА ДАННЫХ ОБ АБСТРАКТНЫХ ГЕОПРОСТРАНСТВЕННЫХ ОБЪЕКТАХ

Рябухин С.И., E-mail: poleem@mail.ru

Федеральный исследовательский центр ДВО РАН, Институт горного дела

Аннотация: Язык XML и его предметные расширения в области описания геопозиционированных объектов и явлений широко используется для структурирования данных о таких объектах (далее геоданных) в целях их передачи и представления. Программная совместимость таких геоданных обеспечивается за счет их изоморфности унифицированным предметным метаданным, прикладным онтологиям или структурам баз данных, построенным на известных дискурсных корпусах типов геоданных, классов геообъектов и отношений между ними. Вместе с тем, между элементами геоданных могут иметь место неявные связи, семантика которых не предусмотрена в известных корпусах и, как следствие, не распознаваема детерминированными алгоритмами программных анализаторов XML-документов. Такие ситуации характерны для абстрактных геопозиционированных объектов, в т.ч. появляющихся в результате обработки геоданных о наблюдаемых физических объектах. Гетерогенность данных о таких объектах при отсутствии их семантической совместимости существенным образом затрудняет их сохранение в репозиториях и базах данных. Одной из форм проявления неявных связей между элементами геоданных служит их морфологическое сходство, однако для его обнаружения требуется применение специальных методов анализа машиночитаемых геоданных.

Ключевые слова: расширяемый язык разметки, географические научные данные, модели и структуры геоданных, онтологии геоданных, семантический анализ данных, методы искусственного интеллекта, логика предикатов первого порядка.

MODELS AND METHODS OF STRUCTURING AND ANALYZING DATA ON ABSTRACT GEOSPATIAL OBJECTS

S I Ryabukhin, E-mail: poleem@mail.ru

Federal Research Center FEB RAS, Institute of Mining

Annotation: The XML language and its subject extensions in the field of description of geo-located objects and phenomena are widely used for structuring data on such objects (hereinafter referred to as geodata) for the purposes of their transmission and presentation. Software compatibility of such geodata is ensured by their isomorphism with unified subject metadata, applied ontologies or database structures built on known discourse corpora of geodata types, geo-object classes and relations between them. At the same time, there may be implicit links between the elements of geodata, the semantics of which is not provided for in the known corpora and, as a result, is not recognized by the deterministic algorithms of software parsers of XML

documents. Such situations are typical for abstract geographic objects, including those appearing as a result of processing geodata on observed physical objects. The heterogeneity of data on such objects in the absence of their semantic compatibility significantly complicates their storage in repositories and databases. One of the forms of manifestation of implicit connections between elements of geodata is their morphological similarity, but its detection requires the use of special methods for analyzing machine-readable geodata.

Key words: Extensible Markup Language, geographic scientific data, geodata models and structures, geodata ontologies, semantic data analysis, artificial intelligence methods, first-order predicate logic.

Введение

Обеспечение программной и семантической интероперабельности данных на стадиях их обработки, хранения, передачи и представления – является одной из важнейших задач информатики [1]. Частный случай такой задачи – создание условий для программной и семантической совместимости геоданных, т.е. данных о геопозиционированных объектах и явлениях.

В настоящее время совместимость геоданных на уровне программных приложений обеспечивается путём унификации их структур в соответствии с эталонными метаданными. В частности, для сохранения геоданных в базах данных геоинформационных систем используются эталонные реляционные структуры, известные под аббревиатурой SHP (Shapefile Technical Description) [1], основанные на формате «.dbf». При передаче и представлении геоданных в качестве эталонных метаданных используются расширения языка XML в виде схем XSD (XML Schema Documentation) для географических метаданных, в частности, грамматики языков KML (Keyhole Markup Language) [3] и GML 3.2.1 (Geography Markup Language) [4]. Спецификации таких грамматик основаны на абстрактной модели географических данных консорциума OpenGIS Abstract Specification [5] и объектной модели DOM (Document Object Model) [6]. Такая унификация геоданных позволяет осуществлять обработку XML/GML-документов посредством детерминированных алгоритмов и выявлять наборы экземпляров типов данных в соответствии с предметными грамматиками

В отличие от программной интероперабельности геоданных, обеспечение их семантической совместимости на уровне отношений между типами данных и связей их экземпляров, – представляет собой более сложную задачу. Для решения такой задачи необходим расширенный анализ машиночитаемых геоданных, позволяющий обнаруживать неявные отношения между типами геоданных, непредусмотренные грамматиками метаданных и известными предметными онтологиями. Проблема семантической совместимости геоданных отчетливо проявляется в случае аккумулирования сведений об абстрактных геопозиционированных сущностях или их категориях. Такие сведения могут быть получены, например, в ходе наблюдений за реальными объектами и явлениями или посредством обработки результатов наблюдений, в т.ч. путём мультиплексирования данных дистанционного зондирования и т.д.

В целом, проблемы выявления семантических связей между экземплярами типов геоданных в машиночитаемых XML-документах обусловлены несколькими причинами. Во-первых, язык XML имеет существенные ограничения в части передачи смыслового содержания данных и метаданных. Как известно, связи в языке XML представлены двумя видами явных отношений, а именно иерархией элементов (тегов) и ассоциациями элементов посредством ссылок. Прочие виды отношений выявляются исключительно путём интерпретации, в том числе с применением эвристических методов лингвистического и синтаксического анализа элементов XML-документов [7].

Во-вторых, применимость обобщенных онтологических моделей геопространственных объектов и явлений [8] существенным образом зависит от наличия ассоциированных с ними прикладных онтологий и спецификаций метаданных, например, таких как OpenStreetMap [9], CityGML [10], GeoSciML[11], GeoNames [11] и т.д. Вместе с тем, несмотря на то, что такие прикладные онтологии находятся на этапе развития, их роль не столь существенна, так как полноту всякой онтологии по определению нельзя считать исчерпывающей [13].

В-третьих, неявные связи между экземплярами геоданных могут иметь хронологическую или холистическую природу, проявляющуюся, например, в

общности морфологии геоданных [14]. При этом для автоматического выявления признаков такой общности требуется применение методов интеллектуального анализа геоданных.

Например, не следует ожидать точного установления местоположения таких абстрактных объектов, как например антропогенные ландшафтные зоны, сообщества видов растений, биоценозы [15,16] и т.д., если поиск границ таких виртуальных сущностей осуществляется исключительно на основе статистических критериев схожести групп их элементов. С одной стороны, для установления состава и локализации сообществ видов растений более продуктивным будет оценивание сродства групп видов растений на основе совокупности знаний, в т.ч. сведений об условиях совместимости отдельных видов растений [17]. С другой стороны, возможны ситуации, когда совокупность подлежащих учету условий, ограничений и прочих знаний об «отношениях» между представителями живой природы может быть неполной.

Другим примером абстрактного геообъекта с нечеткими границами, неполными знаниями о составе и гетерогенным набором данных служит вечная мерзлота. Для локализации зон вечной мерзлоты учитывают множество косвенных данных и их взаимосвязей, как-то: состояние поверхностного слоя грунта в разное время года; температура воздуха; состояние снежного покрова; состояния растительности и т.д. Вместе с тем, остаётся открытым вопрос о составе, структуре и содержании данных, наиболее полным образом описывающих явление вечной мерзлоты [18].

В целом, понятие «абстрактный геообъект» соответствует известному классу географических метаданных категории «GeoScientificInformation» – научных данных о Земле, в т.ч. геофизических особенностях и процессах, полезных ископаемых, землетрясениях, вулканической активности, оползнях, почвах, эрозиях [19].

Нетрудно видеть, что в целом при обработке унифицированных данных об указанных выше абстрактных геообъектах решаются задачи анализа и классификации машиночитаемых данных, при этом для решения таких задач необ-

ходимо применение интеллектуальных методов распознавания структуры и содержания геоданных. В целом проблемы обработки таких геоданных обусловлены следующими факторами: 1) гетерогенностью данных, проявляющейся в широком спектре типов данных, в т.ч. в наличии категориальных данных; 2) неконечномерностью наборов атрибутов геообъектов, затрудняющей создание полных метрических пространств [20], обеспечивающих возможность сопоставления таких объектов путём исчисления мер их схожести; 3) необходимостью обоснования получаемых результатов анализа и классификации геоданных.

Перечисленные выше факторы делают затруднительным применение известных методов нечеткой классификации данных с интерпретируемым механизмом получения решений, в частности, классификационных деревьев с обучением [21], а также методов кластеризации категориальных данных с использованием статистических критериев схожести кластеров [22]. Сказанное объясняет необходимость разработки иных методов анализа структур данных об абстрактных геопозиционированных объектах. Актуальность разработки таких методов объясняется востребованностью способов обеспечения совместимости геоданных при их сохранении в репозиториях и базах данных без каких-либо потерь сведений, сопутствующих унификации структур данных.

Объектами исследования послужили, во-первых, спецификации географических метаданных, в частности актуальные в настоящее время грамматики GML, во-вторых, известные онтологии географических данных, в-третьих, методы и алгоритмы интеллектуального анализа и классификации разнородных по составу совокупностей данных.

Структуры геоданных для абстрактных объектов в грамматике GML

Как известно, грамматика GML 3.2.1 и её дополнения [23], содержат XSD-схемы структур данных для предопределённого набора классов, используемых в целях описания геопозиционированных объектов [5]. В частности, в такой набор классов входят: 1) геометрические примитивы; 2) геометрические

комплексы; 3) системы координат; 4) топологии; 5) хронологии; 6) единицы измерения; 7) направления; 8) наблюдения; 9) покрытия.

В контексте нашего исследования мы рассматриваем понятие «абстрактный геопозиционированный объект» как продукт обработки геоданных о реальных объектах и явлениях. Мы полагаем, что среди указанных выше классов, именно концепция покрытия (coverages) является наиболее пригодной для адекватной и полной передачи сути понятия «абстрактный геопозиционированный объект», а также для описания соответствующих ему структур данных.

Строго говоря, назначение класса покрытий состоит в спецификации способов связи свойств геообъектов с их пространственным распределением [4], при этом грамматика GML содержит предопределённый набор наименований способов такой связи, включая ассоциативную и аналитическую связь. В то же время, в грамматике GML постулируется, что покрытия могут служить результатом ассоциативной связи не только экземпляров указанных выше классов (например: векторов; растров; различного вида триангуляционных сеток; точечных и других геометрий и т.д.), но и абстрактных объектов, в том числе представленных в виде структур данных. Например, одним из таких абстрактных объектов может являться совокупность данных об инфраструктуре автомобильных дорог, другим – данные о транспортируемых грузах. Таким образом, концепция покрытий может выполнять более значительную роль с точки зрения представления сведений об абстрактных геообъектах, а также с точки зрения анализа структур данных о таких объектах на предмет выявления их связанности.

Формальное описание сведений об абстрактных геообъектах

Анализ содержания грамматики GML, в т.ч. анализ текстового описания глоссария, ограничений на естественном языке, а также структур данных на языке XSD, – в целом позволил составить формализованное, предикатное представление о составе сведений, входящих в определение класса покрытий (см. выражение 1). С одной стороны, такое представление является гипотетическим, с другой стороны оно необходимо для более полного понимания роли покры-

тий в описании абстрактных геообъектов как результатов отношений между другими реальными и абстрактными объектами.

$$\exists (x, y, c, p, f, v) [(x, y) \in \{O\}, c \in \{c_1, \dots, c_8\}, p \in \{p_1, \dots, p_n\}, f \in \{f_1, \dots, f_n\}, v \in \{\{V_1\}, \dots, \{V_n\}\}, R^1(x, y), R^2(x, c), R^3(x, p), R^4(x, p, f, v)] \quad (1)$$

В выражении 1 используются следующие обозначения: $\{O\}$ – геообъекты; $\{c_1, \dots, c_8\}$ – классы грамматики GML; $\{p_1, \dots, p_n\}$ – атрибуты геообъектов; $\{f_1, \dots, f_n\}$ – способы связи местоположения объекта и значений атрибутов; $\{\{V_1\}, \dots, \{V_n\}\}$ – диапазоны значений атрибутов; предикаты: $R^1(x, y)$ – отношения между геообъектами; $R^2(x, c)$ – представление геообъекта посредством класса грамматики GML; $R^3(x, p)$ – свойства геообъекта; $R^4(x, p, f, v)$ – поле покрытия.

Следует заметить, что предикат $R^1(x, y)$ служит обобщенной формой бинарных семантических отношений между реальными геообъектами и именно наличие таких отношений является определяющим для введенного нами понятия «абстрактный геообъект». При этом специфика отношений вида $R^1(x, y)$ не позволяет осуществить их идентификацию посредством детерминированного алгоритмического исчисления, например путём расчёта пространственной близости геопозиционированных объектов x и y с применением функционала SpatialSQL. Другими словами, для исчисления отношений $R^1(x, y)$ необходим аппарат, основанный на концепции «открытых данных», т.е. на модели знаний.

Предположим, что существуют отношения $R_i^1(x, y)$, которые проявляются в морфологическом сходстве структур данных S^x и S^y , описывающих объекты x и y . Тогда предикат $R_i^1(x, y)$ и структуры данных для объектов x и y допустимо связать импликацией, вида: $\{S^x, S^y, R_i^1(x, y) \rightarrow \text{МераСходства}(S^x, S^y)\}$. Например, в случае покрытия для одного отдельно взятого объекта его структура будет описываться предикатами $R^2(x, c)$, $R^3(x, p)$, $R^4(x, p, f, v)$ (см. выражение 1).

Указанная выше импликация представляет собой аксиоматическую продукционную модель, для практического использования которой необходимо уточнить содержание функционального предиката $\text{МераСходства}(S^x, S^y)$, а также конкретизировать виды отношений R^2 , R^3 , R^4 .

Для уточнения указанных выше видов отношений составлена концептуальная модель связей элементов в схемах XSD, постулированных в грамматике GML (см. рис.1). Предикатная иллюстрация графических элементов такой модели позволяет детализировать набор отношений R^1 , R^2 , R^3 , R^4 , входящих в структуры данных для абстрактных геообъектов. Концептуальная модель на рисунке 1 визуализирует явные и неявные отношения между элементами схем XSD в грамматике GML. При этом явные отношения элементов следуют из синтаксиса языка XML/XSD, а неявные отношения элементов установлены на основании высказываний на естественном языке, используемом для расширения синтаксиса языка XML, в том числе для описания ограничений на участие элементов в отношениях.

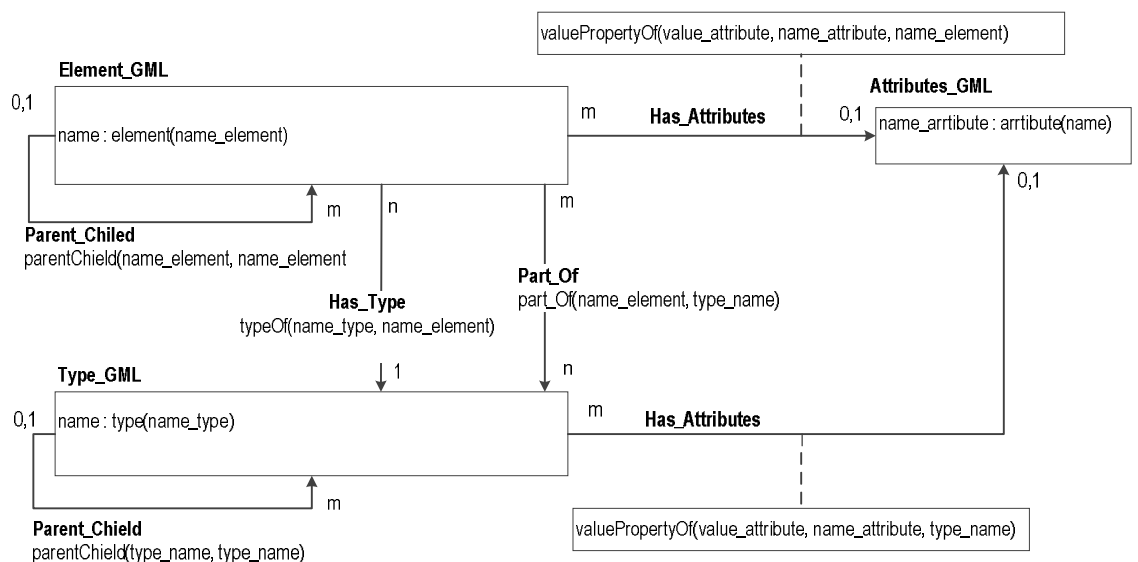


Рисунок 1 – Концептуальная модель элементов GML 3.2.1.

На рисунке 1 используются следующие обозначения: *Element_GML* – простые элементы грамматики GML; *Type_GML* – составные элементы грамматики GML; *Attributes_GML* – атрибуты элементов; *Has_Type* – ассоциативная связь элементов; *Part_Of* – вложенность простых элементов в составные; *Has_Attributes* – ассоциативная связь элементов и атрибутов; *Parent_Child* – наследуемость элементов схем XSD в грамматике GML. Указанные выше элементы концептуальной модели объясняют содержание отношений R^2 , R^3 , R^4 следующим образом:

- R^2 : *element()*, *type()*, *part_Of()*, *parentChild*;
- R^3, R^4 : *valuePropertyOf()*, *name_attribute()*.

Расхождение в составе переменных предиката R^4 и *valuePropertyOf* объясняется наличием связи *typeOf()*, которая определяет тип элемента схемы XSD как способ связи местоположения объекта и значений атрибутов (примечание – в дальнейшем предикатная форма связи «элемент – свойство элемента» будет уточнена).

Кардинальность отношений на рисунке 1 позволяет квантифицировать соответствующие предикатные переменные (см. выражения 2-8).

$$\exists x \{x \in \{name_element\}, element(x)\} \quad (2)$$

$$\exists v \forall (z,x) \{valuePropertyOf(v, z, x)\} \quad (3)$$

$$\forall z \{z \in \{attribute\}, name_attribute(z)\} \quad (4)$$

$$\exists (x,y) \{element(x), type(y), part_Of(x, y)\} \quad (5)$$

$$\forall (x,y) \{element(x), type(y), typeOf(y,x)\} \quad (6)$$

$$\exists y \{y \in \{name_element\}, type(y)\} \quad (7)$$

$$\exists ! x \exists y \{(x,y) \in \{name_element\}, parentChild(x,y)\} \quad (8)$$

Совокупность предикатных выражений 2-8 позволяет уточнить состав аксиоматической продукционной модели $\{S^x, S^y, R^l(x,y) \rightarrow МераСходства(S^x, S^y)\}$, в частности сделать замену обозначений S^x и S^y структур абстрактных объектов.

Для составления базы знаний для указанной выше продукции была построена дополнительная сетевая интерпретация грамматики GML. Семантическая сеть включает более 200 узлов – элементов схем XSD грамматики GML, и более 300 дуг – связей между такими элементами. Полезность семантической сети заключается в следующем. Во-первых, при её построении был уточнён набор отношений, полученных на основании концептуальной модели (см. рисунок 1), в частности выявлено дополнительное отношение *partOfSet()*. Приведём пример такого отношения в схеме XSD грамматики GML – *partOfSet("aggregationType", "gml:AggregationAttributeGroup")*. Во-вторых, сетевая модель обеспечила возможность визуального контроля соблюдения огра-

ничений на допустимые виды отношений между теми классами в грамматике GML. В-третьих, сетевая модель является машиночитаемой, а именно выполнена в формате «*.vsd», что позволяет редактировать её структуру и содержание, а также автоматически получать спецификацию её элементов, в том числе формировать базу знаний аксиоматической системы в виде означенных предикатов или, иначе атомарных аксиом, например: *typeOf*("gml:AbstractGMLType", "gml:AbstractGML"); *propertyOf*("location", "gml:AbstractFeatureType"); *parentChild*("gml:AbstractFeature", "gml:AbstractTopology") и т.д.

Метрики сходства абстрактных геообъектов

Назначение функционального предиката *МераСходства()* состоит в исчислении морфологической близости структур абстрактных объектов. Как ранее было сказано, сложность такого исчисления обусловлена гетерогенностью таких данных, в т.ч. наличием категориальных данных, а также неконечностью наборов атрибутов абстрактных геообъектов. Другими словами, близость геообъектов O^1 и O^2 невозможно оценить точным образом и без потери значимой информации, если их сопоставление осуществляется исключительно на основе матрицы разностей значений атрибутов, входящих в базис метрического пространства. В противном случае, в контексте принятой продукционной модели $\{S^x, S^y, R^l(x,y) \rightarrow \text{МераСходства}(S^x, S^y)\}$, – *МераСходства()* является следствием морфологической связи структур данных и такая связь является результатом предикатного логического исчисления. Говоря иначе, для точной оценки близости объектов помимо известных метрик следует применять псевдометрики вида $F(|S^x, S^y|)$, где: $|S^x, S^y|$ – это матрица морфологического сходства структур S^x, S^y ; F – аксиоматическая модель дедуктивного вывода о сходстве объектов O^1 и O^2 .

Концепция абстрактных геообъектов охватывает весьма широкий дискурс и было бы неправильным говорить о создании универсального набора псевдометрик, позволяющего оценивать сходство любых геообъектов. Мы локализовали область поиска способов установления совместимости унифицированных геоданных и ограничились использованием лишь определённого класса мета-

данных, а именно покрытий. При этом мы абстрагируемся от проблем разнородности предметных характеристик геообъектов, образующих покрытия, а также от геометрии областей покрытий и акцентируем внимание на анализе морфологии данных, представленных сложными типами данных и их структурами. Именно типы и структуры данных представляют важность, так как проявляются посредством наборов видов отношений между элементами данных. С учетом сказанного, задача формирования псевдометрик сводится к установлению набора видов отношений и признаков таких отношений. В тоже время, мы ограничим группу допустимых видов отношений областью типов и структур данных, характерных для так называемого диапазона охвата покрытий, отвечающего за описание измеренных, смоделированных или наблюдаемых параметров геообъектов и их значений [24]. Еще более сужая предметную область, мы остановимся на тех структурах данных, которые являются неотъемлемой частью всякого экземпляра покрытия, а именно отражают сведения о типах шкал для параметров геообъектов. Например, в грамматике GML в структуре метаданных о покрытиях за указание типов применяемых шкал при описании покрытий отвечают элементы «gml:rangeSet», «gml:MeasureType», а также структура данных «gml:UomIdentifier» [4]. Содержание указанных структур данных является унифицированным [25], то есть распознаваемым алгоритмами программ-анализаторов (парсеров) и, следовательно, может быть использовано в качестве основания для идентификации наборов отношений, ассоциированных с каждой, в том числе категориальной шкалой. Таким образом, при наличии формальных, предикатных спецификаций видов отношений, характерных для каждой шкалы, появляется возможность понятным образом декомпозировать структуру функционального предиката *МераСходства()*. Аппарат исчисления данного предиката основан на попарном сопоставлении структур абстрактных геообъектов приведён в выражении 9.

$$\begin{aligned}
 \text{МераСходства}(S^x, S^y) \leftarrow \{ & [R^2(x, c), R^3(x, p), R^4(x, p, f, v), (Scale_i(s^x_i) \rightarrow R^l_{ij}(s^x_i))], \\
 & [R^2(y, c), R^3(y, p), R^4(y, p, f, v), (Scale_i(s^y_i) \rightarrow R^l_{ij}(s^y_i))], \\
 & \text{Equivalence}(R^l_{ij}(s^x_i), R^l_{ij}(s^y_i)) \}
 \end{aligned} \tag{9}$$

В выражении 9 используются следующие обозначения: $Scale_i(s^x_i)$, $Scale_i(s^y_i)$ – шкалы для элементов частных геообъектов, представленных структурами s^x_i и s^y_i ; $R^l_{ij}(s^y_i)$ – виды отношений, характерные для шкал; предикаты $R^2(x,c)$, $R^3(x,p)$, $R^4(x,p,f,v)$ и их содержание упоминались нами ранее (см. выражения 2-8); $Equivalence(R^l_{ij}(s^x_i), R^l_{ij}(s^y_i))$ – функциональный предикат, поставляющий количественную оценку сходства видов отношений, по очевидным причинам мы опустим описание его реализации. Таким образом, раскрыв содержание аппарата исчисления предиката $МераСходства()$, мы считаем полностью выполненной задачу описания содержания продукционной модели $\{S^x, S^y, R^l(x,y) \rightarrow МераСходства(S^x, S^y)\}$ логического исчисления схожести абстрактных геообъектов.

Заключение

В статье приведены результаты анализа известной унифицированной спецификации географических метаданных, а именно грамматики GML, в том числе определены структуры данных, которые наиболее полным образом соответствуют концепции абстрактных геопозиционированных объектов. Цель такого анализа заключалась в выявлении признаков, обеспечивающих как программную так и семантическую совместимость данных о таких объектах. Полученные результаты позволили создать, во-первых, формальное предикатное описание структуры унифицированных метаданных, во-вторых, сформировать концепцию псевдометрик, необходимых для сопоставления гетерогенных географических данных, представленных так называемыми полями покрытий, в-третьих, разработать структуру аксиоматической продукционной модели, пригодную для логического исчисления сродства абстрактных географических объектов на основании морфологического сходства структур данных о таких объектах. В дальнейшем, с точки зрения практической полезности разработанной аксиоматической модели исчисления сродства геообъектов предстоит сформировать уточненный набор логических правил разработанной аксиоматической модели, выполнить её верификацию по критериям непротиворечивости и выводимости, а также разработать алгоритм и исполняемый код логической

программы, выполняющей функции анализа и классификации геоданных, представленных в формате «.gml», структура которых унифицирована в соответствии с грамматикой GML.

Список литературы

1. Макаренко С. И., Олейников А. Я., Черницкая Т. Е. Модели интероперабельности информационных систем // Системы управления, связи и безопасности. 2019. №4.
2. Shapefile Technical Description – Режим доступа: <https://www.esri.com/content>. – дата обращения: 01.10.2024.
3. Keyhole Markup Language – Режим доступа: <http://earth.google.com/kml/2.2> – дата обращения: 01.10.2024.
4. Geography Markup Language (GML 3.2.1). – Режим доступа: <https://docs.ogc.org/as/04-084r4/04-084r4.html>. – дата обращения: 01.10.2024.
5. OpenGIS Abstract Specification – Режим доступа: <https://www.ogc.org/ru/publications/standard/abstract-specification>. – дата обращения: 01.10.2024.
6. Document Object Model – Режим доступа: <https://www.w3.org/TR/REC-DOM-Level-1>. – дата обращения: 01.10.2024.
7. Ammara F., Aqsa M., Kiran Q., Saleem M. An Approach to Map Geography Mark-up Language Data to Resource Description Framework Schema // Communications in Computer and Information Science. - 2020. P. 343-354. –DOI: 10.1007/978-981-15-5232-8.
8. W3C Geospatial Ontologies – Режим доступа: <http://www.w3.org/2005/Incubator/geo/XGR-geo-ont-20071023>. – дата обращения: 01.10.2024.
9. OpenStreetMap – Режим доступа: <http://www.openstreetmap.org>. – дата обращения: 01.10.2024.
10. CityGML – Режим доступа: <https://www.ogc.org/ru/publications/standard/citygml>. – дата обращения: 01.10.2024.
11. Geoscience Markup Language – GeoSciML. – Режим доступа: <http://geosciml.org/schemas>. – дата обращения: 01.10.2024.
12. GeoNames – Режим доступа: <https://www.geonames.org>. – дата обращения: 01.10.2024.
13. Gruber, T.R. The role of common ontology in achieving sharable, reusable knowledge bases / J.A. Allen, R. Fikes, E. Sandewell (eds.) // Principles of Knowledge Representation and Reasoning. Proceedings of the Second International Conference. – Morgan Kaufmann. – 1991. – P. 601-602.
14. Kai Sun, Yunqiang Zhu, Peng Pan (et al) Geospatial data ontology: the semantic foundation of geospatial data integration and sharing // Big Earth Data. – 2019. – DOI: 10.1080/20964471.2019.1661662.

15. Шитиков В. К., Розенберг Г. С. Оценка биоразнообразия: Попытка формального обобщения / Количественные методы экологии и гидро-биологии (сборник научных трудов, посвященный памяти А.И. Бака-нова) \ Отв. ред. чл.-корр. РАН Розенберг Г С (Тольятти: СамНЦ РАН) – 2005.
16. Возможности использования спутникового сервиса ВЕГА для реше-ния различных задач мониторинга наземных экосистем Барталев С.А., Ершов Д.В., Лупян Е.А., Толпин В.А. // Современные проблемы дис-танционного зондирования Земли из космоса, 2012. Т. 9. № 1. С.49-56.
17. Ryabukhin S. I. Method of analysis of floral lists of higher plants // IOP Conference Series: Earth and Environmental Science, Volume 547, SCI-ENTIFIC SUPPORT FOR THE SUSTAINABLE DEVELOPMENT OF AGRO-INDUSTRIAL COMPLEX 16-17 July 2020, Khabarovsk, Russian Federation. 2020 IOP Conf. Ser.: Earth Environ. Sci. 547 012032.
18. Буффар Трой, Урюпова Екатерина, Доддс Клаус и др. Научное со-трудничество: мониторинг вечной мерзлоты циркумполярной зоны и обмен данными // АиС.- 2021. №45. - DOI: 10.37482/issn2221-2698.2021.45.184.
19. ISO 19115-2:2019 – Географическая информация. Метаданные. Часть 2. Расширения для сбора и обработки данных. – Режим досту-па:<https://www.gostinfo.ru/catalog/Details/?id=6443793>. – дата обраще-ния: 01.10.2024.
20. Петровский А.Б. Пространства множеств и мультимножеств. – М.: Едиториал УРСС, 2003. – 248 с.
21. Полин Я.А., Зудилова Т.В., Ананченко И.В. и др. Деревья решений в задачах классификации: Особенности применения и методы повыше-ния качества классификации // Современные наукоемкие технологии. – 2020. – № 9. – С. 59-63. – DOI: <https://doi.org/10.17513/snt.38215>.
22. Yang Y, Guan X and You J CLOPE: a fast & effective clustering algorithm for transactional data // Proceedings of the eighth ACM SIGKDD interna-tional conference on Knowledge discovery & data mining (ACM).– 2002. – pp. 682-687.
23. Geography Markup Language (GML) – Extended schemas and encoding rules. – Режим доступа: <http://www.ogc.org/ru/publications/standard/gml>. – дата обращения: 01.10.2024.
24. Spatial Data on the Web Best Practices - W3C Editor's Draft 09 December 2024. – Режим доступа: <https://www.w3.org/2021/sdw/bp/#spatial-things-features-and-geometry>. – дата обращения: 01.10.2024.
25. The Unified Code for Units of Measure – UCUM v.2.2.– Режим доступа: <https://ucum.org/ucum>. – дата обращения: 01.10.2024.