

УДК 519.711.3

## ОНТОЛОГИЯ ИНТЕГРАЦИИ ЗНАНИЙ НА ОСНОВЕ ТЕХНОЛОГИИ IPSE В ИНТЕЛЛЕКТУАЛЬНОЙ СРЕДЕ «ОБЛАЧНОЙ» МОДЕЛИ

А.В. Бухановский<sup>1</sup>, С.В. Иванов<sup>2</sup>, С.В. Ковальчук<sup>3</sup>, Ю.И. Нечаев<sup>4</sup>

Санкт-Петербургский национальный исследовательский университет информационных технологий,  
механики и оптики

<sup>1</sup>avb\_mail@mail.ru, <sup>2</sup>svivanov@mail.ifmo.ru, <sup>3</sup>kovalchuk@mail.ifmo.ru, <sup>4</sup>nechaev@mail.ifmo.ru

### Аннотация

Рассматривается формализованная модель онтологической системы интеграции знаний на основе технология iPSE (*Intelligent Problem Solving Environment* – iPSE). Модель реализована в рамках концепции открытых систем в интеллектуальной среде «облачных» вычислений. Разработанная модель онтологии учитывает расширение функциональных возможностей технологии iPSE за счет модификации потока событий (*WorkFlow*). Указаны направления практических приложений разработанной онтологической системы знаний при реализации концепции представления и обработки информации в задачах контроля экстремальных ситуаций в сложной динамической среде.

**Ключевые слова:** онтология интеграции знаний, интеллектуальная технология, «облачная» модель, поток событий, экстремальная ситуация, сложная динамическая среда.

### Введение

Разработка интеллектуальных приложений на базе высокопроизводительных вычислений представляет собой одно из перспективных направлений при создании систем поддержки принятия решений (ППР) [1-5]. Концепция iPSE (*Intelligent Problem Solving Environment*) [1] определяет принципы построения инструментальных технологических платформ для разработки композитных приложений в форме интеллектуальной оболочки управления параллельными вычислительными процессами в распределенной иерархической среде, включающей в себя вычислительные системы различных архитектур [2]. Такой подход расширяет достоинства традиционных проблемно-ориентированных инструментальных оболочек (PSE, *Problem Solving Environment*) [4] обеспечения эффективного (с точки зрения производительности) исполнения композитных приложений за счет использования для управления распределенными вычислениями интегрированных баз знаний об особенностях предметной области и специфике вычислительного процесса.

В рамках iPSE на основе программных кодов формализуются не только методы и вычислительные алгоритмы, но и экспертные знания об организации процесса вычислений применительно к специфике предметной области. Технология iPSE реализует функции интеллектуальной системы (ИС) ППР при разработке композитных приложений, предоставляя единый интерфейс взаимодействия для предметно-ориентированных программных модулей и компонентов, которые могут разрабатываться различными коллективами на разных языках и иметь различные условия распространения и использования. Реализация технологии iPSE ориентирована не только на поддержку высокопроизводительных вычислений для суперкомпьютерных систем с традиционной (кластерной) архитектурой, но и для неоднородных систем, включая гиперкластеры (в рамках модели метакомпьютинга) и Грид. При этом управление эффективностью выполнения сценария вычислений, задаваемого потоком дан-

ных (*WorkFlow*), позволяет избежать конфликтных ситуаций при разделении ресурсов между различными вычислительными модулями и пользователями.

## 1 Концептуальная модель онтологии iPSE

Концептуальные основы создания платформ класса iPSE базируются на теоретических принципах, определяющих архитектуру системы и уровни ее управления. Наряду с традиционными математическими методами построения расписаний здесь широко применяются достижения искусственного интеллекта (ИИ), определяющие принципы функционирования систем, основанных на знаниях. Среди них важная роль принадлежит принципу открытости, позволяющему обеспечить наиболее сложные уровни иерархической структуры системы – самоорганизацию и самообучение. В результате открываются возможности интерпретации сложных процессов при принятии решений, а также при моделировании действий системы в процессе решения задачи и «обучении» на своем опыте. На рисунке 1 приведена концептуальная схема онтологии iPSE, определяющая основные принципы функционирования.

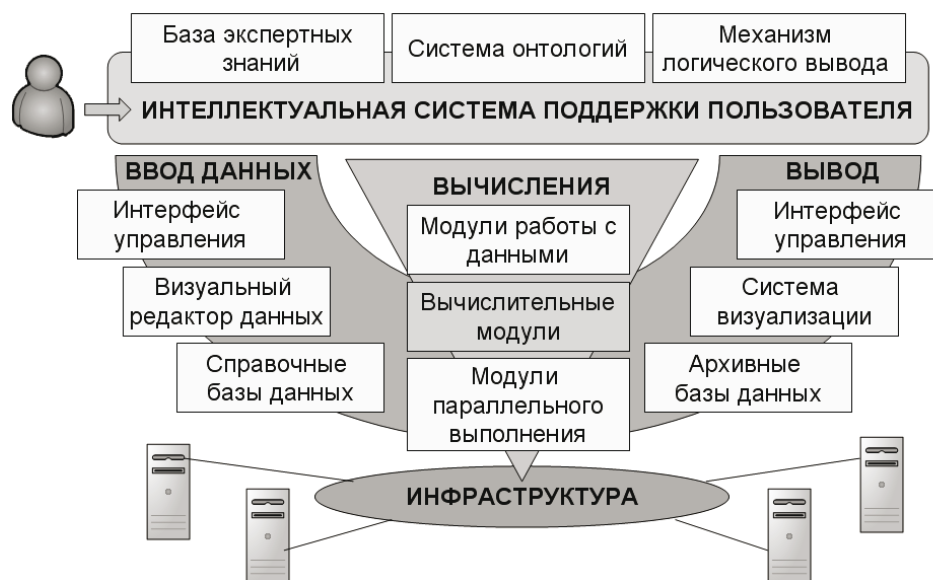


Рисунок 1 – Концептуальная модель онтологии iPSE

Из рисунка 1 видно, что пользователь взаимодействует с многопрофильной интеллектуальной технологической платформой разработки интеллектуальных приложений через ИС ППР, которая обеспечивает возможность формирования приложения на языке предметной области, а также рациональное использование инструментария ввода и вывода данных, визуализации, хранения и пр. Предметно-ориентированные вычислительные модули, представленные соответствующим репозиторием, являются основной частью системы, однако построение композитных приложений и параллельного исполнения на распределенной вычислительной инфраструктуре реализованы независимо и являются прерогативой самой системы, гибко учитывающей возможности внутреннего распараллеливания, заложенные в каждом из компонентов.

## 2 Архитектура и онтология информационных потоков в iPSE

На рисунке 2 приведена онтология основных информационных потоков в iPSE. Структура содержит подсистемы логического вывода, управления знаниями, человеко-

компьютерного взаимодействия (ЧКВ), оболочку параллельного исполнения вычислительных модулей, хранилище данных, информационный портал.

Подсистема ЧКВ предоставляет пользователю диалоговый интерфейс, который включает в себя модуль интервьюирования пользователя, конструктор сценариев разрабатываемых композитных приложений и интеллектуальный редактор данных, посредством которого формируется входная информация для расчетов с использованием понятийной базы и справочной информации предметной области. В состав интерфейса входит интеллектуальный инструктор – модуль, отслеживающий и анализирующий действия пользователя с целью возможной коррекции или оптимизации. Модуль научной визуализации и модуль валидации и верификации результатов расчетов на основе знаний предметной области формируют тестовые задания для оценки работоспособности композитных приложений.

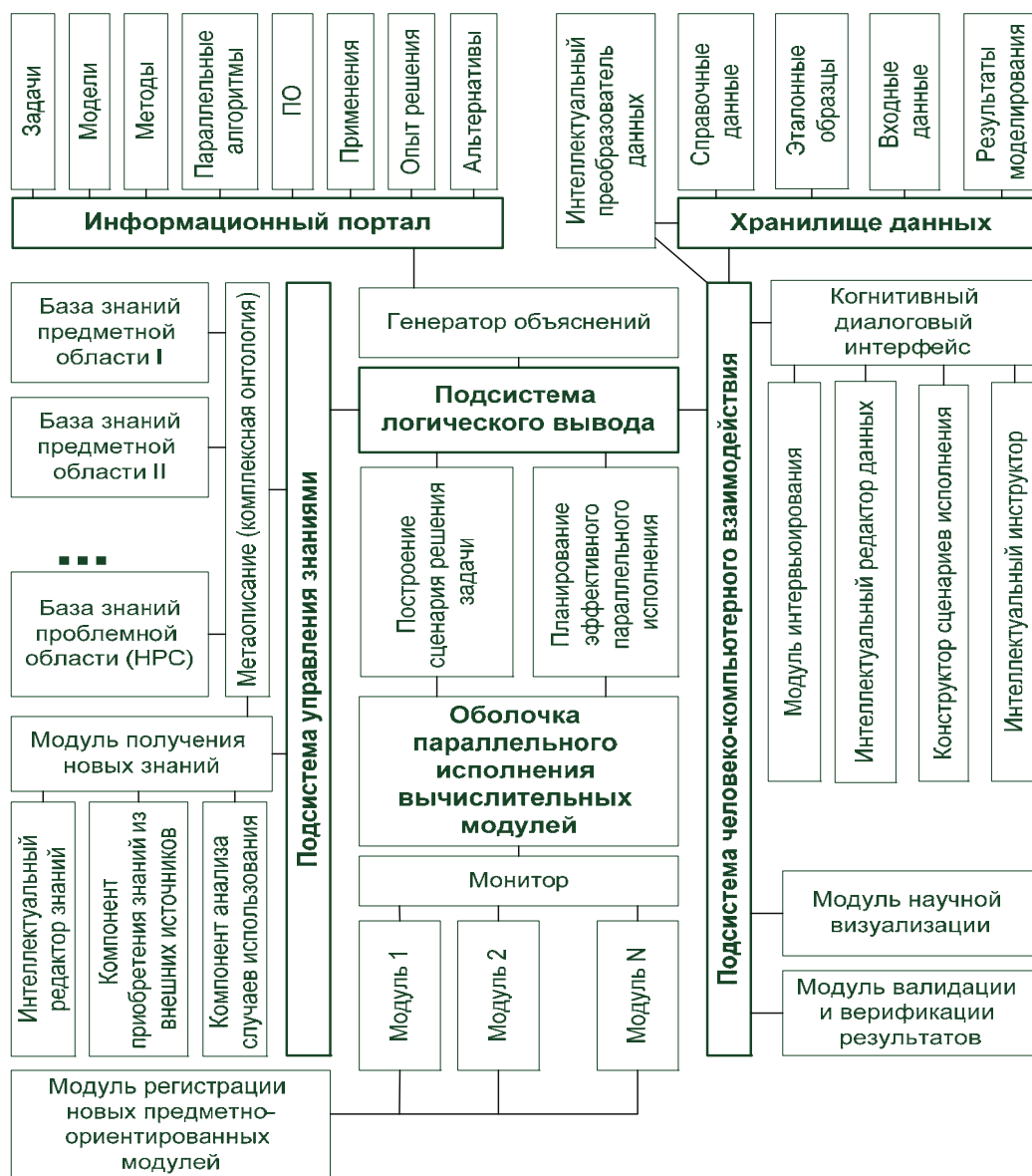


Рисунок 2 – Онтология программной архитектуры инструментальной технологической платформы iPSE

Подсистема управления знаниями включает в себя набор баз знаний предметной и проблемной области. Взаимоотношения между понятиями из разных баз знаний регламентиру-

ются посредством метаописания их структуры и состава в форме комплексной онтологии. В рамках онтологии определены формы представления параметров, включая вид и структуру моделей параллельной производительности. Адаптивные функции данной подсистемы реализуются посредством модуля получения новых знаний.

Подсистема логического вывода является основным механизмом интеллектуальной оболочки, реализующей вывод на знаниях, ассоциированных с основными программными компонентами в составе управляющего ядра программного комплекса. Она включает в себя компоненты построения сценариев решения задачи на основе знаний предметной области и планировщик параллельного исполнения, использующий знания о производительности. На основе этих знаний подсистема интерпретирует задание (метаописание, предоставленное пользователем через систему ЧКВ) и путем формирования активных фактов предметной области определяет набор вычислительных сервисов и сценарий их взаимодействия.

Оболочка параллельного исполнения вычислительных модулей является основным содержательным элементом инструментальной платформы. Она представляет собой самостоятельную систему распределенных вычислений, которая обеспечивает распределенный запуск и мониторинг исполнения заданий в распределенной среде. В ходе работы компонентов блока транслируются управляющие команды, поступающие от ИС конструирования заданий. Оболочка параллельного исполнения осуществляет унифицированный доступ к вычислительным предметно-ориентированным модулям в составе комплекса. Каждый такой модуль помимо процедурной части (программного модуля) имеет декларативную составляющую, содержащую информацию о математических моделях, методах, алгоритмах и условиях применения, как фрагмент соответствующей базы знаний. Процедурная, и декларативная составляющие компонента могут храниться на разных узлах распределенной вычислительной системы.

Хранилище данных связано с остальными подсистемами посредством преобразователя данных. Интеллектуальная функция этого модуля реализуется путем преобразования форматов входных и выходных файлов для различных модулей и обеспечения их эквивалентности на уровне решаемых задач. Хранилище данных содержит собственно входные и выходные данные (результаты моделирования), справочные данные, а также эталонные образцы, необходимые для решения задач верификации и валидации.

Информационный портал является специфической подсистемой iPSE, которая заменяет собой традиционное «Руководство пользователя». Портал представляет собой динамическую гипертекстовую информационную систему с перекрестной классификацией содержания по следующим категориям понятий: задачи, модели, методы, алгоритмы, программное обеспечение, применение, опыт использования и альтернативы. Пользователь может осуществлять навигацию в различных направлениях, в итоге получая рекомендации по конкретным возможностям комплекса для решения собственной задачи. Дополнительным компонентом является генератор объяснений, который связывает действия, выполняемые в процессе логического вывода, с содержимым информационного портала, что позволяет представить пользователю аргументированную информацию обо всем процессе рассуждений системы.

### **3 Формальная модель и уровни описания онтологии предметно-ориентированных сервисов**

Формально слой классов онтологии определяется как граф

$$(1) \quad O = \langle C, R \rangle,$$

где  $C$  – множество классов,  $R$  – множество абстрактных отношений, связывающих классы. Аналогично слой индивидов онтологии определяется как граф

$$(2) \quad \tilde{O} = \langle \tilde{C}, \tilde{R} \rangle,$$

где  $\tilde{C}$  – множество индивидов, а  $\tilde{R}$  – множество отношений между индивидами. При этом для каждого элемента слоя индивидов определены:

а) отношение генерализации

$$(3) \quad gn^{(C)} : \tilde{C} \rightarrow C, \quad gn^{(R)} : \tilde{R} \rightarrow R,$$

определяющее связь индивидов и связей между ними с соответствующими классами и связями классов;

б) «сторожевые условия» (*guard condition*), определяющие применимость элементов в данных условиях

$$(4) \quad gc^{(C)}(F) : \tilde{C} \rightarrow \{0,1\}, \quad gc^{(R)}(F) : \tilde{R} \rightarrow \{0,1\},$$

где  $F$  – множество активных фактов, определенных для текущей задачи;

в) функция критериальной оценки

$$(5) \quad k^{(C)}(F) : \{ \tilde{c} \in \tilde{C} \mid gc^{(C)}(\tilde{c}) = 1 \} \rightarrow \Psi^{(C)},$$

$$(6) \quad k^{(R)}(F) : \{ \tilde{r} \in \tilde{R} \mid gc^{(R)}(\tilde{r}) = 1 \} \rightarrow \Psi^{(R)},$$

где  $\Psi^{(C)}$  и  $\Psi^{(R)}$  – пространство критериев оценки индивидов и отношений между ними.

Следует отметить, что в общем случае «сторожевые условия» могут рассматриваться как один из критериев оценки. Тем не менее, в данном документе они выделены в отдельный класс в целях упрощения работы критериальной оценки, оперирующей множествами  $\tilde{C}$  и  $\tilde{R}$ .

Целью работы блока логического вывода является определение способа решения поставленной задачи. При этом способ решения определяется как кортеж

$$(7) \quad S = (s_1, s_2, \dots, s_N)$$

фиксированной структуры,  $i$ -м элементом которого являются множества вида

$$(8) \quad s_i = \{ \tilde{c} \in \tilde{C} \mid gn^{(C)}(\tilde{c}) = c_i \},$$

где последовательность классов  $c_i \in C$  и требований к множествам  $s_i$  определяет общую структуру решения. Для оценки построенного решения по системе критериев применяется анализ графа

$$(9) \quad \tilde{O}' = \langle \tilde{C}', \tilde{R}' \rangle : \quad \tilde{C}' = \bigcup_i s_i \cup \tilde{C}_s,$$

где  $\tilde{C}_s$  – присоединенная система классов

$$(10) \quad \tilde{C}_s = \left\{ \tilde{c}_s \mid \tilde{c}_s \notin \bigcup_i s_i, \exists \tilde{c}_1 \in \bigcup_i s_i : rch(\tilde{c}_s, \tilde{c}_1) \right\},$$

где  $rch(\tilde{c}_1, \tilde{c}_2)$  – отношение достижимости на графе  $\tilde{O}'$ . Оценка  $\tilde{O}'$  осуществляется в пространстве критериев  $\Psi$ , определяемом пересечением множеств критериев, описывающих пространства  $\Psi^{(C)}$  и  $\Psi^{(R)}$ . Задача ранжирования способов запуска требует определения отображения  $\psi: \Psi \rightarrow R$ , для вычисления интегрального критерия качества решения.

На практике принципы реализации онтологий в соответствии с рассмотренным формализмом базируются на подходах и методологии проектирования систем, основанных на знаниях. Среди принципов онтологии выделяют ясность, согласованность, расширяемость. Жизненный цикл онтологий включает процедуры управления приложениями, их разработку

и поддержание. Процесс построения онтологий распадается на серию подпроцессов. Сначала строится глоссарий терминов, затем деревья классификации концептов и диаграмма бинарных отношений и только после этого – остальные промежуточные представления. Формируя структуру онтологии, необходимо учитывать ряд особенностей, проистекающих как из требований объекта исследования (компаративными приложениями с динамической компонентной архитектурой на основе прикладных сервисов), так и определяемых необходимостью универсализации структуры онтологии в предположении о возможности ее адаптации для задач произвольной предметной области. При этом производится декомпозиция онтологии на несколько базовых уровней, различающихся по степени абстракции, с которой рассматривают анализируемое композитное приложение.

#### 4 Онтологии семантического описания прикладных сервисов в распределенной среде iPSE

Композитное приложение состоит из отдельных сервисов, *отчуждаемых* в процессе исполнения от своих разработчиков. Потому для их использования необходимо также приобретать и формализовывать знания об их использовании в предметной области. Для представления знаний о предметно-ориентированных сервисах в составе композитных приложений, формируемых в рамках концепции iPSE, используется модель комплексной онтологии. Являясь аналогом понятия «модель», онтология обеспечивает взаимодействие между программными компонентами и служит средством коммуникации между специалистами, имеющими различный взгляд на одни и те же проблемы. Благодаря формальной спецификации, унификации представления различных моделей и автоматизации проверки корректности, онтологии уже зарекомендовали себя в качестве интеллектуальной основы средств инструментальной поддержки разработки информационных систем. При использовании онтологического подхода в рамках iPSE строится нормативная база (формальная система), концепция которой учитывает принципы системного подхода и средства описания системных отношений (рисунок 3).



Рисунок 3 – Концептуальная модель представления знаний в рамках онтологии iPSE



Расширение семантики онтологической системы достигается на основе использования адаптивного алфавита, знакам которого ставятся в соответствие понятия на основе некоторого правила. В рамках iPSE онтология играет роль не только инструмента для структурирования знаний о предметной области, но и предоставляет основу для реализации логического вывода в рамках гибридного подхода, модифицированного с учетом рассматриваемой проблемы и поставленной цели – автоматизации построения эффективных высокопроизводительных композитных приложений сервисной архитектуры.

Специфичность задач разработки приложений требует проведения работ по адаптации онтологии к условиям использования в составе систем разработки и управления приложениями в распределенной среде «облачных» вычислений. При этом основными направлениями развития являются:

- 1) Реализация работы с «неидеальными» (нечеткими, вероятностными, неоднозначными, субъективными) знаниями. Такая структура знаний учитывается как при взаимодействии с пользователем, так и при построении логического вывода. Алгоритм взаимодействия реализуется таким образом, чтобы по возможности устранить неоднозначность высказываний при описании проблематики задачи, так и ответных сообщений системы. Все знания в рамках системы, полученные от субъективных экспертов, обладают мерой неопределенности, субъективности и взаимной противоречивостью.
- 2) Четкое логическое и функциональное разграничение знаний на базовые, предметно-независимые и определенные предметной областью. При этом должна сохраняться унифицированность представления знаний, обеспечивающая активное использование всей базы знаний в процессе логического вывода. Указанный подход ориентирован на реализацию универсального механизма адаптации базы знаний к произвольной предметной области и соответствующей ей группе задач.
- 3) Междисциплинарный характер онтологической структуры, ориентированный на совмещение точек зрения отдельных групп экспертов, знания которых положены в основу реализуемой базы. При этом, с одной стороны, должны учитываться знания, позволяющие составлять суждения относительно решаемых задач предметной области, а с другой стороны, осуществляться анализ состояния распределенной среды как системы взаимосвязанных программно-аппаратных вычислительных сервисов.
- 4) Нацеленность онтологической структуры на поддержку определения оптимального процесса решения задачи, стоящей перед пользователем (целеориентированный подход).

Таким образом, допустимыми считаются модификации канонических требований к онтологии, направленные на повышение удобства представления знаний с точки зрения их использования в процессе достижения цели. На рисунке 4 приведена структура предлагаемой онтологии, применимой для описания совокупности экспертных знаний, используемых в процессе создания и исполнения композитного приложения. В составе онтологии можно выделить два основных слоя: описание концептов (классов) и описание индивидов, реализующих концепты. При этом индивиды могут быть связаны отношениями, определенными на уровне концептов. Кроме того, допустимы отношения между отдельными концептами (отношение генерализации). В простейшем случае множество отношений может быть ограничено двуместными отношениями. Тем не менее, этот вариант часто приводит к необоснованному усложнению структуры множества отношений за счет декомпозиции семантически связанных многомерных отношений на совокупность двуместных отношений. Еще одним элементом онтологии являются атрибуты (характеристики) индивидов, детализирующие их описание. Одним из возможных расширений является ассоциация характеристик не только с

индивидами (как реализациями классов), но и со связями между ними (как реализации классов допустимых связей). Такое расширение применимо при введении операций 2-го порядка.

## Заключение

Таким образом, на основе комплексной онтологии разработан подход, методы и модели концепции iPSE, обеспечивающей построение проблемно-ориентированных сред распределенных вычислений, а также интеллектуальных технологий создания и управления композитными приложениями. В рамках данной концепции обосновано использование онтологических представлений знаний для описания композитных приложений, формализации структуры приложений на основе языка EasyFlow, построения оптимального расписания исполнения композитного приложения и визуализации больших объемов данных в распределенной среде. Разработана структурная схема и алгоритм функционирования интеллектуальной модели, осуществляющий преобразование композитного приложения с последующим его исполнением на распределенной инфраструктуре «облачных» вычислений. На основе концепции iPSE реализован ряд проблемно-ориентированных программных систем, в частности, высокопроизводительный программный комплекс для квантово-механических расчетов и моделирования наноразмерных систем HPC-NASIS и высокопроизводительный программный комплекс моделирования динамики корабля в экстремальных условиях эксплуатации ShipX-DS [1].

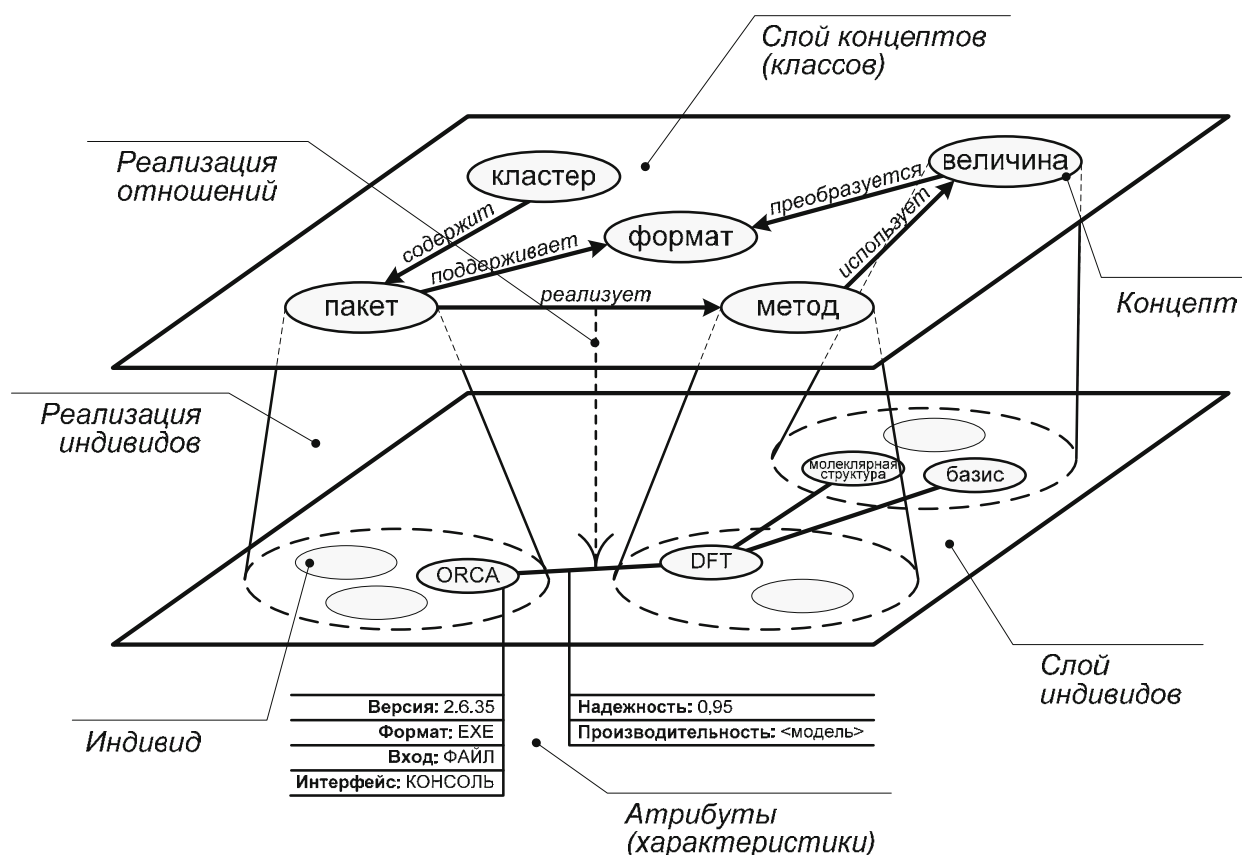


Рисунок 4 – Типовая структура онтологии представления априорных знаний о предметно-ориентированных сервисах, доступных в распределенной среде «облачных» вычислений



Работа выполнена в рамках комплексных НИОКР:

- «Создание распределенной вычислительной среды на базе облачной архитектуры для построения и эксплуатации высокопроизводительных композитных приложений», выполняемого в рамках реализации постановления Правительства РФ №218 (платформа облачных вычислений на базе концепции iPSE);
- «Распределенные экстренные вычисления для поддержки принятия решений в критических ситуациях», выполняемого в рамках реализации постановления Правительства РФ №220 (интеллектуальные технологии ППР с использованием баз знаний).

### Список источников

- [1] Бухановский, А.В. Интеллектуальные высокопроизводительные программные комплексы моделирования сложных систем: концепция, архитектура и примеры реализации / А.В. Бухановский, С.В. Ковальчук, С.В. Марьин // Известия вузов. Приборостроение. - 2009. - Т. 52, №10. - С. 5–24.
- [2] Всероссийская суперкомпьютерная конференция «Научный сервис в сети Интернет: решение больших задач» [Электронный ресурс] <<http://agora.guru.ru/display.php?conf=abrau2008>>.
- [3] Нечаев, Ю.И. Искусственный интеллект: концепции и приложения / Ю.И. Нечаев. - СПб: ГМТУ, 2002.
- [4] Rice, J. R. From Scientific Software Libraries to Problem-Solving Environments / J.R. Rice, R. Boisvert // IEEE Computational Science & Engineering. - 1996. - Vol.3. N 3. - P. 44-53.
- [5] Urgent Computing Workshop 2007. Argonne National Lab, University of Chicago, April 25-26, 2007 [Электронный ресурс]: <<http://spruce.teragrid.org/workshop/urgent07.php>>.

### Сведения об авторах



**Бухановский Александр Валерьевич**, доктор технических наук, профессор, директор НИИ наукоемких компьютерных технологий Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики.

**Boukhanovsky Alexandr Valerievich**, D. Sc., professor, director of scientific research institute of the high computer technology of the St.-Petersburg National Research University Information Technologies, Mechanics and Optics.



**Иванов Сергей Владимирович**, к.т.н., старший научный сотрудник НИИ наукоемких компьютерных технологий Санкт-Петербургского национального университета информационных технологий, механики и оптики.

**Ivanov Sergei Vladimirovich**, Ph.D., main scientific employee of scientific research institute of the high computer technology of the St.-Petersburg National Research University Information Technologies, Mechanics and Optics.



**Ковальчук Сергей Владимирович**, к.т.н., старший научный сотрудник НИИ наукоемких компьютерных технологий Санкт-Петербургского национального университета информационных технологий, механики и оптики.

**Kovalchuk Sergei Vladimirovich**, Ph.D., main scientific employee of scientific research institute of the high computer technology of the St.-Petersburg National Research University Information Technologies, Mechanics and Optics.



**Нечаев Юрий Иванович**, заслуженный деятель науки РФ, академик РАН, доктор технических наук, профессор, ведущий научный сотрудник НИИ наукоемких компьютерных технологий Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики, Международный эксперт в области высокопроизводительных вычислений и интеллектуальных систем.

**Nechaev Yury Ivanovich**, Academician of RANS, Russian Federation Science Honoured Figure, the main scientific employee of scientific research Institute of the high technology computer technologies of the St.-Petersburg National Research University Information Technologies, Mechanics and Optics. International expert in the field of high-performance computing and intelligence systems.

## OXFORD · 2012

*Дорогой Николай Михайлович!*

Выражаю Вам глубокое признание за большой вклад в развитие проблемы интерпретации знаний и возможность публикации статей в престижном журнале «Онтология проектирования» по вопросам теории и практики интеллектуальных систем новых поколений. Методы и модели онтологии позволяют формализовать структуры данных и построить алгоритм преобразования информации при взаимодействии объектов в сложных динамических средах.

Я недавно вернулся из Оксфорда (Англия), где меня награждали орденом за вклад в развитие международных отношений и в связи с присуждением престижной премии имени Альберта Эйнштейна за разработку современной теории катастроф. От имени моих английских и американских друзей и коллег передаю Вам и членам редколлегии журнала «Онтология проектирования» пожелание успехов в научной работе.



*Искренне Ваш Ю.И.Нечаев,  
Выдающийся ученый XXI века:  
Кэмбридж (Великобритания),  
Международный эксперт:  
Вашингтон (США) – Лондон  
(Великобритания),  
Заслуженный деятель науки РФ,  
академик РАН, д.т.н., профессор,  
Лауреат премии имени  
Альберта Эйнштейна(США)*