

ОНТОЛОГИЧЕСКИЙ ПОДХОД К ПРОЕКТИРОВАНИЮ БАЗЫ ДАННЫХ ДЛЯ ОЦЕНКИ ВЛИЯНИЯ ЭНЕРГЕТИКИ НА ОКРУЖАЮЩУЮ СРЕДУ

Ворожцова Татьяна Николаевна

К.т.н., вед. инженер, e-mail: tnn@isem.irk.ru

Макагонова Надежда Николаевна

К.т.н., гл. специалист, e-mail: mak@isem.irk.ru

Массель Людмила Васильевна

Д.т.н., заведующий лабораторией, e-mail: massel@isem.irk.ru

Лаборатория информационных технологий в энергетике

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

664130 г. Иркутск, ул. Лермонтова 130

Аннотация. В настоящее время очевидной является проблема влияния энергетики на окружающую среду как одного из антропогенных факторов. В рамках исследований по оценке влияния энергетики на геоэкологию региона предполагается разработка Web-ориентированной информационной системы, интегрирующей математические и семантические методы, инструментальные средства, базу данных и базу знаний рассматриваемой предметной области. В статье рассматривается предлагаемый онтологический подход к структурированию и интеграции информации, необходимой для исследований. Онтологии предлагается рассматривать как прообраз инфологической модели данных при проектировании базы данных и базы знаний.

Ключевые слова: геоэкология, антропогенный фактор, инфологическая модель, семантическое моделирование, онтология, метаонтология, база данных.

Цитирование: Ворожцова Т.Н., Макагонова Н.Н., Массель Л.В. Онтологический подход к проектированию базы данных для оценки влияния энергетики на окружающую среду // Информационные и математические технологии в науке и управлении. 2019. № 3 (15). С. 31–41. DOI: 10.25729/2413-0133-2019-3-03

Введение. В ИСЭМ СО РАН совместно с коллективами ученых Беларуси и Армении при поддержке фондов ЕАПИ¹ и РФФИ² в рамках международного проекта «Методы и технологии оценки влияния энергетики на геоэкологию региона» [7] выполняются исследования по оценке влияния объектов энергетики на экологию региона. Основной задачей геоэкологии является изучение изменений геосферных оболочек под влиянием природных и антропогенных факторов. Влияние энергетики на природную среду, как одного из антропогенных факторов, очевидно, но сложно выявить вклад той или иной отрасли промышленности в формирование загрязнений элементов природной среды. В настоящее время оценка влияния энергетики на экологию производится путем постоянного мониторинга эмиссии загрязняющих веществ, проведения замеров, а также расчета

¹ ЕАПИ – Евразийская Ассоциация поддержки научных исследований, учрежденная в июле 2016 г. по инициативе РФФИ совместно с партнерскими организациями Беларуси, Армении, Киргизии и Монголии.

² Результаты получены в рамках выполнения проекта по госзаданию №АААА-А17-117030310444-2, отдельные аспекты прорабатывались в рамках проектов, поддержанных грантами РФФИ № 18-57-81001 и № 19-07-00351

выбросов, сбросов загрязняющих веществ, образования отходов. Для оценки влияния энергетических объектов на природную среду рассматриваются характерные загрязняющие вещества, поступающие при функционировании этих объектов – водяной пар, сажа (углерод), взвешенные вещества, пыль неорганическая, золошлаковые отходы, диоксид серы, оксид углерода, оксиды серы и другие, замеряемые при обследовании объектов. Кроме замеров и расчетов, используются статистическая информация, отчетные данные конкретных предприятий, государственные доклады «О состоянии и об охране окружающей среды» [1].

1. Задачи проекта. Задачами выполняемых исследований по оценке влияния объектов энергетики на экологию региона являются, в частности, анализ существующих методов оценки выбросов загрязняющих веществ от энергетических объектов и моделей распространения загрязнений, а также определение состава информации, необходимой для использования рекомендуемых методов.

Одной из задач проекта является разработка Web-ориентированной информационной системы (WIS), которая интегрирует математические, семантические методы, инструментальные средства, базу знаний, базу данных и геоинформационную систему.

Архитектура WIS представлена на рис. 1 и включает четыре уровня:

- уровень семантического моделирования,
- уровень представления знаний,
- уровень представления данных,
- уровень математического моделирования.

База данных проектируется и реализуется на уровне представления данных, с использованием онтологий, реализуемых на уровне представления знаний.

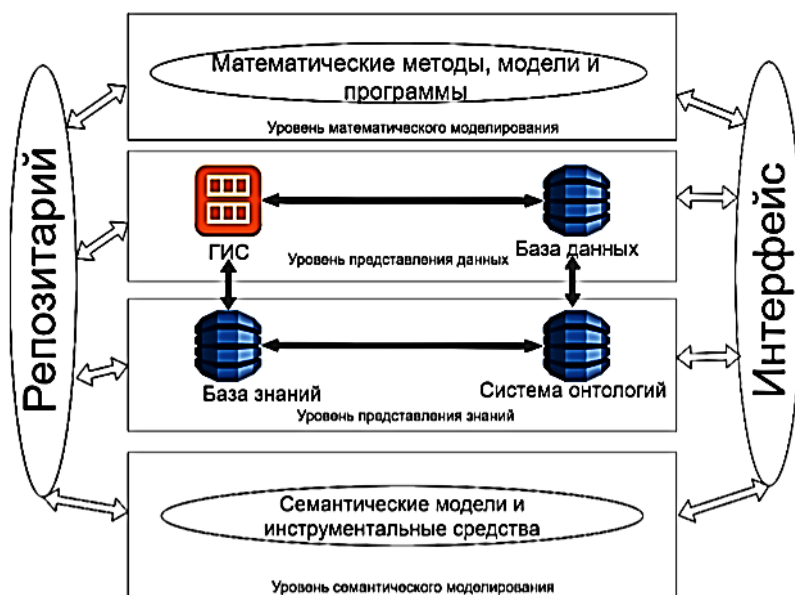


Рис. 1. Архитектура Web-ориентированной информационной системы

2. Состав необходимой информации. Исследователи ИСЭМ СО РАН, специалисты в области экологии, давно занимаются экологическими проблемами Байкальского региона и оценками негативного влияния объектов энергетики на окружающую среду, изучая выбросы загрязняющих веществ (сера, азот, сажа и др.) предприятий топливно-энергетического комплекса (ТЭК) [в частности, 5]. Методической основой оценки такого воздействия является системный подход, в котором учитываются все взаимосвязанные процессы

производства – добыча топливно-энергетических ресурсов, их преобразование в конечный продукт (тепловую или электроэнергию, продукты переработки углеводородного топлива) и передача его конечному потребителю. При этом используются как непосредственные натурные наблюдения и данные мониторинга текущего состояния природной среды, так и статистическая и отчетная информация предприятий.

Для выполнения расчетов требуется широкий перечень сведений об энергоисточниках, функционирующих в Байкальском регионе, их составе, типах оборудования, мощности, производстве тепло- и электроэнергии, выбросах загрязняющих веществ и т.д. Энергоисточники подразделены на три группы объектов, сжигающих различные виды органического топлива, характерные для исследуемого региона:

- тепловые электрические станции,
- дизельные электростанции,
- котельные.

В таблице 1 в качестве примера приведены основные показатели по тепловым электрическим станциям с указанием наименования, принятого обозначения единицы измерения, источника информации и примечания. Значения некоторых показателей можно получить из статистических отчетов, но большая часть показателей обычно собирается специалистами во время проведения экспедиций при обследовании энергопредприятий, изучении годовых технических отчетов, паспортов предприятий, справочных книжек энергетика.

Таблица 1. Основные показатели тепловой электрической станции

Показатель	Принятое обозначение	Единица измерения	Источник информации	Примечание
Мощность электрическая	N	МВт	Формы статистической отчетности 6-ТП	
Мощность тепловая	N _т	Гкал/ч	Официальные сайты, годовые отчеты генерирующих компаний	
Тип и марка котлоагрегатов			Паспорта энергопредприятий, справочная книжка энергетика	
Паро-, тепло-производительность котлов		т/ч, или ГДж/ч	Официальные сайты, годовые отчеты генерирующих компаний, паспорта котельных агрегатов	
Производство электроэнергии	W	млн. кВт·ч/год	Формы статистической отчетности 6-ТП	
Отпуск тепловой энергии	Q	тыс. Гкал/год	Формы статистической отчетности 6-ТП	
Вид и марка топлива			По крупному – формы статистической отчетности 6-ТП. С указанием добывающего предприятия – годовые отчеты энергокомпаний	Необходимо максимально подробно с указанием месторождения или добывающего предприятия
Суммарный расход топлива на выработку энергии	B	тыс. т/год	В условном исчислении – формы статистической отчетности 6-ТП	Перевод в натуральные объемы с использованием теплотворной способности видов топлива конкретных месторождений и добывающих предприятий

Топочное устройство			Обследование, годовые технические отчеты энергоснабжающих организаций, паспорта предприятий, справочная книжка энергетика	
Конструкция горелок				
Тип и марка очистного оборудования				
Тип шлакоудаления				Сухое или мокрое
Высота дымовой трубы	h	м		

Кроме основных технико-производственных показателей предприятий для оценки выбросов загрязняющих веществ в атмосферу необходима подробная информация о технических характеристиках топлива, способе его сжигания, конструкции и характеристиках топочного и очистного оборудования, составе загрязняющих веществ.

Для использования такого разнородного набора данных необходимо выполнить их структуризацию и спроектировать базу данных (БД) для хранения данных и их использования в моделях для оценки влияния энергопредприятий на геоэкологию Байкальского региона.

3. Семантическое описание предметной области. Семантическое описание предметной области является начальным этапом представления знаний, позволяющим выявить основные понятия и их взаимосвязи. Одним из способов описания семантических аспектов предметной области является онтология – «формализованная спецификация концептуализации» [14 - 16], т.е. формализованное описание терминов (концептов) предметной области и отношений между ними. Формализация позволяет применять программные средства для работы с онтологиями.

Онтологии используются в различных областях информатики и искусственного интеллекта, так как дают возможность представления и структурирования знаний. Существуют разные направления и технологии использования онтологий при разработке баз данных и информационных систем [10 - 13], такие, как:

- OBDI (Ontology Based Data Integration) – интеграция данных, базирующаяся на онтологиях;
- OBDA (Ontology Based Data Access) – доступ к данным, базирующийся на онтологиях;
- OBIS (Ontology Based Information Systems) – информационные системы, базирующиеся на онтологиях.

Как правило, при разработке баз данных онтологии используются в качестве концептуальных схем [2 - 4].

Проведенный онтологический инжиниринг предметной области данных исследований позволил выявить основные взаимосвязи базовых понятий [6] и разработать систему онтологий, включающую несколько уровней. При этом использовался авторский фрактальный подход [8] к структурированию знаний, который предполагает построение взаимосвязанных разномасштабных информационных объектов с использованием общих принципов и методов. На первом, верхнем уровне представлена метаонтология, отражающая взаимосвязи энергетики и геоэкологии.

Онтологии следующих, более детальных уровней, отражают этапы производственного процесса в энергетике и связанные с ними антропогенные факторы влияния на природную среду. На рис. 2 показана онтология процесса преобразования топлива, которая показывает зависимость выделенных антропогенных факторов (отходов и выбросов) от вида и качества топлива, технологии его сжигания и применяемой технологии очистки.

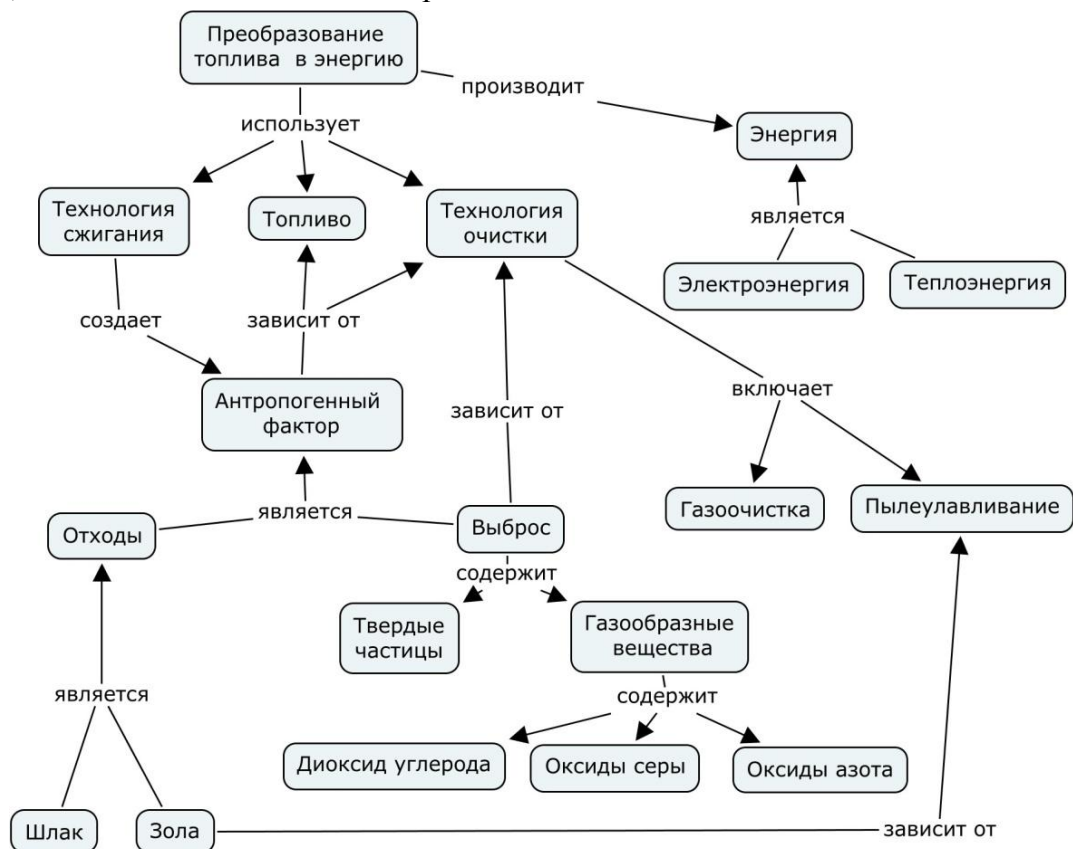


Рис. 2. Онтология этапа преобразования топлива

На уровне описания типовых энергетических объектов онтологии отражают свойства объекта, тип объекта (тепловая электростанция, дизельная электростанция, котельная), а также взаимосвязи с видом производимой энергии, типом топлива и характеристиками используемого оборудования (рис. 3).

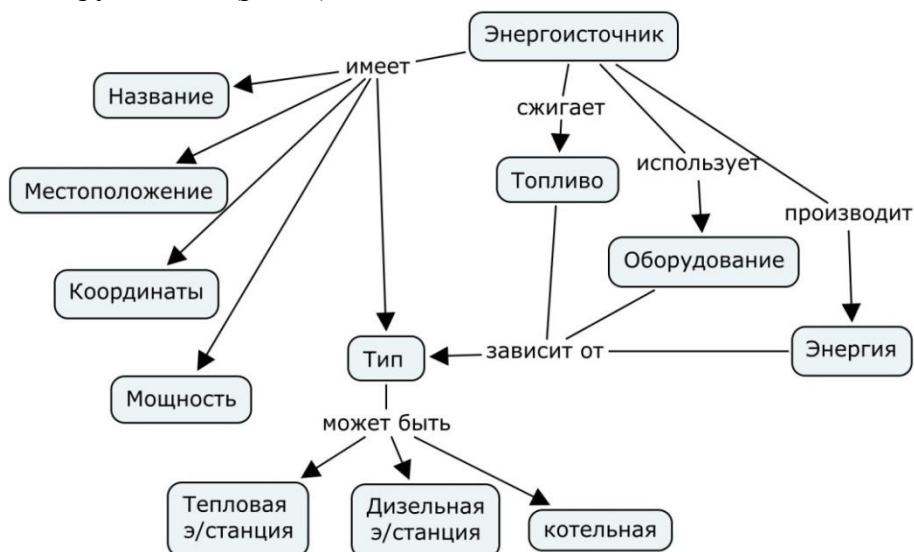


Рис. 3. Онтология типового энергетического объекта

На следующем, еще более детальном уровне онтологической модели энергообъекта, отражены основные технические характеристики и производственные показатели конкретных типов энергообъектов, численные значения которых хранятся в базе данных. Онтология показателей тепловой электростанции представлена на рис. 4.

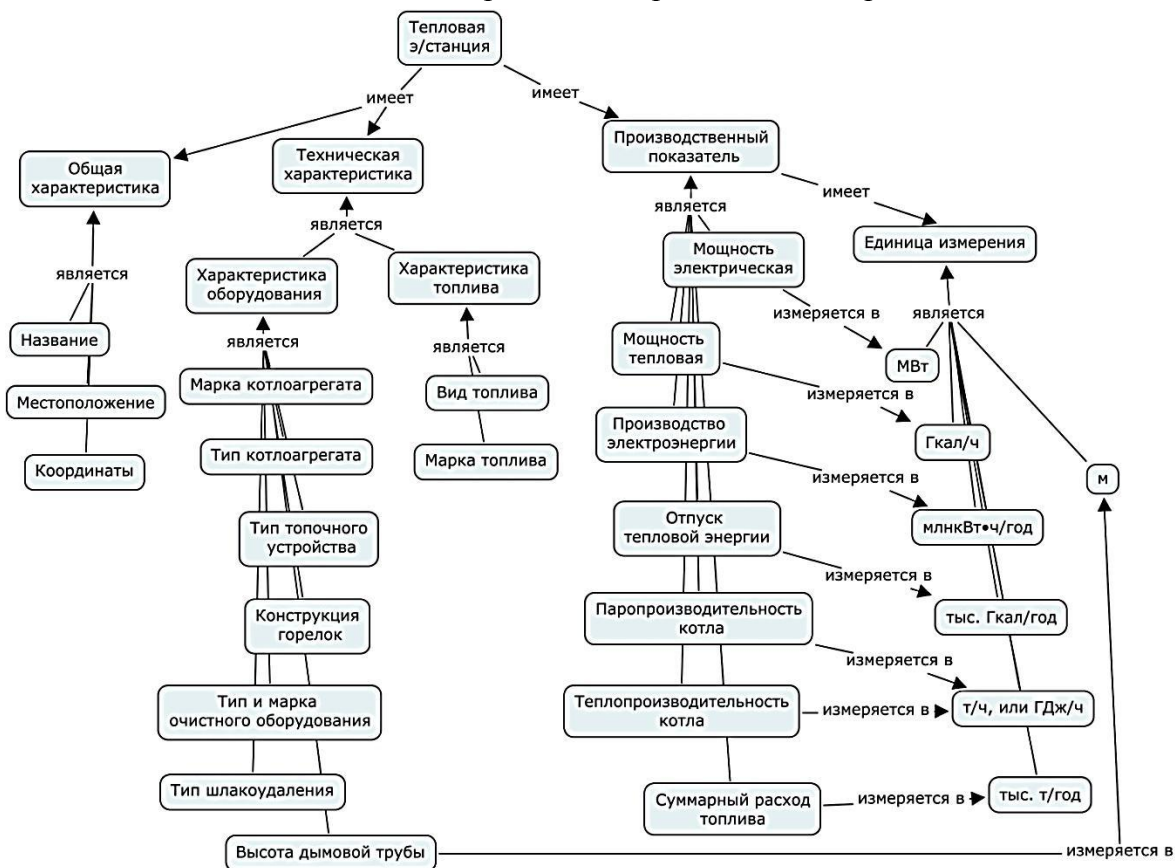


Рис. 4. Онтология тепловой электростанции

Разработанную систему онтологий предлагается использовать для формирования баз данных и знаний. Онтологические модели рассматриваются как прообраз инфологических моделей данных, на основе которых проектируются базы данных и базы знаний.

4. Построение инфологической модели данных на основе онтологий. Процесс разработки баз данных обычно начинается со стадии построения инфологической модели предметной области. Инфологическое моделирование также является способом построения семантической модели предметной области, то есть информационной модели наиболее высокого уровня абстракции. Инфологическая модель предметной области, предложена Петером Пин-ШенЧеном (PeterPin-ShanChen) в 1976 году [9, 17].

Основные элементы инфологической модели – это сущности и связи между ними:

- Сущность (Entity) – предмет, который может быть идентифицирован некоторым способом, отличающим его от других предметов. Сущность соответствует некоторому классу однотипных объектов и предполагается, что в системе существует множество экземпляров данной сущности. Объект, которому соответствует понятие сущности, имеет набор характеристик, определяющих свойства данного представителя класса, называемых атрибутами.
- Связь (Relationship) – ассоциация, устанавливаемая между сущностями. Связь может существовать между двумя разными сущностями или между сущностью и

ею же самой (рекурсивная связь). Она показывает, как связаны экземпляры сущностей между собой.

Диаграмма «сущность-связь» позволяет в графическом виде представить инфологическую модель предметной области. Такая модель разрабатывается без ориентации на какую-либо конкретную СУБД.

Инфологическая модель создается по результатам проведения исследований предметной области и используется для построения логической и физической моделей данных. Физическая модель данных непосредственно используется для построения базы данных.

Для разработки требуемой базы данных предлагается следующая методика построения инфологической модели, базирующаяся на анализе онтологий:

1. Анализ онтологий и выбор терминов, описывающих объекты предметной области, т.е. сущностей на языке инфологической модели.
2. Анализ отношений между выбранными терминами, выбор иерархических и структурных отношений, таких, как «состоит из», «входит», «является частью», «включает», «делится на» и др.
3. Анализ отношений между выбранными терминами и выбор таких, которые связывают сущность с ее свойствами, например, «описывает», «характеризует».
4. Выбор концептов в онтологиях, описывающих свойства объекта, т.е. являющихся признаками или атрибутами соответствующих сущностей.
5. Анализ свойств объекта и выбор таких, которые однозначно идентифицируют сущности, т.е. ключевых атрибутов.

На рис. 5 представлен фрагмент инфологической модели, основанной на анализе онтологий, описывающих энергообъекты трех типов (тепловые электрические станции, дизельные электростанции и котельные).

Заключение. В статье рассмотрен онтологический подход к построению базы данных для оценки влияния энергетики на геоэкологию региона. Проектируемая база данных является частью разрабатываемой Web-ориентированной информационной системы, которая интегрирует математические, семантические методы, инструментальные средства. Используемый системный подход предполагает учет всех взаимосвязанных процессов производства от добычи топливно-энергетических ресурсов до передачи тепловой и/или электрической энергии конечному потребителю. Примененный в рамках системного подхода семантический анализ предметной области и построенная система онтологий позволили структурировать имеющуюся информацию, построить инфологическую модель предметной области и спроектировать соответствующую базу данных.

Результаты получены в рамках выполнения проекта по госзаданию №АААА-А17-117030310444-2, отдельные аспекты прорабатывались в рамках проектов, поддержанных грантами РФФИ № 18-57-81001 и № 19-07-00351.

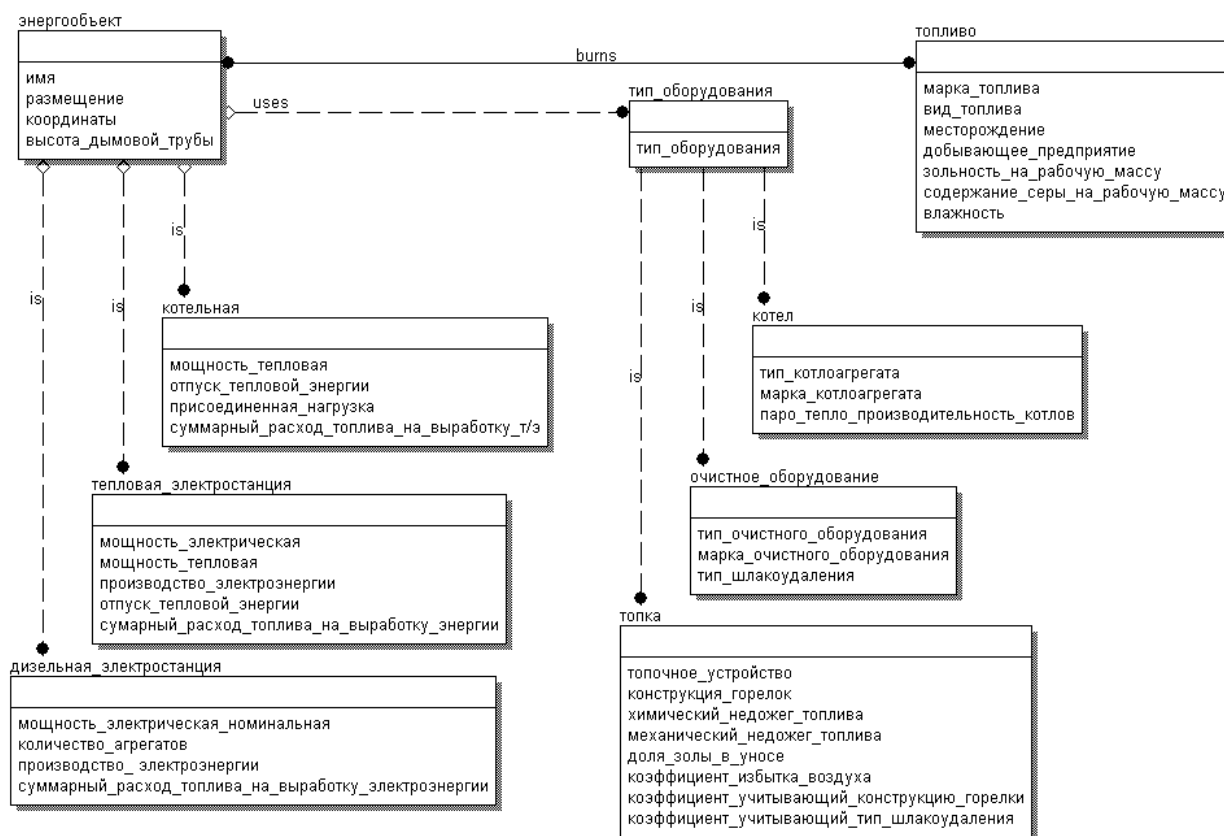


Рис. 5. Инфологическая модель описания энергообъектов

СПИСОК ЛИТЕРАТУРЫ

1. Государственный доклад «О состоянии и об охране окружающей среды Российской Федерации в 2017 году». Режим доступа: http://www.mnr.gov.ru/docs/o_sostoyanii_i_ob_okhrane_okruzhayushchey_sredy_rossiyskoy_federatsii/gosudarstvennyy_doklad_o_sostoyanii_i_ob_okhrane_okruzhayushchey_sredy_rossiyskoy_federatsii_v_2017/ (дата обращения: 10.04.2019)
2. Калиниченко Л.А. Эффективная поддержка баз данных с онтологическими зависимостями: реляционные языки вместо дескриптивных логик // Программирование. 2012. № 6. С. 45–62.
3. Когаловский М.П. Метаданные в компьютерных системах // Программирование. 2013. № 4. С. 23–46.
4. Когаловский М.П. Системы доступа к данным, основанные на онтологиях // Программирование. 2012. № 4. С. 55–77.
5. Майсюк Е.П., Иванова И.Ю. Анализ существующих методов оценки воздействия энергетических объектов на окружающую среду // Информационные и математические технологии в науке и управлении. 2018. №4 (12). С. 113–127. DOI: 10.25729/2413-0133-2018-4-12.
6. Массель Л.В., Иванова И.Ю., Ворожцова Т.Н., Майсюк Е.П., Ижбулдин А.К., Зорина Т.Г., Барсегян А.Р. Онтологические аспекты исследования взаимовлияния энергетики и геоэкологии // Онтология проектирования. 2018. Т.8. №4 (30). С. 550–561. DOI: 10.18287/2223-9537-2018-8-4.

7. Массель Л.В. Проблема оценки влияния энергетики на геоэкологию региона: постановка и пути решения // Информационные и математические технологии в науке и управлении. 2018. №2 (10). С. 5–21. DOI:10.25729/2413-0133-2018-2-01.
8. Массель Л.В. Фрактальный подход к структурированию знаний и примеры его применения // Онтология проектирования. 2016. №2(20). С. 149–161. DOI: 10.18287/2223-9537-2016-6-2-149-161.
9. Петер Пин-Шен Чен. Модель «сущность-связь» — шаг к единому представлению о данных / пер. с англ. М.Р. Когаловский // Системы управления базами данных. 1995. № 03.
10. A Direct Mapping of Relational Data to RDF / W3C Working Draft 24 March 2011. Available at: <http://www.w3.org/TR/rdb-direct-mapping> (accessed 10.04.2019)
11. A. Poggi, M. Ruzzi. Ontology-based data access with MASTRO. Available at: http://www.webont.org/owled/2007/PapersPDF/submission_33.pdf (accessed 11.04.2019)
12. Dejing Dou, Han Qin, Paea Lependu. Ontograte: towards automatic integration for relational databases and the semantic web through an ontology-based framework // International Journal of Semantic Computing. Vol. 4, № 1 (2010). Pp. 123–151.
13. D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, A. Poggi, and R. Rosati. Ontology-based database access // Proc. of SEBD 2007. 2007. Pp. 324–331.
14. Gruber T.R. A Translation Approach to Portable Ontology Specifications // Knowledge Acquisition. 1993. Vol. 5. № 2. Pp. 199–220.
15. Gruber T.R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing // International Journal Human-Computer Studies. Vol. 43. Pp. 907–948.
16. Guarino N. Formal Ontology in Information Systems: Proceedings of the 1st International Conference, Trento, Italy. FOIS. 1998. Vol. 46.
17. Peter Pin-Shan Chen. The Entity-Relationship Model-Toward a Unified View of Data // ACM Transactions on Database Systems. Volume 1. Number 1. 1976.

UDK 004.8

AN ONTOLOGICAL APPROACH TO THE DESIGN OF DATABASE FOR ASSESSING THE IMPACT OF ENERGY SECTOR ON THE ENVIRONMENT

Tatiana N. Vorozhtsova

PhD., main engineer, e-mail: tnn@isem.irk.ru

Nadezhda N. Makagonova

PhD., main specialist, e-mail: mak@isem.irk.ru

Liudmila V. Massel

Dr., Professor, Chief Researcher, Head of the Laboratory, e-mail: massel@isem.irk.ru

The laboratory of information technologies in power engineering

Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences

130, Lermontov Str., 664033, Irkutsk, Russia

Abstract. Currently, the problem of the energy sector influence on the environment as one of the anthropogenic factors is obvious. As part of studies to assess the impact of energy sector on the geocology of the region, it is proposed to develop a Web-based information system integrating mathematical and semantic methods, instrumental tools, a database and a knowledge base of the subject area under consideration. The article discusses the proposed ontological approach to the structuring and integration of data and knowledge necessary for research. Ontologies are proposed to be considered as a prototype of an infological data model when designing a database and knowledge base.

Keywords: geocology, anthropogenic factor, infological model, semantic modeling, ontology, metaontology, database.

References

1. Gosudarstvennyy doklad «O sostoyanii i ob okhrane okruzhayushchey sredy Rossiyskoy Federatsii v 2017 godu» [State report "On the state and environmental protection of the Russian Federation in 2017"]. Available at: http://www.mnr.gov.ru/docs/o_sostoyanii_i_ob_okhrane_okruzhayushchey_sredy_rossiyskoy_federatsii/gosudarstvennyy_doklad_o_sostoyanii_i_ob_okhrane_okruzhayushchey_sredy_rossiyskoy_federatsii_v_2017/ (accessed 10.04.2019) (in Russian)
2. Kalinichenko L.A. Effektivnaya podderzhka baz dannykh s ontologicheskimi zavisimostyami: relyatsionnyye yazyki vmesto deskriptivnykh logik [Effective support of databases with ontological dependencies: relational languages instead of descriptive logics.] // *Programmirovaniye = Programming*. 2012. № 6. Pp. 45–62. (in Russian)
3. Kogalovsky M.P. Metadannyye v komp'yuternykh sistemakh [Metadata in computer systems] // *Programmirovaniye = Programming*. 2013. № 4. Pp. 23–46. (in Russian)
4. Kogalovskiy M.P. Sistemy dostupa k dannym, osnovannyye na ontologiyakh [Ontology-based data access systems] // *Programmirovaniye = Programming*. 2012. No. 4. Pp. 55–77. (in Russian)
5. Maysyuk Ye.P., Ivanova I.YU. Analiz sushchestvuyushchikh metodov otsenki vozdeystviya energeticheskikh ob'yektov na okruzhayushchuyu sredu [Analysis of existing methods for assessing the impact of energy facilities on the environment] // *Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii = Information and mathematical technologies in science and management*. 2018. №4 (12). Pp. 113–127. DOI: 10.25729/2413-0133-2018-4-12 (in Russian)
6. Massel' L.V., Ivanova I.YU., Vorozhtsova T.N., Maysyuk Ye.P., Izhbuldin A.K., Zorina T.G., Barsegyan A.R. Ontologicheskiye aspekty issledovaniya vzaimovliyaniya energetiki i geokologii [Ontological aspects of the study of the interaction of energy and geocology] // *Ontologiya proyektirovaniya = Ontology of Designing*. 2018. V.8. №4 (30). Pp. 550–561. DOI: 10.18287/2223-9537-2018-8-4. (in Russian)
7. Massel' L.V. Problema otsenki vliyaniya energetiki na geokologiyu regiona: posta-novka i puti resheniya [The problem of assessing the impact of energy on the regional geo-ecology: formulation and solutions] // *Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii = Information and mathematical technologies in science and management*. 2018. №2 (10). Pp. 5–21. DOI:10.25729/2413-0133-2018-2-01 (in Russian)

8. Massel' L.V. Fraktal'nyy podkhod k strukturirovaniyu znaniy i primery yego pri-meneniya [Fractal approach to knowledge structuring and examples of its application] // *Ontologiya proyektirovaniya = Ontology of Designing*. 2016. №2(20). Pp. 149–161. DOI: 10.18287/2223-9537-2016-6-2-149-161. (in Russian)
9. Peter Pin-Shen Chen. Model' «sushchnost'-svyaz'» – shag k yedinomu predstavleniyu o dannykh / per. s angl. M.R. Kogalovskiy [Model "entity-relationship" - a step to a single view of the data] // *Sistemy upravleniya bazami dannykh = Database management systems*. 1995. № 03. (in Russian)
10. A Direct Mapping of Relational Data to RDF / W3C Working Draft 24 March 2011. Available at: <http://www.w3.org/TR/rdb-direct-mapping> (accessed 10.04.2019)
11. A. Poggi, M. Ruzzi. Ontology-based data access with MASTRO. Available at: http://www.webont.org/owled/2007/PapersPDF/submission_33.pdf (accessed 11.04.2019)
12. Dejing Dou, Han Qin, Paea Lependu. Ontograte: towards automatic integration for relational databases and the semantic web through an ontology-based framework // *International Journal of Semantic Computing*. Vol. 4. № 1 (2010). Pp. 123–151.
13. D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, A. Poggi, and R. Rosati. Ontology-based database access // *Proc. of SEBD 2007*. 2007. Pp. 324–331.
14. Gruber T.R. A Translation Approach to Portable Ontology Specifications // *Knowledge Acquisition*. 1993. Vol. 5. № 2. Pp. 199–220.
15. Gruber T.R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing // *International Journal Human-Computer Studies*. Vol. 43. Pp. 907–948.
16. Guarino N. Formal Ontology in Information Systems: Proceedings of the 1st International Conference, Trento, Italy. FOIS. 1998. Vol. 46.
17. Peter Pin-Shan Chen. The Entity-Relationship Model-Toward a Unified View of Data // *ACM Transactions on Database Systems*. Volume 1. Number 1. 1976.