

УДК 519.711.3

ОНТОЛОГИЯ ЦЕНТРОВ КОМПЕТЕНЦИИ НА ОСНОВЕ СОВРЕМЕННОЙ ТЕОРИИ КАТАСТРОФ В ИНТЕЛЛЕКТУАЛЬНОЙ СРЕДЕ «ОБЛАЧНОЙ» МОДЕЛИ

А.В. Бухановский¹, В.Н. Васильев², Ю.И. Нечаев³

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

¹avb_mail@mail.ru, ²vasilev@mail.ifmo.ru, ³nechaev@mail.ifmo.ru

Аннотация

Рассматривается формализованная модель комплексной онтологии центра компетенции на основе интеллектуальной технологии iPSE (Intelligent Problem Solving Environment) и высокопроизводительных вычислений в среде «облачной» модели. Разработанная модель онтологии учитывает расширение функциональных возможностей технологии iPSE за счет модификации потока событий (WorkFlow – WF). Указаны направления практических приложений онтологической системы знаний при реализации концепции представления и обработки информации в задачах контроля чрезвычайных ситуаций в сложной динамической среде.

Ключевые слова: онтология, центр компетенции, современная теория катастроф, интеллектуальная технология, высокопроизводительные вычисления, «облачная» модель, поток событий, чрезвычайная ситуация, сложная динамическая среда.

Введение

Разработка интеллектуальных приложений в задачах контроля катастрофических явлений представляет собой одно из перспективных направлений интеграции данных в сложной динамической среде [1-8]. Концепция программного инструментария центра компетенции (ПИЦК) разработана на основе технологии iPSE (Intelligent Problem Solving Environment) [1] и определяет принципы построения инструментальной технологической платформы композитных приложений анализа и прогноза текущих ситуаций в интеллектуальной среде «облачной» модели. Платформа ПИЦК организована в виде интеллектуальной оболочки управления параллельными вычислительными процессами в распределенной иерархической среде, включающей в себя вычислительные системы различной архитектуры [2]. Такой подход расширяет возможности традиционных проблемно-ориентированных инструментальных оболочек PSE (Problem Solving Environment) [5], обеспечивая эффективную производительность композитных приложений за счет использования для управления распределенными вычислениями симбиотических знаний об особенностях чрезвычайных ситуаций в рассматриваемой предметной области и специфике вычислительного процесса.

Реализация ПИЦК ориентирована не только на поддержку высокопроизводительных вычислений для суперкомпьютерных систем с традиционной архитектурой, но и для неоднородных систем, включая гиперкластеры (в рамках модели метакомпьютинга) и Грид-системы. Управление эффективностью выполнения сценария вычислений, задаваемого потоком данных WF, позволяет избежать конфликтных ситуаций при разделении ресурсов между вычислительными модулями и пользователями. Таким образом, система ПИЦК может быть отнесена к классу информационных систем (ИС), обладающих сложной распределенной структурой (структурная сложность), многоцелевым характером преобразования информа-

ции (функциональная сложность), а также ориентацией на учет и формализацию неопределенности (информационная сложность).

1 Концептуальная модель онтологии центра компетенции на основе современной теории катастроф

Концептуальные основы создания платформы ПИЦК базируются на достижениях искусственного интеллекта (ИИ) [3]. Важная роль принадлежит принципу открытости, позволяющему обеспечить наиболее сложные уровни иерархической структуры системы – самоорганизацию и самообучение. В результате открываются возможности интерпретации процессов при анализе альтернатив и принятии решений, а также при моделировании действий системы в процессе решения задачи и «обучении» на своем опыте. На рисунке 1 приведена концептуальная схема ПИЦК, определяющая онтологические принципы функционирования системы при контроле чрезвычайных ситуаций на основе методов современной теории катастроф [4].



Рисунок 1 – Концептуальная модель онтологии современной теории катастроф, реализуемая в ПИЦК

Взаимодействие с многопрофильной интеллектуальной технологической платформой ПИЦК осуществляет оператор, реализующий оперативный контроль вычислительного процесса в рамках «облачной» модели. Предметно-ориентированные вычислительные модули представлены соответствующим репозиторием и являются основной частью системы. Композитные приложения и их параллельное исполнение на распределенной вычислительной инфраструктуре реализованы независимо. При этом система гибко учитывает возможности внутреннего распараллеливания, заложенные в каждом из компонентов.

2 Формальная модель онтологии центра компетенции

Формальная модель онтологии ПИЦК определяет последовательность обработки информации в соответствии с функциональным назначением системы. Поток информации, реализующий онтологию чрезвычайных ситуаций как системообразующий элемент, представлен на рисунке 2. Здесь указаны все элементы ПИЦК в рамках концептуальной модели предметной области на основе комплексной онтологии проектирования сложных систем. Модель он-

тологии инвариантного ядра ПИЦК, как программной системы, включает в себя предметную область, базу знаний и базу данных. В функциональном аспекте эта модель содержит следующие компоненты [1, 2]:

$$(1) \quad \langle S(F), S(M), S(W), P_F, T_F, A, X, Y \rangle.$$

Первые три компонента кортежа (1) являются структурами, заданными на множестве элементов $S_i \in S$, которые соответствуют основным компонентам программной системы. Так, $S(F) = \{S_{F1}, \dots, S_{FN}\}$ – совокупность функциональных подсистем, определяющих основные возможности программной системы. $S(M)$ – структурная схема системы, включающая множество $M \equiv \text{Set}(S(M))$ ее компонентов и имеющая собственную организацию $\text{Org}(S(M))$; $S(W)$ – условия формирования целостной системы (цели функционирования, принципы и алгоритмы управления, качество результата решения задачи и эффективность).



Рисунок 2 – Формальная модель онтологии чрезвычайных ситуаций, реализуемая в ПИЦК

Структурная схема системы $S(M)$ в (1) определяется компонентами:

$$(2) \quad S(M) = \langle M, C, T_M \rangle,$$

где C – множество, определяющее совокупность связей между элементами в соответствии с организацией $\text{Org}(S(M))$, а T_M – множество моментов времени в виде последовательности взаимодействия между компонентами. Управляя составом T_M , можно строить на основе указанных компонентов различные динамические архитектуры. В рамках семиотического подхода множество элементов M в структуре S можно описать следующим образом:

$$(3) \quad M = \langle \{m_i\}, I(M), F(M), Q(S) \rangle.$$

Здесь $\{m_i\}$ - множество формальных или логико-лингвистических моделей, реализующих заданные интеллектуальные функции в рамках данного элемента; $I(M)$ – функция выбора необходимой модели (совокупности моделей) в текущей ситуации; $F(M)$ – множество функций модификации моделей m_i ; $Q(S)$ – функция модификации системы S и ее базовых компонентов $I(M), F(M)$. Правила задания (3) определяют характеристики адаптивной составляющей, отвечающей за расширение функциональности, получение новых знаний и приобретение знаний на основе накопленных данных.

Условия формирования системы $S(W)$ представляются совокупностью

$$(4) \quad S(W) = \langle G, R, U_R, K_R, E_G \rangle,$$

где G – цели, обеспечивающие реализацию задачи R , определяющей онтологию системы; U_R – принципы и алгоритмы управления объектом разработки; K_R – качество результата решения задачи; E_G - эффективность, определяющая, какой ценой достигается цель G .

На основе концепции ПИЦК задача R формулируется пользователем на языке предметной области. Целью системы, обеспечивающей решение задачи, является выполнение вычислений при фиксированном наборе параметров (точности расчетов, задаваемых в K_R) за время $T \in T_F$. Принципы U_R заключаются в создании композитного приложения, которое при принятом наборе параметров эффективно выполняется на заданной архитектуре, а в качестве показателя E_G используется классическое понятие параллельной эффективности.

В кортеж (1) также входят параметры, определяющие динамические характеристики ПИЦК: P_F – множество функциональных параметров и T_F – множество моментов времени, инвариантных объектам моделирования, уровню их организации и предметной области. Множество P_F представляет собой наборы функциональных параметров для совокупностей (X, Y, A)

$$(5) \quad P_F^x = \{\pi_F^x\}, P_F^y = \{\pi_F^y\}, P_F^A = \{\pi_F^A\}$$

и подразделяется по области действия на локальные и глобальные параметры. Локальные параметры (P_F^x, P_F^y) являются численными характеристиками отдельного блока (функциональной подсистемы S_{Fi}), сфера действия которых ограничена математической моделью этого блока. Глобальные параметры P_F^A представляют собой характеристики всей системы либо ее нижних уровней иерархии (макроблоков). Параметры любого типового блока задаются в виде выражений, содержащих глобальные параметры.

Оператор $A: X \rightarrow Y$ определяет процесс интерактивного взаимодействия «Оператор – ЭВМ» в процессе функционирования системы $S(F)$; $X = X_j (j = \overline{1, n})$ и $Y = Y_i (i = \overline{1, m})$ – вектор-множества входных и выходных данных. Эти данные включают в себя требования к структуре и функционалу разрабатываемого композитного приложения, представляемые на языке предметной области, а также входные и выходные данные – результаты вычислений.

Формальная модель (1)–(5) позволяет привести описание онтологии ПИЦК на основе iPSE как программной оболочки в соответствие со стандартом IDEF0 информационных систем общего плана. При этом концепция ПИЦК расширяет спецификацию стандарта, поскольку является системой распределенного ИИ. Эта система сочетает строгие формальные методы с эвристическими методами и моделями, базирующимися на знаниях экспертов, моделях рассуждений, имитационных моделях, накопленном опыте эксплуатации.

3 Архитектура и информационные потоки центров компетенции в интеллектуальной среде «облачной» модели

Модель (1)–(5) определяет онтологию, формализующую основные информационные потоки в ПИЦК и может быть отображена на архитектуру программной системы. На рисунке 3 приведена типовая архитектура системы ПИЦК, интегрирующая информационную среду анализа чрезвычайных ситуаций. Система содержит функциональные компоненты, обеспечивающие логический вывод на основе динамической базы знаний, реализуемый под управлением программного модуля. ПИЦК предоставляет пользователю когнитивный диалоговый интерфейс, который включает в себя конструктор сценариев исполнения, посредством которого формируется входная информация для расчетов с максимальным использованием понятийной базы и справочной информации соответствующей предметной области. В состав ПИЦК входит модуль, отслеживающий и анализирующий действия оператора с целью возможной коррекции или оптимизации его действий.



Рисунок 3 – Общая программная архитектура, определяющая онтологию информационной среды ПИЦК

Важную роль при контроле чрезвычайных ситуаций выполняет модуль научной визуализации, а также модуль валидации и верификации результатов расчетов. Его назначение – на основе знаний предметной области формировать соответствующие тестовые задания для оценки работоспособности разрабатываемых пользователем композитных приложений. Таким образом, в рамках ПИЦК на основе комплексной онтологии реализуются разнообразные функции, обеспечивающие функционирование программного комплекса.

Взаимоотношения в системе ПИЦК регламентируются посредством метаописания структуры и состава в форме комплексной онтологии. В рамках онтологии определены формы представления параметров, включая вид и структуру моделей параллельной производительности. Адаптивные функции реализуются посредством модуля получения новых знаний. Он объединяет в себе интеллектуальный редактор знаний для непосредственного взаимодействия с экспертами, компонент приобретения знаний из внешних источников посредством технологий web-mining, а также компонент ретроспективного анализа результатов предыдущих запусков с целью самообучения системы.

Подсистема логического вывода является основным механизмом интеллектуальной оболочки, реализующей вывод на знаниях, ассоциированных с основными программными компонентами управляющего ядра программного комплекса. На основе этих знаний подсистема интерпретирует задание (метаописание, предоставленное пользователем) и путем формирования активных фактов предметной области определяет набор вычислительных сервисов и

сценарий их взаимодействия. После этого ведется построение последовательности конкурирующих оптимальных расписаний параллельного выполнения по данным мониторинга текущего состояния вычислительного комплекса.

4 Формальная процедура онтологии управления производительностью композитного приложения ПИЦК

Принятие решений о режиме исполнения композитного приложения осуществляется на основе априорной информации, математических моделей и структурированной базы знаний с использованием различных методов обработки результатов мониторинга вычислительной среды и учетом неполноты и неопределенности данных, вызываемых немонопольным доступом к вычислительным ресурсам. Онтология, формализующая контроль чрезвычайных ситуаций, представлена на рисунке 4. В рамках этой парадигмы регламентируется возможность *интеллектуального управления* производительностью композитных приложений. Эффективное взаимодействие сервисов организуется подсистемой управления, которая выполняет операцию логического вывода (строит оптимальное расписание) на основе знаний о производительности, заложенных в функциональных сервисах, и данных о функционировании распределенной системы в целом посредством ее мониторинга в режиме реального времени.

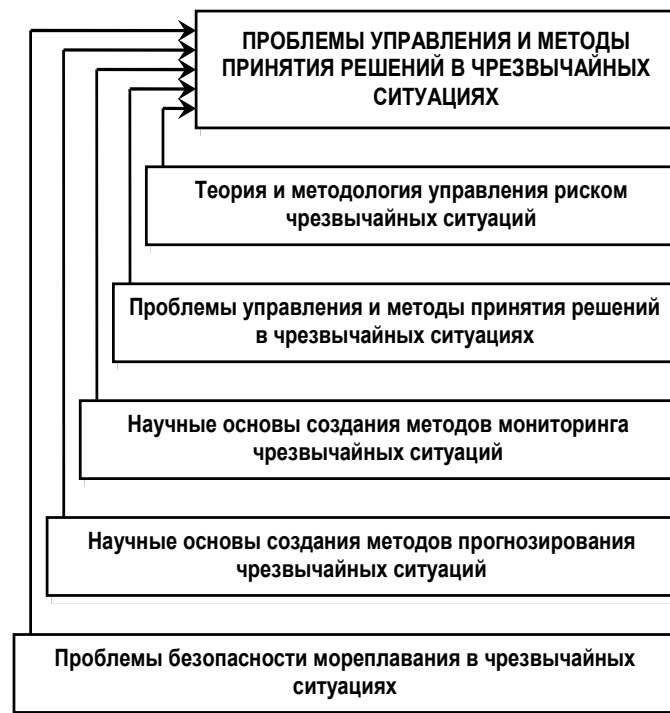


Рисунок 4 – Онтология, определяющая управление и методы принятия решений в чрезвычайных ситуациях

Так как композитное приложение описывается в форме потока задач (WF), формальный механизм построения описания композитного приложения сводится к последовательности преобразований AWF в конкретный (или частично конкретный) CWF. В качестве модели AWF выступает ориентированный ациклический граф $W_a = \{w_a = (V_a, E_a)\}$, где множество вершин V_a – решаемые подзадачи, а множество ребер E_a – зависимости между ними. Промежуточным этапом организации схемы построения приложения является частично конкретный WF, представимый в форме:

$$(6) \quad \begin{aligned} W_i &= \{(w_i = (V_i, E_i), state, resource)\}, \\ state: V_i &\rightarrow \{done, running, scheduled, not_scheduled\}, \\ resource: V_i &\rightarrow C \cup \{\emptyset\}, \end{aligned}$$

где $state$ – функция отображения множества решаемых подзадач на множество состояний планирования: «выполнено», «запущено», «спланировано», «не спланировано»; $resource$ – функция отображения множества решаемых задач на множество доступных ресурсов C .

Для составления расписания используется процедура планирования:

$$(7) \quad sched: W_i \times T_0' \times H \rightarrow W_i,$$

где T_0' – множество, содержащее характеристики времени исполнения основных сервисов в составе WF, H – характеристики распределенной среды. Ход исполнения WF в целом может быть представлен в виде последовательности частично конкретных WF (CWF):

$$(8) \quad \begin{aligned} W_c(w_a \in W_a, sched, t_0', h \in H) &= \{(w_i)\}, \\ w_0 &= \{w_a, state(v) = not_scheduled, resource(v) = \emptyset\}, \\ w_i &= sched(w_{i-1}, t_0', h), i > 0. \end{aligned}$$

При этом функция оценки времени окончания счета на i -м вычислительном ресурсе t_0' (основная характеристика процесса синхронизации) представляет собой отображение вида:

$$(9) \quad t_0' : H \times N \rightarrow R^+.$$

Значения t_0' определяются рамках концепции ПИЦК и интерпретируются как априорные знания предметной области, формой представления которых являются параметрические модели производительности, ассоциированные с доступными вычислительными сервисами предметной области, входящими как базовые структурные блоки в (7), (9).

5 Реализация онтологии центров компетенции в интеллектуальной среде «облачной» модели

На рисунке 5 приведена типовая архитектура ПИЦК на базе технологии iPSE, реализующая особенности разработанной концепции. Программная система функционирует в режиме реального времени в интеллектуальной среде «облачной» модели.

Анализ альтернатив и принятие решений осуществляются на основе данных динамических измерений, априорной информации, математических моделей и структурированной базы знаний с использованием различных методов обработки результатов мониторинга вычислительной среды с учетом неполноты и неопределенности данных. При этом система совмещает различные подходы к представлению знаний – декларативный и процедурный. Декларативная часть системы обеспечивает описание допустимых возможностей, а процедурная – организацию доступа к данным, реализацию вычислительных алгоритмов, интерфейс пользователя.

На основе концепции ПИЦК реализован ряд проблемно-ориентированных программных систем, в частности, высокопроизводительный программный комплекс для контроля чрезвычайных ситуаций, связанных с наводнениями Санкт-Петербурга [8].

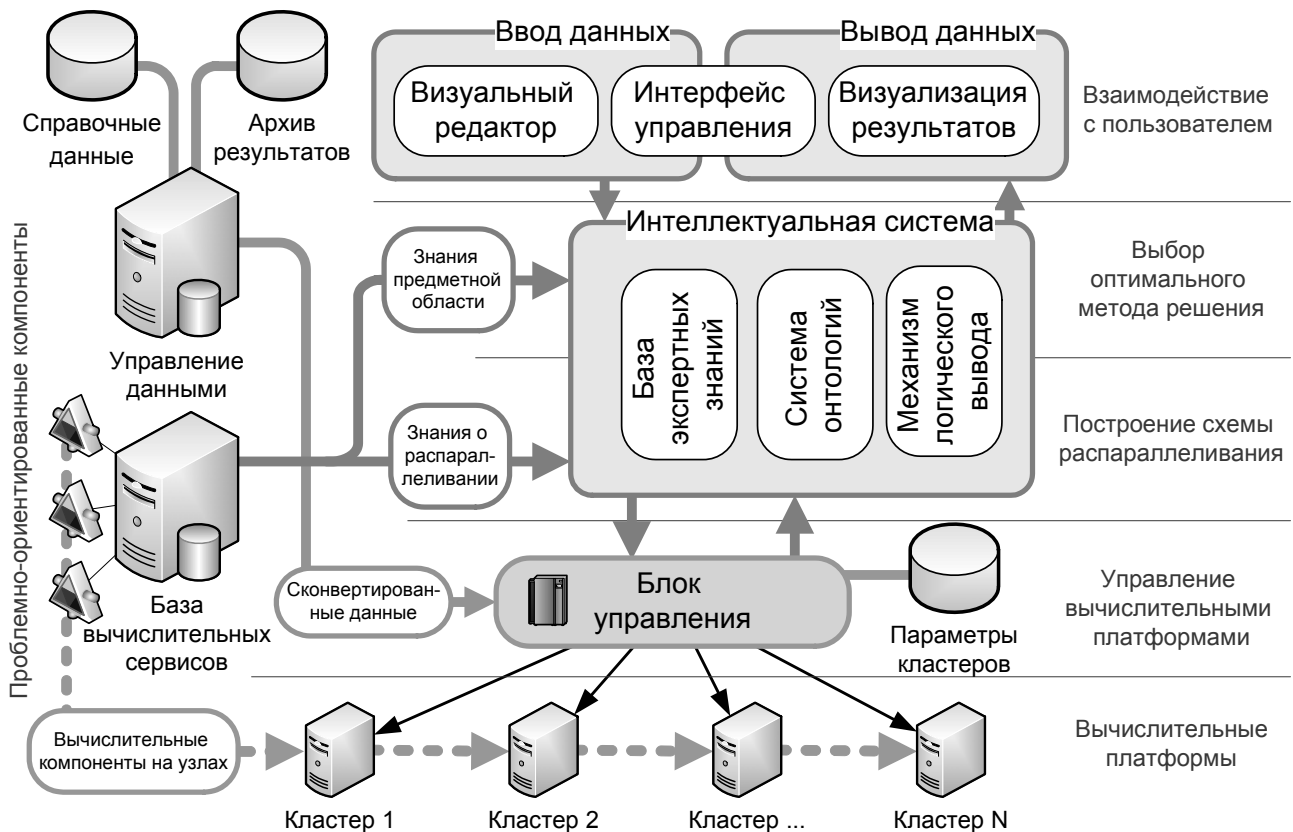


Рисунок 5 – Практическая реализация онтологии центра компетенции в интеллектуальной среде «облачной» модели

Заключение

Таким образом, на основе проведенного исследования разработан подход, методы и модели концепции ПИЦК, обеспечивающей построение проблемно-ориентированной среды распределенных вычислений на основе интеллектуальных технологий создания и управления композитными приложениями. В рамках данной концепции обосновано использование онтологических представлений знаний для описания композитных приложений, формализации структуры приложений, 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 в формах MWF-AWF-CWF с последующим его исполнением на распределенной инфраструктуре «облачных» вычислений.

Работа выполнена в рамках комплексных НИОКР: «Создание распределенной вычислительной среды на базе облачной архитектуры для построения и эксплуатации высокопроизводительных композитных приложений», выполняемой в рамках реализации постановления Правительства РФ №218 (платформа облачных вычислений на базе концепции iPSE); «Распределенные экстренные вычисления для поддержки принятия решений в критических ситуациях», выполняемой в рамках реализации постановления Правительства РФ №220 (интеллектуальные технологии ППР в рамках современной теории катастроф).

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] **Нечаев, Ю.И.** Теория катастроф: современный подход при принятии решений / Ю.И. Нечаев – СПб.: Арт-Экспресс, 2011.
- [5] **Rice, J.R.** From Scientific Software Libraries to Problem-Solving Environments / J.R. Rice, R.F. Boisvert // IEEE Computational Science & Engineering. - 1996. - Vol.3, N.3. - P. 44-53.
- [6] *GTSI Cloud Computing Maturity Model* [электронный ресурс] <http://www.gtsi.com/sms/documents/White-Papers/Cloud-Computing.pdf>
- [7] *Urgent Computing Workshop 2007*. Argonne National Lab, University of Chicago, April 25-26, 2007 [Электронный ресурс]: <<http://spruce.teragrid.org/workshop/urgent07.php>>.
- [8] **Бухановский, А.В.** Высокопроизводительные технологии экстренных вычислений для предотвращения угрозы наводнений / А.В. Бухановский, А.Н. Житников, С.Г. Петросян, П.М.А. Слоот // Известия вузов. Приборостроение. - 2011. - Т. 54, № 10. - С. 14-20.

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.



Васильев Владимир Николаевич, член-корреспондент РАН, доктор технических наук, профессор, ректор Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики.

Vasilev Vladimir Nikolaevich, Korr. Member RAS, D. Sc., professor, rector of the St.-Petersburg National Research University Information Technologies, Mechanics and Optics.



Нечаев Юрий Иванович, заслуженный деятель науки РФ, академик РАЕН, 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.