



**ДИСКУССИОННЫЙ КЛУБ**

УДК 007:681.3

**А. А. Запевалин**<sup>1</sup> канд. техн. наук, профессор-исследователь, **А. В. Гаевская**<sup>2</sup>, докт. биол. наук, проф., зав. отд.

<sup>1</sup> Микстекский Технологический Университет, Уахуапан де Леон, штат Оахака, Мексика  
(Universidad Tecnológica de la Mixteca, Huajuapán de León, Oaxaca, Mexico)

<sup>2</sup> Институт биологии южных морей им. А. О. Ковалевского Национальной академии наук Украины,  
Севастополь, Украина

**КОНЦЕПТУАЛИЗАЦИЯ ПАРАЗИТОЛОГИИ  
ДЛЯ РАЗРАБОТКИ ИНФОРМАЦИОННОЙ СИСТЕМЫ ПОДДЕРЖКИ ИССЛЕДОВАНИЙ**

Для построения универсальной информационной системы представления паразитологических данных и знаний рассмотрены особенности этапа концептуального проектирования. Представлены результаты онтологического анализа существующей терминологической системы, которые являются основой для построения концептуальной модели системы.

**Ключевые слова:** паразитология, информационные системы, проектирование, концептуальный анализ, онтологии, концепт

Разработка современных информационных систем (ИС) включает несколько этапов, каждый из которых представляет собой преобразование моделей:

модель реального мира (области исследований, предметной области) преобразуется в концептуальную модель – этап концептуального моделирования;

концептуальная модель преобразуется в модель логики компьютерного приложения – этап логического моделирования;

наконец, разработанная логическая модель в виде структур и форматов данных и отобранных алгоритмов преобразуется в физическую модель, которая и воплощается в конкретной операционной среде, на конкретном сочетании аппаратного и программного обеспечения.

В настоящее время преобразование модели реального мира в концептуальную прово-

дится на основе так называемых онтологий [3] – словарей основных концептов предметной области и смысловых связей между ними, принятыми сообществом исследователей.

Неотъемлемой частью любого исследования является накопление, сохранение и обработка данных об интересующих исследователя объектах и процессах в виде простейших записей или в обычных таблицах наблюдений. Применение для этих целей более сложных приложений типа СУБД влечет за собой разработку системы имен таблиц, полей и связей между ними, что ставит перед необходимостью концептуального проектирования базы данных. Разработка полномасштабной информационной системы для конкретной области исследований обязывает пройти все этапы, чтобы в дальнейшем избежать проблем с развитием, модификацией и ре-использовани-

ем накопленной информации в новых приложениях.

Хорошо разработанная концептуальная модель позволяет объединить уже накопленные архивы и базы данных и расширить возможности их использования в рамках экспертных систем и систем, базирующихся на знаниях, объединить усилия многих исследователей на базе современных интеллектуальных информационных систем – следующем эволютивном этапе развития информационных технологий.

Сущность концептуализации состоит в следующем. Носитель знаний использует терминологическую систему как базу для построения соответствующей концептуальной системы и обоснования ассоциативных связей между концептами. Эти связи должны рассматриваться как основа знаний, их представление в процессе передачи знаний составляет наибольшую трудность, поскольку сложнее всего поддается вербализации. Связи присутствуют в любых терминологических системах, но равнозначны для неопита. Для перехода к концептам необходимо оценить важность и предпочтительность терминов и связей между ними, «проложить рекомендованный маршрут» между концептами, построить их иерархию, а лучше – сеть взаимосвязей. Именно на основе такой системы концептов и связей между ними – онтологии предметной области, – проводится концептуализация.

Знание следует рассматривать как совокупность информационных фрагментов на сети смысловых связей. Можно считать, что связи между терминами, явные и неявные, очевидные и нет, общепринятые и спорные, составляют скелет знаний специалистов. Выявление и сохранение максимально возможного числа связей – цель онтологического этапа.

Для любой предметной области не существует единственно правильной онтологии. Проектирование онтологии – это творческий процесс и две онтологии, разработанные разными людьми, никогда не будут одинаковыми.

Потенциальные приложения онтологии, а также понимание разработчиком предметной области и его точка зрения на нее будут, несомненно, влиять на принятие решений при создании онтологии. “Цыплят по осени считают” – мы можем оценить качество нашей онтологии, только используя ее в приложениях, для которых мы ее разработали. Всесторонний анализ предметной области, выявление концептов и связей между ними из принятых терминологических систем, их обсуждение и верификация должны содействовать успешности разработки концептуальной модели. И только концептуальная модель как документ может обеспечить корректность логической, а впоследствии и физической модели проектируемой ИС.

Описание предметной области на основе соответствующей терминологической системы (ТС) является традиционным. Словари, тезаурусы, справочники – сертифицированные источники необходимой для построения любой онтологии информации. Сам же переход от ТС к онтологиям является в некоторой степени вынужденной мерой для эффективного представления модели реального мира в виде структур данных компьютерных систем. Эта вынужденность объясняется необходимостью формализации закладываемой в фундамент ИС информации.

Любая ТС не является формализованной моделью, в ней присутствуют синонимия и полисемия, она подвержена особенностям вербализации как единственного способа представления знаний специалистом. Неоднозначность ТС призвана удерживать многосмысленность реального мира. Онтология, в свою очередь, пытается отразить многообразие и многозначность, но на основе строгих формальных методов, в виде спецификаций на формальном языке, которые отличаются от описаний на естественном языке. Это делается для решения задач представления информации в цифровых

системах, с трудом поддерживающих много-смысленность реального мира. Создание ТС является эволютивным, историко-культурным процессом. Онтология не создает ничего нового, она просто пытается структурировать, формализовать существующее. Любая ТС принята для общения среди специалистов, онтология же служит для создания концептуальных моделей ИС. Онтология – это формальная спецификация концептуализации, дескриптивная форма в рамках специальных систем, типа Protégé-2000, OntoLingva, OntoEdit, и т.п. [4].

В данной работе рассмотрены вопросы, предшествующие формальной спецификации, внимание акцентируется на принципиальных неформальных аспектах.

В качестве исходного материала взята структура взаимосвязанных терминологических статей – терминологический справочник [1], включающий 2400 статей по различным аспектам интересующей нас предметной области – паразитологии.

Паразитология [2], как наука, и как экземпляр концепта **наука**, соотносится с двумя концептами высшего уровня – **процессом** и **объектом** (здесь и далее жирным шрифтом выделяются имена концептов). С одной стороны имеем процессы исследования, сбора данных, их накопления, анализа, обобщения, создания нового знания и его применения. С другой стороны – наука как **объект** – совокупность идей и данных, взаимосвязанных и согласованных, выраженных в текстах, моделях, базах данных и т.п.

При этом объект науки – реальный мир (физические и/или биологические системы со всеми их взаимосвязями) на базе абстрактной системы – науки, как множества идей, концепций и связей между ними, воплощается в цифровые системы баз данных и знаний, конечные наборы и структуры данных в памяти компьютеров.

**Паразитология** также может быть представлена как взаимосвязанная совокуп-

ность следующих высокоуровневых концептов:

- **объект (организм) и система объектов;**
- **явление и феномен;**
- **процесс, его состояния, условия и события;**
- **вещество и среда, их признаки и характеристики;**
- а также **субъекты** – исследователи и организации, модели, методы и инструменты, ими используемые, и, конечно же, результаты, например, статьи (см. рис. 1. – связи не показаны, ввиду их очевидности и многочисленности).

Рис. 1 Высокоуровневые концепты паразитологии  
Fig. 1. Top-level concepts of parasitology



Эти концепты могут рассматриваться как названия основных подсистем предполагаемой информационной системы, совокупность которых позволит охватить данные и знания разных типов из гетерогенных источников. Кроме того, они присутствуют во многих онтологиях высшего уровня известных и признанных сегодня онтологических систем (WordNet, Cyc, DOLCE, SUO и др. [4 – 6]), что обеспечивает преемственность, согласованность и ре-использование проектных решений.

Центральный концепт паразитологии как предметной области – **паразит** (синоним – антиген) – прямо или опосредованно взаимосвязан со всеми остальными концептами, такими как **организм**, **гидробионт**, **экосистема**, **среда обитания**, **хозяин**, **пищевые цепи (источник пищи)** и т. п. Концепт определяет таксономическую систематику всех известных паразитов, которые служат так называемыми примерами или экземплярами концепта (термины, принятые в онтологическом анализе).

Концепт **организм** и его специфическое проявление (пример) **гидробионт** связан с концептами **паразит** и **хозяин**, **внешняя среда** и **среда обитания** как **системы (биологические)**, **жизненный цикл** как **процесс**, **функции** организма и **состояния** процессов функционирования и взаимодействия. Синонимы концепта – особь, индивидуум, живое существо (см. рис. 2).



Рис. 2. Концепты паразит и хозяин  
Fig. 2. Parasite and host concepts

Концепт **хозяин**, как **организм** и как **гидробионт**, подобно концепту **паразит**, имеет собственную таксономическую иерархию. Между таксономиями **хозяев** и **паразитов** может быть описано **отношение**, порождающее или определяющее существование паразито-хозяинных **систем**.

Термин **среда обитания** для образования концепта должен быть переформулирован как «**система условий**» взамен «совокупности» для отражения особенностей взаимодействия биотических и абиотических условий. Ранжирование **среды обитания** (рис. 3) определяет

концепты **особь**, **популяция**, **вид**, **род**, **семейство**, как субклассы концептов **организм** и **экосистема**. Основой подобного ранжирования служат и концепты **паразит** и **хозяин**.

Как подкласс концепта **система**, концепт **биосистема** и ее пример **паразито-хозяинная система** опирается на концепты **процесс** и **отношение** в их специфических разновидностях, таких как **болезни** и **заболевания**, **подавление паразита** или **угнетение хозяина**.

В свою очередь, концепт **экосистема** базируется на концептах **явление** и **процесс** в их биотическом и абиотическом проявлениях. Системы биотических и абиотических факторов описывают разнообразие биосистем.



Рис. 3. Отношения концепта среда обитания  
Fig. 3. Concept Environment relations

Концепт **процесс** (и/или система процессов) представлен в данной онтологии такими примерами (целесообразность их рассмотрения в качестве концептов нижнего уровня следует изучить отдельно), как **жизненный цикл** (различных типов, с фазами и их длительностями), **метаболизм**, уже упоминавшиеся **болезнь** и **заболевание**, иммунитет, аллергия, и проявления других **функций** (и **дисфункций**) живого организма. Примерами

процессов служат взаимодействия веществ (абсорбция), взаимодействия систем (аварийный сброс), взаимодействия между организмами и многие другие.

Концепты **болезнь** и **заболевание** (**процессы**) могут быть использованы как синонимы, но в общем случае требуют разделе-

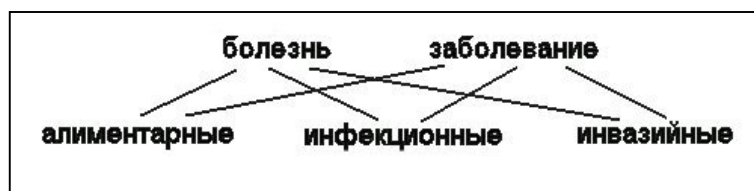


Рис. 4 Концепты болезнь и заболевание  
Fig. 4. Illness and disease concepts

Существенным, с точки зрения заявленной предметной области, будет уточнение концепта **болезнь** до концепта **болезнь рыб**, с его опорой на концепты **состояние** (процесса функционирования организма) и собственно **процесс** (физиологического функционирования). Концепт **возбудитель** (организм или процесс) определяет таксономию болезней, включающую и инвазийные, т.е. вызываемые **паразитами**.

Концепт **пищевые цепи** – это и **процесс** преобразования **веществ**, неотъемлемая часть функций жизнедеятельности **организма**, **гидробионта**, и **феномен**, определяющий существование **биосистем**.

Особь, как пример организма состоит из **тканей**, **органов**, **отделов**, **систем** (нервной, пищеварительной...), т.е. из частей **тела**, в границах которого взаимодействуют **вещества**

ния, где **болезнь** – понятие более широкое, включающее изменения **поведения (процесс)** и эмоциональные отклонения. Дальнейшая концептуализация приводит к общеизвестной таксономии болезней и заболеваний по их возбудителям (см. рис. 4).

(ферменты, витамины, продукты питания и распада) и паразитирующие **организмы**; происходят **процессы** – физиологические, химические, физические...) и их частные проявления – **реакции**.

Наконец, концепт **субъект** во взаимосвязи с изучаемыми примерами других концептов онтологии может быть описан общепринятыми биографическими (исследователи и организации), библиографическими (книги, статьи и т.п.) или фактологическими (модели, методы, инструменты...) системами.

Все вышеизложенное позволяет перейти к следующему этапу проектирования – верификации построенной онтологии и переход к концептуальной модели на основе ER-диаграммам, с их последующей нормализацией и воплощением в структуры взаимосвязанных таблиц конкретной СУБД.

1. Гаевская А. В. Паразитология и патология рыб: Энциклопедический словарь-справочник. – М.: Изд-во ВНИРО, 2003. – 231 с.
2. Гаевская А. В., Мачкевский В. К. Проблемы морской паразитологии Азово-Черноморского бассейна. Концептуальный подход // Экология моря. – 2001. – Вып. 57. – С. 36 – 43.
3. Schulze-Kremer S. Ontologies for molecular biology and bioinformatics // Silico Biology. – 2002. – 2, 3. – P. 179 – 193.
4. Mizoguchi R. Tutorial on ontological engineering - Part 2: Ontology development, tools and languages

// New Generation Computing. - OhmSha & Springer, 2004. – 22, No.1. – P. 61 – 96.

5. Doerr M., Hunter J., Lagoze C. Towards a Core Ontology for Information Integration. // J. Digital Information. – 2003. – 4, 1. - Article No. 169.
6. Kingston J., Merging Top Level Ontologies for Scientific Knowledge Management // Proc. AAAI-02 Workshop, August 2002.

Получено 24 октября 2005 г.

**Концептуалізація паразитології для розробки інформаційної системи підтримки досліджень. О. О. Запевалін, А. В. Гаєвська.** Для побудови універсальної інформаційної системи представлення паразитологічних даних і знань розглянуті особливості етапу концептуального проектування. Представлено результати онтологічного аналізу існуючої термінологічної системи, що є основою для побудови концептуальної моделі системи.

**Ключові слова:** паразитологія, інформаційні системи, проектування, концептуальний аналіз, онтології, концепт

**Conceptualization of Parasitology for development of the information system of the researches support. A. A. Zapevalin, A. V. Gaevskaya.** The special features of the stage of conceptual design for constructing of the universal information system of representation of parasitological data and knowledge are examined. The results of ontological analysis of the existing terminological system, which are basis for constructing the conceptual model of system, are represented.

**Key words:** parasitology, information system, constructing, conceptual design, ontology, concept