

УДК 004.822

MSC 68T30, 68T35

DOI:10.35330/1991-6639-2020-2-94-31-39

БАЗОВЫЕ ПРИНЦИПЫ И СТРАТЕГИЯ ПРИОБРЕТЕНИЯ И УПРАВЛЕНИЯ ЗНАНИЯМИ НА ОСНОВЕ ЭВОЛЮЦИОННОЙ ПАРАДИГМЫ[□]

Ю.А. КРАВЧЕНКО

Федеральное государственное автономное образовательное учреждение высшего образования
«Южный федеральный университет»
344006, Ростовская обл., г. Ростов-на-Дону, ул. Большая Садовая, 105/42
E-mail: info@sfedu.ru

Статья посвящена описанию теоретических основ приобретения и управления знаниями, отличающихся новой трактовкой концепции, стратегии и принципов решения задач семантического поиска, классификации, структурирования и интеграции знаний в условиях неопределенности и информационного переполнения на основе эволюционной парадигмы.

Согласно основным положениям эволюционной эпистемологии знание растет путем эволюции. Эволюция в классическом понимании осуществляется путем вариации и селекции репликаторов. В живой природе репликаторы – гены, которые подвержены процессам мутации и рекомбинации и селекция которых зависит от успеха репродукции. Репликаторами в эволюции знания становятся идеи, трансформирующиеся через человеческое воображение, фантазии и эмоции. Научное сообщество значительно преуспело в эффективной реализации эволюционных процессов развития современной науки. Развитие научных школ и коммуникаций между ними способствует процессам репликации, вариации и селекции идей. Таким образом, в основе эволюционного развития знания лежат процессы изменчивости, репликации и селекции идей, продуцируемых на множестве распределенных неоднородных источников знаний.

Применение описанных в работе теоретических основ позволяет увеличить объемную долю структурированных знаний в глобальном информационном пространстве и интенсифицировать их совершенствование и применение. При оценке релевантности элементов знания учитывается семантика анализируемого информационного ресурса, что позволяет идентифицировать неявные зависимости и закономерности на множестве системно значимых отношений.

Ключевые слова: приобретение и управление знаниями, семантический поиск, классификация, структурирование, интеграция, механизм эволюции.

ВВЕДЕНИЕ

Как показал анализ литературных источников, задачи приобретения и управления знаниями являются весьма актуальными и в течение многих десятилетий находятся в центре внимания и остаются предметом многих научных исследований [1-7].

Одной из актуальных проблем современной науки является «информационное переполнение». В последнее время данная проблема усугубилась, т.к. значительную часть информации стали создавать не только люди, но и различные технические системы и средства, взаимодействующие друг с другом.

Основой нового знания служит информация. **Знание** – это сложная сетевая иерархия элементов информации с выявленными зависимостями и/или существенными связями между фактами, событиями, явлениями и процессами. По последним прогнозам, уже в

□ Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 18-07-00055 и № 19-07-00099

ближайшее время объем информации в глобальной сети будет удваиваться ежегодно. Сегодня анализируется только 0,5% накопленной в цифровом виде информации [8]. По мнению ученых, одним из основных способов решения проблемы информационного переполнения является переход от хранения и обработки данных к приобретению и управлению знаниями.

Под **приобретением знаний** будем понимать получение информации извне и ее систематизацию, а под **управлением знаниями** – систематические процессы совершенствования, распределения, применения и сохранения знаний, необходимых для успеха определенных видов научной и профессиональной деятельности [8-10].

Особое место в решении данных задач занимают процедуры снижения уровня информационной неопределенности в информационных потоках, сопровождающих процессы приобретения и управления знаниями, а также использование адекватных средств анализа полученной информации для выполнения необходимых процедур оптимизации, прогнозирования и обоснования целесообразности полученных результатов.

1. АКТУАЛЬНАЯ НАУЧНАЯ ПРОБЛЕМА И ГИПОТЕЗА ИССЛЕДОВАНИЯ

Опишем **актуальную научную проблему** информационных технологий, решению которой посвящена данная работа (рис. 1). Обозначим весь объем качественного структурированного знания в глобальном информационном пространстве переменной X , а весь объем необработанной информации – переменной Y , тогда масштабной целью информационных технологий будет являться решение оптимизационной задачи максимизации функции $F = X / Y \rightarrow \max$. При этом необходимо отметить, что в сложившейся ситуации значение данного отношения ежегодно только уменьшается.

Зададим **два базовых уровня инженерии знаний**. Первичный уровень реализуется на основе традиционного понимания процесса появления нового знания, источником которого является человек, а вторичный предусматривает необходимость качественного развития интеллектуальных систем, позволяющих в автоматизированном режиме обрабатывать семантику информации с целью приобретения нового знания.

Гипотезой исследования является следующее утверждение: качественное развитие и интенсификация процессов приобретения и управления знаниями на вторичном уровне инженерии знаний станет решением актуальной научной проблемы увеличения объемной доли структурированных знаний в глобальном информационном пространстве.

Отметим, что эффективность интеллектуальных информационных систем приобретения и управления знаниями зависит от качества решения задач выявления неявных отношений и закономерностей на множестве элементов знания [8-10]. Получение нового знания – это прежде всего результат решения задачи выбора и интеграции множества альтернатив, в том числе на основе **семантического поиска, классификации, структурирования, интеграции:**

– **семантический поиск** – это разновидность информационного поиска, в котором сопоставление смысловых значений элементов знания и создание информационного объекта реализуются на семантическом уровне;

– **классификация знаний** – процесс группировки элементов знаний в соответствии с их системно значимыми признаками;

– **структурирование знаний** – упрощение понимания основных элементов на основе их разделения на группы и подгруппы в соответствии с имеющимися системно значимыми признаками с последующим логическим выстраиванием полученных групп и подгрупп в необходимом порядке;

– **интеграция знаний** – целостный процесс взаимодействия и взаимопроникновения различных систем знаний с применением понятий, теорий и методов одной предметной области в смежных ей.

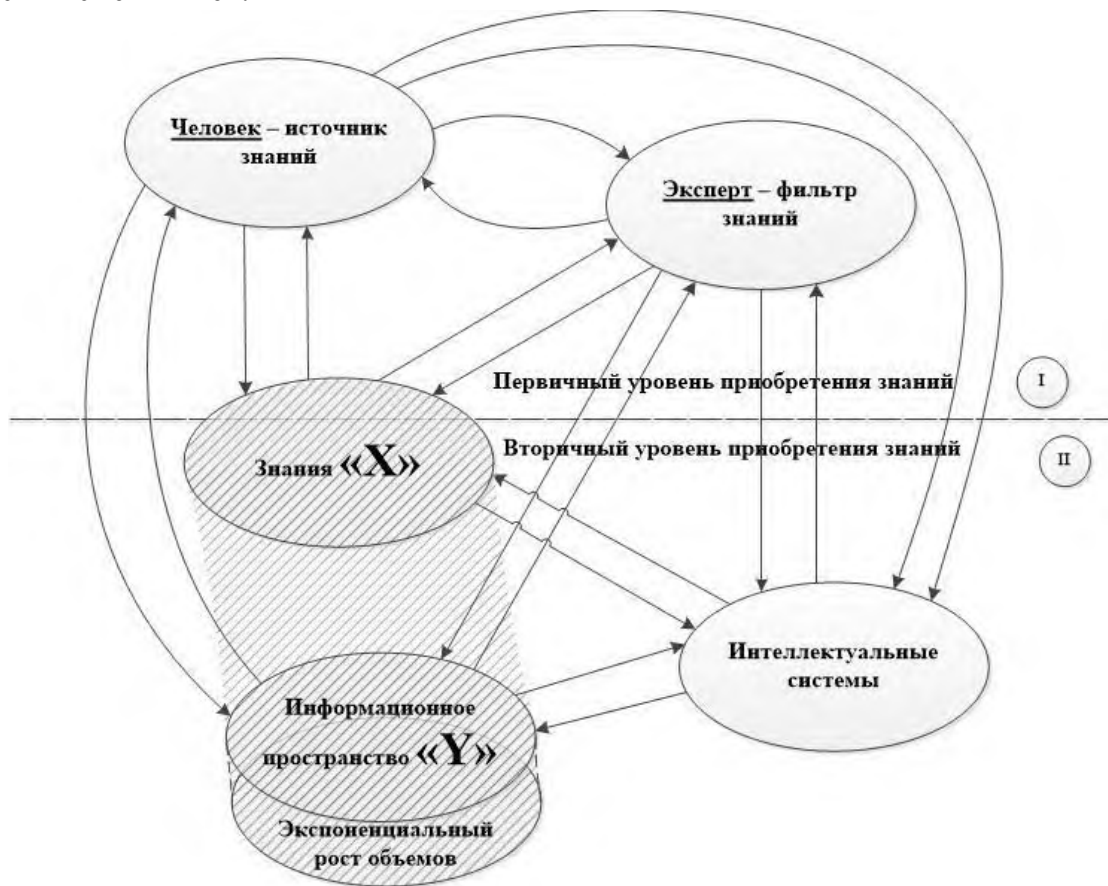


Рис. 1. Описание актуальной научной проблемы информационных технологий

Каждая из перечисленных задач имеет собственные индивидуальные особенности, но в целом все они направлены на идентификацию системно значимых признаков элементов знания для распределения последних по классам, группам и подгруппам, ответов на запросы и онтологиям предметных областей [7-13]. Под **системно значимым признаком**

будем понимать коренное, наиболее важное свойство объекта или события (процесса), без которого теряется смысл.

2. ЭВОЛЮЦИЯ ЗНАНИЙ

Эволюция в классическом понимании осуществляется путем вариации и селекции репликаторов [13]. Репликаторами в эволюции знания будем считать идеи. Идея – это терм, окруженный релевантным знанием. В основе эволюционного развития знаний лежат процессы изменчивости, репликации и селекции идей, продуцируемых на множестве распределенных неоднородных источников знаний.

Каждый из перечисленных процессов имеет собственные особенности и характеристики:

- для активизации процесса **изменчивости** необходимо построение весомых и выразительных связей;
- для развития **репликации** необходимо создать эффективные методы и алгоритмы распределения и индексации анализируемой информации;
- для повышения эффективности процесса **селекции** необходимо разработать средства фильтрации знаний (knowledge sifters).

Эволюцию знаний будем рассматривать с точки зрения модели, предложенной Расселом Акоффом [11], в которой он выделил три основных информационных класса, широко используемых техническими системами, – «**данные – информация – знания**», но, кроме этого, он выделил класс «**понимание**» – определение закономерностей связей на макроуровне между элементами знания в контексте решаемой задачи, реализация которого невозможна без формализации семантики поиска знаний и является целью развития интеллектуальных информационных систем. Также в данной классификации присутствует категория «**мудрость**» – взвешенное осознание закономерностей между элементами знания на метауровне семантической иерархии (рис. 2).



Рис. 2. Иерархия информационных категорий, предложенная Расселом Акоффом

Механизм эволюции с точки зрения управления знаниями – последовательность информационных процессов извлечения, приобретения, обработки, отбора (селекции), совершенствования и передачи знаний о свойствах и признаках отдельных элементов и системы в целом. Описание данного механизма эволюции имеет сходство с механизмом эволюции **Шмальгаузена**, который рассматривал его с точки зрения кибернетики и определял как последовательность нетривиальных информационных процессов: приобретения, выбора, передачи, совершенствования информации о характеристиках (признаках) обособленных элементов (объектов) и всей системы в целом.

В предложенном новом механизме расширена область применения процессов управления знаниями и произведен переход на междисциплинарный уровень сложной сетевой иерархии элементов информации с выявленными зависимостями и существенными связями между фактами, событиями, явлениями и процессами. Тогда как механизм эволюции Шмальгаузена предполагает использование выделенной, упорядоченной и обработанной в соответствии с контекстом части данных, наделенной системой отношений и определенным смыслом (данные о данных, или данные плюс метаданные) [12].

Моделирование рассуждений в задачах приобретения и управления знаниями проводится на основе сравнения **семантических векторов**, построенных на основе векторных репрезентаций слов, в которых смысл имеют только расстояния между векторами. Негативным эффектом указанного свойства векторной репрезентации является быстрая деградация векторов при операциях над ними [3-8].

При построении семантического вектора описывается достаточное количество устойчивых кластеров, отражающих значимые признаки с учетом контекста. В первую очередь вычисляется семантический вектор слова, описывающий расстояние от слова до центра соответствующего кластера в векторном пространстве. После чего, складывая семантические вектора отдельных слов, составляющих ресурс, получаем семантический вектор всего ресурса. Скалярное произведение семантических векторов различных информационных источников позволяет оценить релевантность исследуемых знаний, которая тем выше, чем больше результат произведения.

3. КОНЦЕПЦИЯ И БАЗОВЫЕ ПРИНЦИПЫ

Сформулируем концепцию как систему взглядов, представленную в виде ряда утверждений и интегрирующую их основные принципы, являющиеся элементами теории приобретения и управления знаниями на основе эволюционной парадигмы:

– **Утверждение 1.** *Сравнительный анализ биологической эволюции и эволюции информационных процессов и систем позволяет выявить между ними определенное сходство.*

Пояснение 1. Процессы конкуренции на рынке информационных технологий являются моделью «борьбы за существование». Видовое изменение («мутация») объектов в сфере IT зачастую приобретает ранее непредсказуемые формы. Исчезновение с рынка невостребованных технологий моделирует «естественный отбор». Всем информационным системам свойственна высокая степень «типизации». Более того, информационные системы имеют совершенную «организацию», реализованную в сетевых технологиях. Повторное применение успешных решений моделирует «наследование», а стремление сохранить целостность системы – «гомеостаз». Подобные аналогии характерны также для понятий «обмен веществом и энергией»; «раздражимость»; «онтогенез» и «филогенез» [2, 6, 8, 14].

– **Утверждение 2.** *По отношению к знаниям откажемся от метафоры «здания» и примем метафору «сети», где будем понимать только сети, вложенные в сети.*

Пояснение 2. Декомпозиция элементов знаний позволяет усилить инструктирующий эффект метазнания и идентифицировать неявные отношения на множестве содержательных информационных элементов. Информационные системы имеют схожую с живыми организмами структуру, в которой элементы системы, сами системы и сети систем во многом повторяют принципы строения частей организмов, самих организмов и их сообществ [1, 4, 7]. Причем каждый из элементов данной иерархии также является

интегрированным целым. Онтология представляет собой наиболее формализованную модель представления знаний.

– **Утверждение 3.** *Реальная среда для решения задач приобретения и управления знаниями представляет собой сложный для формализации комплекс взаимосвязанных предметных областей, в котором возникают и актуализируются неструктурированные задачи выявления и идентификации элементов знания.* **Пояснение 3.** Информационное пространство знаний характеризуется свойствами частичной наблюдаемости, динамики, непрерывности, эпизодичности, стохастичности [3-6]. Построение модели подобной среды, пригодной для интерпретации и эффективной обработки, основано на многоагентном подходе.

– **Утверждение 4.** *Повышение эффективности процедур приобретения и управления знаниями тесно связано с двумя основными современными тенденциями развития интеллектуальных систем, свойственными живой природе: наличие у информационных систем интеллектуальных свойств и способность информационных систем к самоорганизации и саморазвитию.* **Пояснение 4.** Самоорганизацию информационных систем будем рассматривать с точки зрения их саморегулирования, которое проявляется в сбалансированном функционировании системы за счет отношений между внутренними и внешними элементами. Согласно классическому определению процесс самоорганизации приводит к значительному росту сложности внутренней организации системы [12-14]. Самоорганизация не является очевидным свойством, проявляющимся при любых обстоятельствах, даже для интеллектуальных информационных систем. Для неизбежности возникновения процессов самоорганизации необходимо выполнение определенных внешних и внутренних условий, связанных не с увеличением количества информации, а с проявлением ее инструктирующих свойств, повышающих ценность данной информации, непосредственно связанную с ее используемостью.

– **Утверждение 5.** *Знания, помимо самостоятельной ценности, создают синергетический эффект по отношению к информационным процессам, используя их, воздействуя на уровень эффективности за счет дополнительного инструктирующего качества, превышающего сумму слагаемых элементов знания.* **Пояснение 5.** Таким образом, одним из условий возникновения самоорганизации в информационной системе является наличие механизмов реализации поиска и отбора информации, имеющей определенную меру ценности, что является упрощенной моделью творческой деятельности человека [1, 8, 13, 14]. Нетривиальность данного подхода к возникновению самоорганизации заключается в том, что ценная информация накапливается в системе медленно и постепенно, давая на протяжении долгого времени лишь несущественные преимущества. Значительные изменения в поведении системы возможны только при условии постоянного анализа и обработки имеющейся ценной информации с целью определения ранее неизвестных зависимостей и закономерностей, выделяющих среди множества элементов информации упорядоченные последовательности составных частей знания нового качества, использование которого позволит системе перейти на новый уровень собственной эволюции. Управляемое создание условий для протекания такого процесса самоорганизации усложняется неопределенностью надлежащего момента времени, в котором интеграция определенных разрозненных элементов ценной информации приведет к формированию знания нового качества. Дальнейшее протекание процесса накопления ценной информации может привести к рассеиванию преимуществ интеграции.

На основе описанной концепции сформулируем базовые принципы приобретения и управления знаниями.

Принцип эволюционного развития знаний – эволюция знаний – совокупность процессов изменчивости, репликации и селекции идей, продуцируемых на множестве распределенных неоднородных источников информации и знаний.

Онтологический принцип знаний – визуальное увеличение каждого узла онтологической структуры знаний позволит увидеть другую сеть, т.е. в целом взаимозависимость всех явлений будет представлена сетями внутри сетей.

Принцип агентности модели информационного пространства – построение модели информационного пространства основано на концепции агентности, предполагающей множественность и автономность одновременно со связанностью.

Принцип создания интеллектуальных систем приобретения и управления знаниями – эволюционная логика отрицает возможность автономной разработки интеллектуальной системы приобретения и управления знаниями, такую систему можно только «вырастить» вместе с информационным процессом, который нуждается в ней.

Принцип системности эволюции знаний – бессистемный сбор знаний при отсутствии исследования отношений между концептами (элементами знания) и индексирования по субъектам поиска тормозит эволюцию знания и препятствует способности продуцирования и выразительности новых идей.

4. СТРАТЕГИЯ ИССЛЕДОВАНИЯ

При разработке любого информационного процесса необходимо реализовать выбор типа субъектов выполнения цепочек операций, построение модели информационной среды, создание инструментов решения основных задач исследования, а также методов использования данных инструментов. В процессах приобретения и управления знаниями данные этапы включают в себя: разработку коалиций интеллектуальных агентов и описание их основных свойств; построение онтологической модели междисциплинарного пространства знаний и информационных моделей решения задач семантического поиска, классификации, структурирования, интеграции знаний; создание биоинспирированных методов приобретения и управления знаниями. Также необходимо учесть возможные отрицательные воздействия внешней среды (большая размерность, информационная неопределенность, избыточность, недостоверность и т.п.) и варианты реакции интеллектуальной информационной системы на эти воздействия.

Представим **стратегию** приобретения и управления знаниями как принципиальную организацию описанных этапов (рис. 3). Данная стратегия направлена на получение достоверного знания. Для Аристотеля **достоверное знание** – результат логического рассуждения, направленного на открытие начал, причин и элементов [8], что в современном понимании трактуется как система значимых отношений на множествах информации мета- и содержательного уровней.

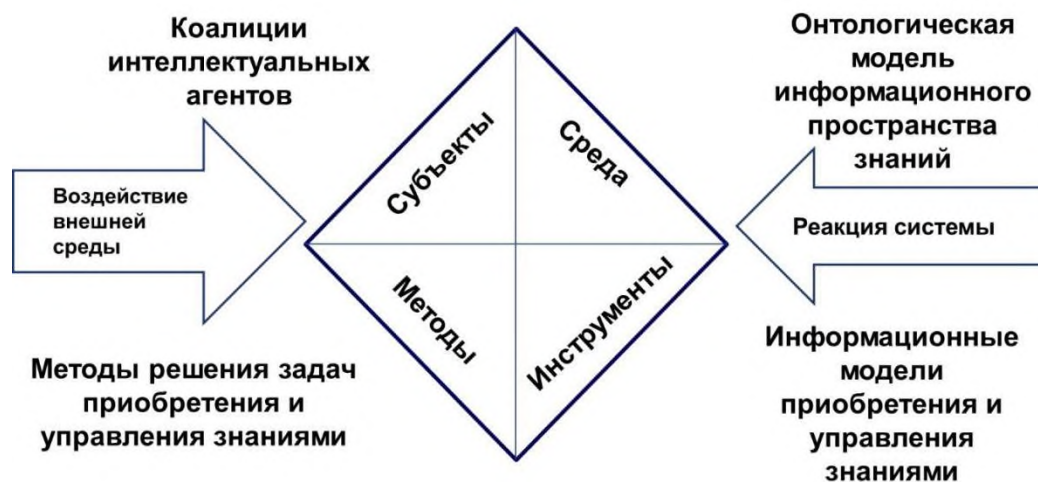


Рис. 3. Стратегия исследования

Под рассуждением будем понимать построение последовательности аргументов, вынуждающих принятие некоторого утверждения, которое и является целью рассуждения [5-8]. Полученный вывод представляет собой систему семантически проанализированных отношений и соответствий, приводящих к обнаружению неявных (скрытых) зависимостей и закономерностей.

Все представленные в стратегии исследования этапы являются комплексными и масштабными. Создание коалиций интеллектуальных агентов направлено на воспроизведение системной организации эксперта на основе интеграции искусственных и естественных (известных из психологии) стилей мышления и способов учения.

Реальное информационное пространство знаний представляет собой сложный для формализации комплекс взаимосвязанных предметных областей, в котором возникают и актуализируются неструктурированные задачи идентификации и декомпозиции элементов знания и их отношений на уровне предметных областей и на междисциплинарном уровне.

Разработка информационных моделей и биоинспирированных методов приобретения и управления знаниями направлена на создание механизмов оценки семантической близости концептов интегрируемых онтологий

ЗАКЛЮЧЕНИЕ

Автор считает, что в данной работе новыми являются следующие положения и результаты: теоретические основы приобретения и управления знаниями, отличающиеся новой трактовкой концепции, стратегии и принципов решения исследуемой в работе актуальной научной проблемы на основе описания эволюции информационных технологий в целом и знаний в частности. Применение данных теоретических основ позволяет увеличить объемную долю структурированных знаний в глобальном информационном пространстве и интенсифицировать их совершенствование и применение.

ЛИТЕРАТУРА

1. Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии) / Под общ. ред. В.З. Ямпольского. Томск: Изд-во НТЛ, 2005. 260 с.

2. *Рассел С., Норвиг П.* Искусственный интеллект: современный подход, 2-е изд.: пер. с англ. М.: Издательский дом «Вильямс», 2006. 1408 с.
3. Kravchenko Yu.A., Kureichik V.V. Knowledge management based on multi-agent simulation in informational systems // (2014) 8th IEEE International Conference on Application of Information and Communication Technologies, AICT 2014. Pp. 264-267.
4. *Кравченко Ю.А.* Синтез разнородных знаний на основе онтологий // Известия ЮФУ. Технические науки. Таганрог: Изд-во ГТИ ЮФУ, 2012. № 11 (136). С. 216-221.
5. *Кравченко Ю.А.* Многоуровневая архитектура сценария управления знаниями на основе онтологического анализа // Известия ЮФУ. Технические науки. Таганрог: Изд-во ЮФУ. 2015. № 2 (163). С. 186-195.
6. *Kureichik V.V., Kravchenko Yu.A., Vova V.V.* Decision support systems for knowledge management // (2015) Advances in Intelligent Systems and Computing, 349. Pp. 123-130.
7. *Осуга С.* Обработка знаний. М.: Мир, 1989. 292 с.
8. *Бородакий Ю.В., Лободинский Ю.Г.* Эволюция информационных систем (современное состояние и перспективы). М.: Горячая линия. Телеком, 2011. 369 с.
9. *Норенков И.П.* Онтологические методы синтеза электронных учебных пособий // Научнопрактический журнал «Открытое образование». М.: CAPITALPRESS, 2010. № 6. С. 39-44.
10. *Кравченко Ю.А.* Задачи семантического поиска, классификации, структуризации и интеграции информации в контексте проблем управления знаниями // Известия ЮФУ. Технические науки. Таганрог: Изд-во ЮФУ, 2016, № 7 (180). С. 5-18.
11. *Акофф Р.* Искусство решения проблем. М.: Книга по требованию, 2012. 221 с.
12. *Колесников А.А.* Основы теории синергетического управления. М.: Фирма «Испос-Сервис», 2000. 256 с.
13. *Курейчик В.В.* Эволюционные, синергетические и гомеостатические методы принятия решений: монография. Таганрог: Изд-во ТРТУ, 2001. 99 с.
14. *Popper, Karl R.* Objective Knowledge: An Evolutionary Approach. Oxford University Press, London (1972). P. 201.

Работа поступила 09.04.2020 г.