

Каныгин Геннадий Викторович
Социологический институт РАН
Российская Федерация, Санкт-Петербург
g.kanygin@gmail.com

Полтинникова Мария Сергеевна
Социологический институт РАН
Российская Федерация, Санкт-Петербург
maria.poltinnikova@gmail.com

КОНСТРУИРОВАНИЕ СОЦИАЛЬНЫХ ОБЪЕКТОВ С ПОМОЩЬЮ КОНТЕКСТНО-ОРИЕНТИРОВАННЫХ ОНТОЛОГИЧЕСКИХ МЕТОДОВ

Аннотация. В статье изложены основания и приемы прикладного применения оригинальных контекстно-ориентированных онтологических методов конструирования социальных объектов. На примере Социологического института РАН показано, за счет каких особенностей и каким образом неподготовленные пользователи с помощью контекстно-ориентированных онтологических методов способны коллективно концептуализировать знание о социальном объекте. Проанализирована зависимость знания о социальном объекте от задачи, для решения которой он концептуализируется. Для демонстрации возможностей методов рассмотрена задача отчетности. В иллюстративных целях отчетность понимается как вычисление показателей по социальным объектам, которые предварительно сконструированы участниками социальной коммуникации. На примере объекта Социологический институт РАН продемонстрированы количественные показатели отчетности, автоматически сгенерированные и вычисленные на основе знания о социальном объекте.

Ключевые слова: контекстно-ориентированные онтологические методы; коллективное построение знания; вычисление показателей; социальные объекты; артефакты

Kanygin Gennady Victorovich
Sociological institute RAS
Russian Federation, Saint-Petersburg
g.kanygin@gmail.com

Poltinnikova Maria Sergeevna
Sociological institute RAS
Russian Federation, Saint-Petersburg
maria.poltinnikova@gmail.com

CONSTRUCTING SOCIAL OBJECTS BY USING CONTEXT-ORIENTED ONTOLOGICAL METHODS

Abstract. The article presents intuitive background for and techniques of application of context-oriented ontological methods for constructing social objects. Taken Sociological institute RAS as an example, we demonstrate how and due to what the context-oriented ontological methods are able to assist inexperienced users to build up collectively knowledge of social objects. Particular emphasis is given to the relationship between a constructive representation of social object and the task for which it is constructed. To demonstrate the capabilities of the methods we consider how the managerial reporting can be organized with their help. For illustrative purposes, the managerial reporting is understood as the calculation of indicators on social objects, under the condition that the objects being collectively constructed by social communication participants. With relevance to

Sociological institute RAS object, there are generated and calculated managerial reporting indicators.

Key words: context-oriented ontological methods; collective construction of knowledge; calculation of indicators; social object; artefact

Введение

В современной социологии принято описывать социальные объекты – организации, общественные процессы, сообщества - с помощью естественного языка. Такой способ описания имеет известные недостатки [Алле, 1994].

В статье рассмотрены контекстно-ориентированные онтологические (КОО) методы [Каныгин, Полтинникова, 2015], предназначенные для концептуализации предметных областей социологии. КОО методы относятся к классу информационных онтологий [Рубашкин, 2012]. КОО методы обладают совокупностью характеристик, которые делают их инструментами решения практических задач описания и конструирования социальных процессов.

КОО методы

- сохраняют для гуманитарного исследователя неограниченную возможность применения привычных естественно-языковых обозначений;
- дают в распоряжение социолога компьютерную функциональность современных информационных технологий;
- облачают компьютерную функциональность в наглядную и доступную для социолога форму;
- дают возможность сообществу неподготовленных в компьютерном отношении пользователей осуществлять концептуализацию социальных процессов в стиле работы команды программистов;
- позволяют организовать и контролировать процесс коллективного описания и конструирования социальных процессов самим участникам этих процессов.

Основной задачей статьи является неформальное изложение инструментальных возможностей методов при их практическом применении неподготовленным в компьютерном отношении пользователем.

Объекты и артефакты

Один из отправных пунктов описания и конструирования социальных объектов состоит в разграничении реальных социальных объектов и артефактов [Каныгин, Полтинникова, 2015]. Поясним, каким образом выглядит такое разграничение на примере

операционализации социологических понятий. В работе [Девятко, 1998, с. 81-82] приведен пример операционализации представления о престижности университета. Результатом такой операционализации указан набор характеристик, значения которых выражают, что такое престижность [ibid.].

На наш взгляд, прежде чем отвечать на вопрос, что такое престижность университета, необходимо построить представление о том, что такое университет. Университет – это функционирующий социальный объект, который может быть разным в зависимости от социальных условий, в которых осуществляется его функционирование. Такое функционирование поддерживается сообществом социальных акторов, имеющих свои цели, ресурсы, мотивы и другие особенности.

Действия акторов направлены на решение задач, стоящих перед университетом, т.е. обучения студентов, развития научных исследований, организации приема и многих других. Вряд ли при постановке и практическом решении этих задач акторы университета имеют в виду престижность их учебного заведения.

Вместе с тем, наряду с акторами, поддерживающими функционирование университета, в обществе существуют другие социальные инстанции. Скажем, министерство по работе с университетами. Вероятно, что именно для такого сорта агентов необходимо конструктивное представление о престижности, с помощью которого они смогут сравнивать университеты. Однако следует признать, что в «социальной природе» такое представление базируется на функционировании университета, которое может быть выражено в огромном числе вторичных показателей или артефактов.

На наш взгляд, современный уровень развития средств социологической концептуализации просто не позволяет следовать этому естественнонаучному пути определения престижности. Используя естественный язык для описания происходящего, социолог лишает себя возможности описать университет сколько-нибудь адекватно его функционированию. В этих условиях социолог вынужден конструировать характеристики описываемого социального объекта в виде мнений экспертов [ibid.].

Такой подход позволяет предложить схемы соотнесения естественно-языковых конструкций, скажем, «университет» с конструктивным выражением мнений экспертов, основанных на этих конструкциях. Однако в таком случае возникают принципиальные трудности. Эксперты, не будучи связанными общим знанием об оцениваемом объекте, никак сами не оцениваются на предмет учета сведений, не имеющих отношения к университету. Иными словами, «социологический подход» провоцирует ангажированность оценки.

В стратегической перспективе мы предлагаем коллективно строить представление о социальных объектах, причем делать это силами самих акторов, которые своей деятельностью поддерживают функционирование таких объектов. Тем самым повышается обоснованность представлений о тех явлениях, которые, как это показывает пример с университетом, на сегодняшний день «моделируются» отдельными словами естественного языка.

Построение таких коллективных представлений позволяет конструктивно различить социальные объекты (далее, просто объекты) и характеристики этих объектов, т.е. артефакты. Каким образом такой подход реализуется с помощью контекстно-ориентированной онтологии показано ниже на примере знания о Социологическом институте РАН (СИ РАН).

Говоря о коллективном построении представлений или знания о социальных объектах, мы имеем в виду две особенности такого построения. Во-первых, практически функционирующее знание о социальных объектах (социальное знание) может быть создано не иначе, как сообществом реальных социальных акторов [Каныгин, Полтеникова, 2015]. Для СИ РАН такими акторами являются: коллектив СИ РАН, ФАНО, Правительство и другие. Состав акторов уточняется в процессе создания социального знания.

Во-вторых, любое коллективное построение знания начинается с усилий отдельных акторов. Но с самого начала своих усилий эти акторы вынуждены предусмотреть необходимость включения других участников социальных процессов в виде соответствующих фрагментов знания, создаваемых этими участниками. Тем самым любое индивидуально создаваемое знание о социальных объектах оказывается коллективным.

Метафора контекстно-ориентированной онтологии

Социальный объект в терминах КОО методов – это связанные между собой естественно-языковые обозначения, называемые понятиями или концептами (далее в тексте они написаны заглавными буквами). Путем введения таких обозначений социолог достигает двух целей. Во-первых, соответствующий социальный феномен вне зависимости от его объема или природы получает соответствующее обозначение. Во-вторых, «из содержательных соображений» строит локальное отношение каждого из введенных им самим обозначений с аналогичными обозначениями, введенными ранее.

Принципиально новым по сравнению с существующими в социологии подходами является то, что любое понятие вводится социологом всегда в связи с другим понятием. Первое понятие такой пары называется термином, второе понятие получило название

контекст. Интуитивно это означает, что социальный феномен, выраженный термином, рассматривается при социальных же условиях, которые всегда необходимо указывать.

Связывая обозначения между собой при ответе на вопрос «что такое?», социолог также должен использовать пары понятий.

На рис. 1 приведен пример начального разъяснения того, что такое СИ РАН. Структура, представленная на этом рисунке, называется контекстно-фиксированное разъяснение или ветвление. С помощью именно этой структуры социолог последовательно отвечает на вопрос «что такое?».

Рис. 1 показывает, что понятие СИ РАН рассматривается не отдельно, а в контексте КОЛЛЕКТИВНО ФОРМИРУЕМОЕ ЗНАНИЕ. И далее, ответом на вопрос, что такое СИ РАН//¹КОЛЛЕКТИВНО ФОРМИРУЕМОЕ ЗНАНИЕ буквально является список пар понятий: СТРУКТУРА СИ РАН// НИКАКОЕ², РУКОВОДЯЩИЕ ДОКУМЕНТЫ// НИКАКОЕ, ПРЕДМЕТНАЯ ОБЛАСТЬ//МИНОБРНАУКИ, СОТРУДНИКИ СИ РАН//НИКАКОЕ.

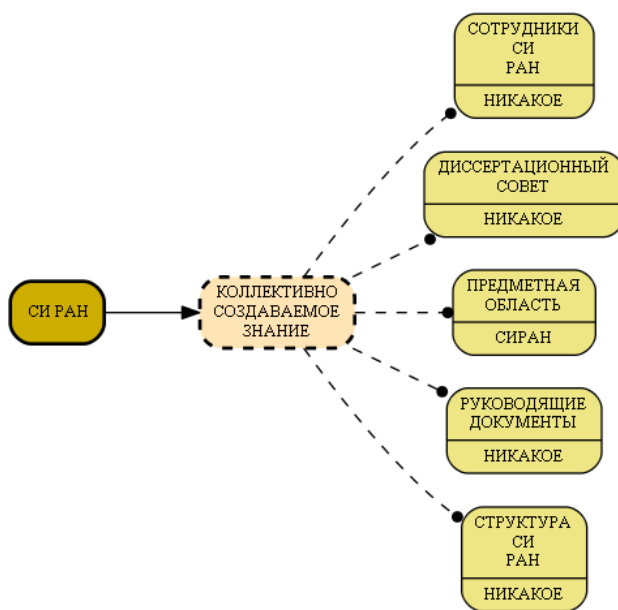


Рис. 1. Ветвление социологических понятий

Откуда взялись эти понятия? Они введены авторами онтологии с целями, во-первых, выражения своего представления о СИ РАН; во-вторых, последующего согласования своего знания с другими социальными акторами, чья деятельность строится на знании о СИ РАН.

Выбор конкретных понятий определяется мнением авторов. Введенные автором концепты могут претерпеть радикальные изменения в ходе дальнейшей концептуализации.

¹ Знаком // мы обозначаем словосочетание «в контексте».

² Контекст НИКАКОЕ используется, когда автор еще не имеет обозначения для контекста, но обязательно введет его позже.

Причиной таких изменений могут быть: учет позиций других соавторов, расширение знания о СИ РАН самих авторов, уточнение естественно-языковых формулировок понятий, изменение круга задач, для решения которых предназначается знание о СИ РАН и многое другое.

КОО методы никак не ограничивают соавторов концептуализации в количестве, составе, естественно-языковых формулировках и других особенностях вводимой терминологии.

Задача социолога при концептуализации, например, СИ РАН состоит в том, чтобы разъяснить в виде ветвления каждое из понятий, составляющих пары, которые он сам вводит по мере необходимости выражения того или иного аспекта функционирования этого социального объекта. Этот процесс начинается с первого шага – ответа на вопрос, что такое СИ РАН//КОЛЛЕКТИВНО ФОРМИРУЕМОЕ ЗНАНИЕ. Второй шаг состоит в разъяснении каждой из пар, возникших на первом шаге - СТРУКТУРА СИ РАН// НИКАКОЕ, РУКОВОДЯЩИЕ ДОКУМЕНТЫ// НИКАКОЕ и т.д. Процесс продолжается до тех пор, пока все аспекты объекта, которые автор онтологии считает важными, не найдут свое обозначение в понятиях.

Так же как естественно-языковой рассказ о СИ РАН, этот процесс разъяснений обусловлен его целями, опытом социолога, стоящими перед ним задачами, привлекаемыми понятиями и другими особенностями. В ходе разъяснений вводимых им понятий социолог создает тезаурус, т.е. словарь естественно-языковых обозначений и множество их контекстно-фиксированных разъяснений. Тезаурус составляется социологом из «содержательных» соображений и в каждый момент своего создания представляет собой множества локальных отношений, структурно подобных КФР и независимых друг от друга.

Тезаурус выражает собой фактологическую базу или наполнение онтологии. Он создается социологом на основе его практического опыта, относящегося к объекту, в данном случае – СИ РАН. Тезаурус разъясняет этот объект в многочисленных «отдельных случаях» его функционирования. Причем, также как конструктивная техника качественного исследования – анализ качественных данных – этот объект описывается только через поименования и связи между ними. Тезаурус не содержит никаких моделей или количественных свойств, имитирующих или измеряющих отдельные процессы, относящиеся к СИ РАН.

Зато в состав КОО методов входят алгоритмы, позволяющие на основе тезауруса, выявлять, насколько социологические концепты, введенные социологом для описания локальных ситуаций функционирования СИ РАН, являются глобально связными.

Ввиду огромного числа таких локальных ситуаций человек не способен «охватить» всю их массу. Контекстно-ориентированные алгоритмы выявляют, возможно ли по разъяснениям отдельных случаев или аспектов знания о каком-либо социальном объекте, скажем, СИ РАН, сконструировать единое связное знание. Такое знание создается в виде терминологического графа (ТГ).

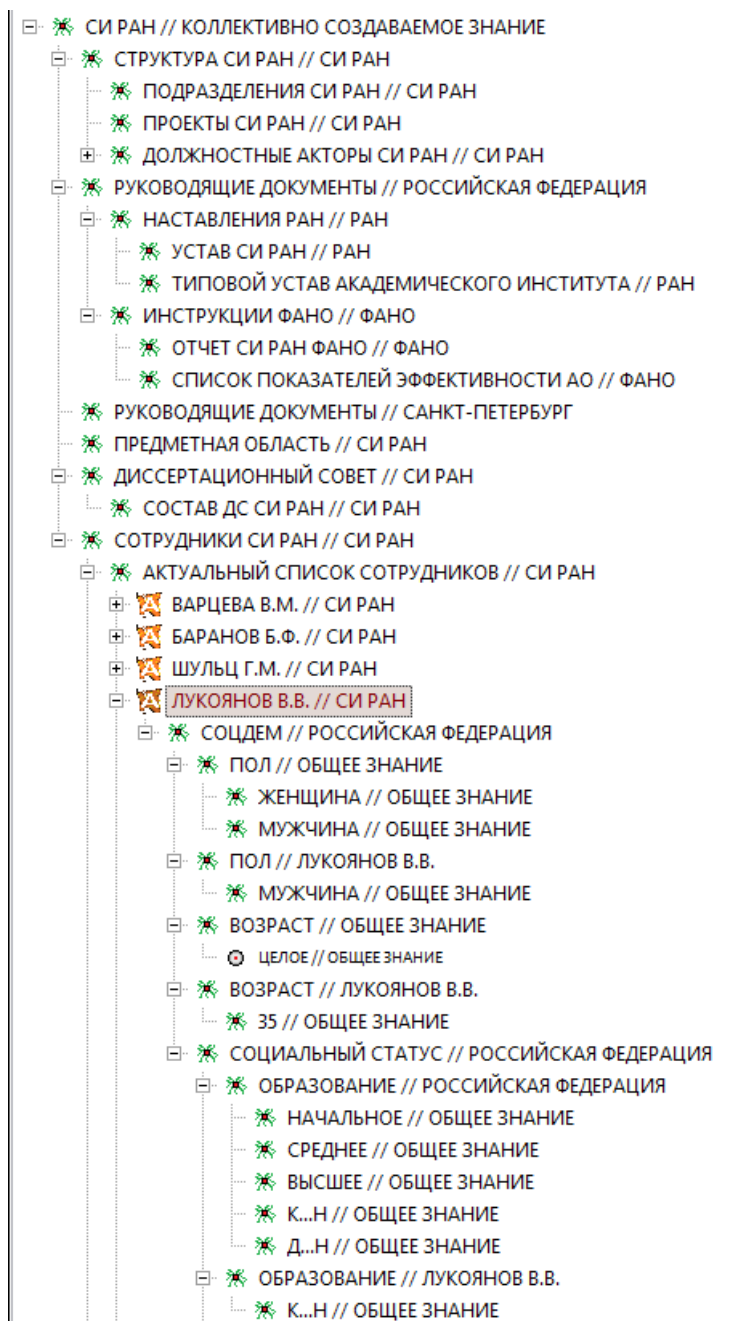


Рис. 2. Фрагмент терминологического графа

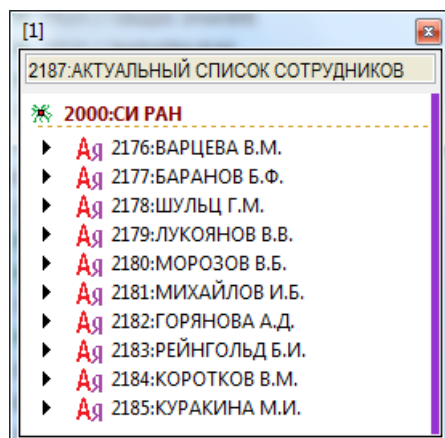


Рис. 3. Актуальный список сотрудников СИ РАН

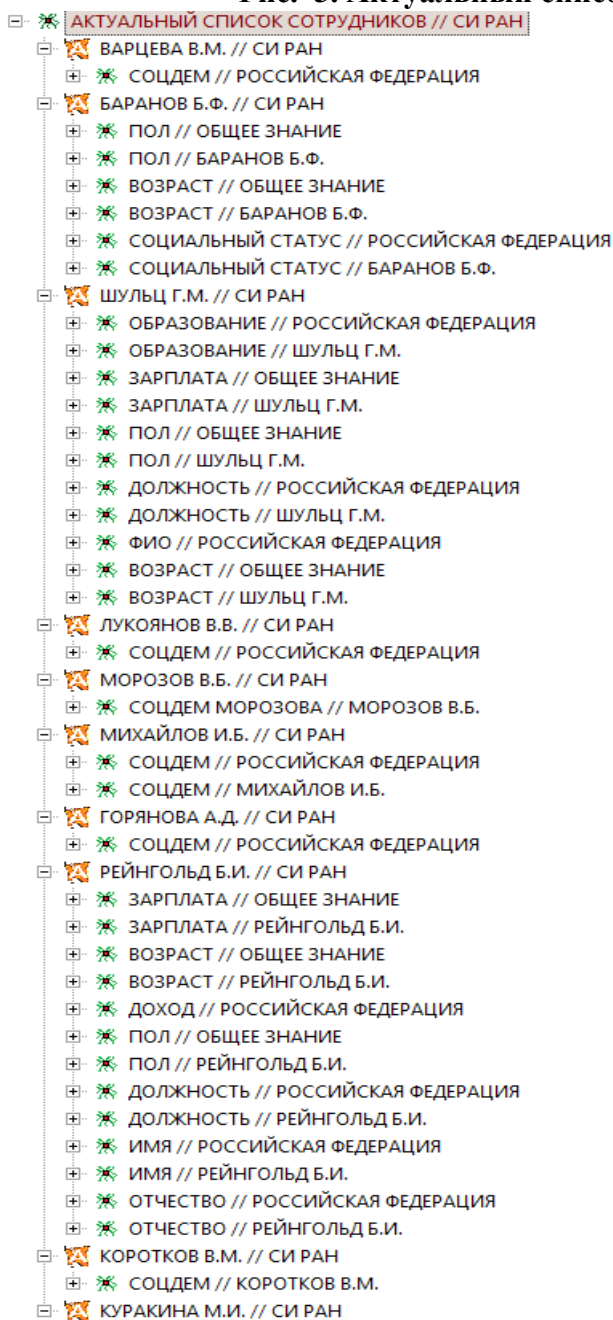


Рис. 4. Описание сотрудника СИ РАН

ТГ для СИ РАН содержит 788 узлов и трудно представим в обозримом виде на печатной странице. На рис. 2-4 приведены различные фрагменты ТГ, выделенные в нем с целью контроля связности отдельных стадий концептуализации.

В частности, рис. 2 показывает, как каждый актер в наглядной и доступной форме может отслеживать в структуре знания отдельные связи, например, свое положение как сотрудника СИ РАН. Тем самым терминологический граф представляет собой средство коллективного контроля за процессом конструирования знания.

Прагматические принципы конструирования социального объекта

Прагматические принципы конструирования социального объекта возникают относительно некоторой задачи по согласованию действий сообщества соавторов. Заранее эта задача неизвестна, она возникает из социальной практики, которую онтологические методы предсказать не могут. Но при наличии такой задачи предлагаемые методы являются инструментом ассистирования соавторам и актерам онтологии в совместном введении, обновлении и понимании терминологии, которая обслуживает решение возникшей задачи. В частности, действий, возникающих как продолжение такой терминологии.

Объект СИ РАН строился с целью автоматизации отчетности этой организации в виде вычисления значений заранее согласованных показателей, которые характеризуют деятельность СИ РАН. Вычисление значений мы будем понимать следующим образом. Знание о СИ РАН поддерживается в актуальном состоянии коллективно, т.е. сообществом актеров, которые регламентируют (например, РАН, ФАНО, Администрация Санкт-Петербурга), осуществляют (например, сотрудники, администрация СИ РАН), пропагандируют (например, авторы статей в журналах SCOPUS) или иными действиями участвуют в функционировании СИ РАН.

Актуальное состояние знания, в частности, предполагает, что его соавторы концептуализируют все практически важные для каждого из них аспекты функционирования СИ РАН. В частности, если кому-то из соавторов необходимы сведения о доходах, возрасте и других характеристиках сотрудников, то эти характеристики должны найти отражение в концептуальном представлении о СИ РАН. В этом случае концептуализация должна быть доведена до числовых значений.

При концептуализации объектов контекстно-ориентированными онтологическими методами задача отчетности, которую решает, например, автоматизированная система управления результатами интеллектуальной деятельности РАН [АСУ РИД РАН], не

возникает как отдельное концептуальное действие, которое отвлекает сотрудников от их прямых обязанностей.

СИ РАН и особенности его определения

В работе [Каныгин, Полтинникова, 2015] мы впервые представили описание социального объекта СИ РАН с помощью контекстно-ориентированной онтологии. В данной работе нас интересует прикладное применение КОО методов. Такое применение мы рассматриваем в виде типичной задачи отчетности СИ РАН перед другими участниками социальной коммуникации.

Для решения этой задачи конкретизируем, какие именно показатели войдут в круг отчетности? Будем считать, что такими показателями будут распределения по полу и должности сотрудников СИ РАН, а также средние значения их возраста и зарплаты.

Однако, чтобы сосчитать эти показатели по объекту СИ РАН, необходимо его концептуализировать адекватно самой задаче. Это означает, что в состав онтологии нужно ввести описания новых объектов, по которым в дальнейшем будут осуществляться вычисления. Судя по названиям показателей, такими объектами являются сотрудники.

Будем считать, что в СИ РАН все сотрудники, по которым нужно вычислять показатели, составляют единое целое, которое мы обозначим понятием **АКТУАЛЬНЫЙ СПИСОК СОТРУДНИКОВ**, отнеся его к контексту СИ РАН. Последнее понятие уже присутствует в тезаурусе. **АКТУАЛЬНЫЙ СПИСОК СОТРУДНИКОВ** вводим заново.

Далее вводим описания самих сотрудников, как показано на рис. 2. Не забываем, что каждое понятие должно быть определено в контексте. Естественно для каждого сотрудника СИ РАН таким контекстом считать СИ РАН. Это означает, что после того, как введено понятие, идентифицирующее сотрудника, например, **ВАРЦЕВА В.М.**, необходимо ввести в тезаурус ветвление вида **ВАРЦЕВА В.М.//СИ РАН**. И так сделать для каждого из сотрудников.

Теперь необходимо концептуализировать представление о каждом сотруднике. Это означает, что каждое из понятий, описывающих сотрудника, должно быть доведено с помощью последовательности разъяснений до значений, которые ассоциированы с «теоретическим понятием», которое выражает показатель.

В рассматриваемом примере мы ввели в состав тезауруса информацию о 10 сотрудниках (см. рис. 3). На рис. 2 показаны результаты такой концептуализации для **АКТУАЛЬНЫЙ СПИСОК СОТРУДНИКОВ//СИ РАН** в виде терминологического графа с раскрытием узлов на глубину в три уровня.

Обратим внимание на ряд особенностей выполненной дополнительной концептуализации. Во-первых, одни и те же понятия употреблены в разных контекстах. Во-вторых, число и состав узлов на нижнем показанном уровне сильно варьируется. Кто-то из сотрудников описан одной парой понятий СОЦДЕМ//РОССИЙСКАЯ ФЕДЕРАЦИЯ, а кто-то (РЕЙНГОЛЬД Б.И.) – тринадцатью. Это говорит о том, что структура описания может конфигурироваться алгоритмами по-разному, в зависимости от тех локальных отношений (кейсов), в которые включен сотрудник.

Теперь раскроем знание о сотруднике, на примере ЛУКОЯНОВ В.В., «до конца» (см. рис. 4). Обратим внимание, что при описании сотрудника использованы как существовавшие в тезаурусе понятия (ПОЛ, ВОЗРАСТ и т.д.), так и новые (ПРИРАБОТОК, 35 и другие). Новые понятия введены в контекстах, которые согласованы с контекстом КОЛЛЕКТИВНО СОЗДАВАЕМОЕ ЗНАНИЕ, в котором определен объект СИ РАН.

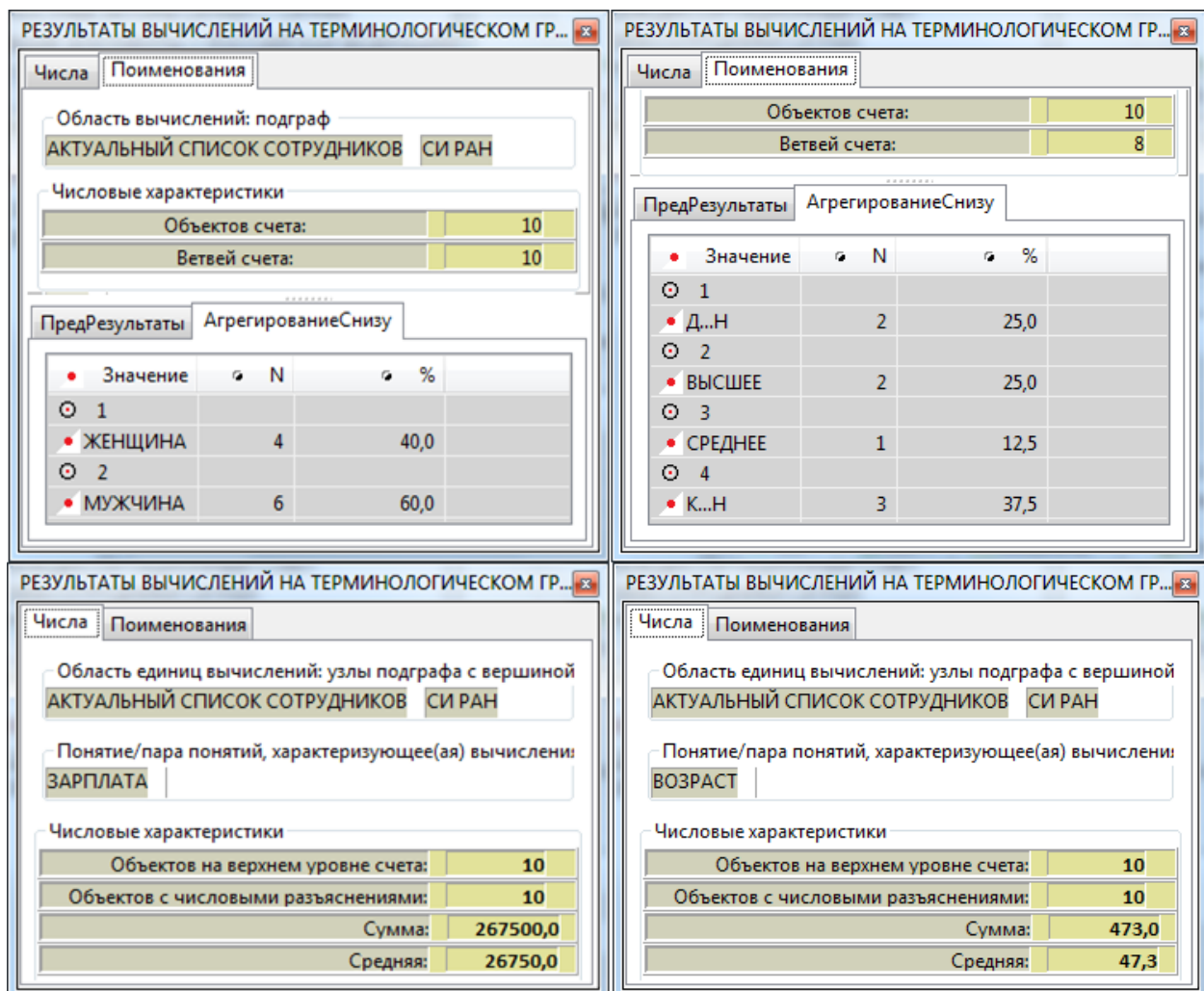


Рис. 5. Значения показателей объекта СИ РАН

Описание сотрудника может иметь избыточный набор ветвлений. Так, например, понятие ПРИРАБОТОК не нужно при решении задачи отчетности. Но это понятие учтено

при генерации ТГ, т.к. разъясняет представление о ДОХОД//РОССИЙСКАЯ ФЕДЕРАЦИЯ. Это представление, в свою очередь, является более общим понятием, через которое введено разъяснение о зарплате сотрудника ЛУКОЯНОВ В.В.

Как избыточное может также восприниматься включенное в терминологический граф ветвление термина ОБРАЗОВАНИЕ в контексте РОССИЙСКАЯ ФЕДЕРАЦИЯ. Для вычислений необходимо указать только ОБРАЗОВАНИЕ//ЛУКОЯНОВ В.В. (что также сделано, см. рис. 4). Однако, во-первых, выявленная многозначность информативна в плане существующих трактовок понятия ОБРАЗОВАНИЕ. Во-вторых, ОБРАЗОВАНИЕ//РОССИЙСКАЯ ФЕДЕРАЦИЯ полезна как шаблон ввода для ОБРАЗОВАНИЕ//ЛУКОЯНОВ В.В., особенно, при работе в условиях большой массы понятий. В-третьих, такая избыточность никак не влияет на решение задачи отчетности.

После того, как объект выражен в виде социального знания, вычисления показателей становится задачей, для решения которой не нужно привлекать экспертов или отвлекать сотрудников СИ РАН. Любой адепт социального знания становится способен получить соответствующий отчет автоматически по своему запросу.

Такие вычисления (назовем их концептуальными) основаны на структуре терминологического графа и ряде соглашений. В этой статье мы приведем только их результаты, собрав на одном рисунке. На рис. 5 представлены количественные характеристики, которые мы считаем решением задачи отчетности средствами контекстно-ориентированных онтологических методов.

Заключение

Отметим, ряд особенностей прикладного применения контекстно-ориентированных методов:

- концептуализация объекта СИ РАН осуществлена в виде социологических понятий с помощью естественного языка;
- социологические понятия проверены на связность и логическую непротиворечивость (см. рис. 2-4);
- графическая форма дала возможность наглядно поэтапно контролировать процесс и результаты концептуализации;
- продемонстрировала возможности модификации социального объекта самими участниками его функционирования.

Библиографический список

Алле М. Современная экономическая наука и факты //THESIS, 1994, Т.2, вып.4, С. 11-19.

АСУ РИД РАН <https://www.asurid.ru/> (дата доступа: 24.04.2015).

Девятко И. Ф. Методы социологического исследования. — Екатеринбург: Изд-во Урал. ун-та, 1998. — 208 с.

Каныгин Г.В., Полтинникова М.С. Контекстно-ориентированная онтология как основа документооборота в социальном управлении // Компьютерная лингвистика и вычислительные онтологии: сборник научных статей. Труды XVIII объединенной конференции «Интернет и современное общество» (IMS-2015), Санкт - Петербург, 23 – 25 июня 2015 г. – СПб: Университет ИТМО. С.94-103.

Рубашкин В. Ш. Онтологическая семантика. Знания. Онтологии. Онтологически ориентированные методы информационного анализа текстов. – М. : Физматлит, 2012. -348 с.