

Теоретические основы моделирования процессов формирования знаний

И.М. Зацман (Россия)

Тезисы доклада

Исследования процессов формирования новых знаний традиционно относились к экономике, в рамках которой использовались и до сих пор используются качественные модели разных технологий «мозгового штурма». Необходимость в количественных моделях стала тем фактором, который сделал доминирующей информационно-компьютерную компоненту этих исследований. Роль и значение этой компоненты рассматривались приглашенными экспертами при подготовке 7-й Рамочной программы ЕС в рамках семинара «Knowledge Anywhere Anytime: “The Social Life of Knowledge”». Материалы этого семинара использовались при формировании программ работ по информационно-коммуникационным технологиям (ИКТ) 7-й Рамочной программы ЕС.

В этих материалах отмечается, что исследование процессов, разработка методов и моделей формирования новых и развития существующих методов и моделей структуризации и представления знания является актуальной проблематикой в XXI в. Участники семинара определили четыре направления этих исследований.

1. Формирование научного понимания того, как знание появляется, каким образом на этот процесс и его результаты влияет совместная деятельность, как формируется конвенциональное знание. Одна из задач этого направления заключается в том, чтобы создать методы и средства описания различий в личностном понимании участниками совместной деятельности смысла одного и того же текста, графика, диаграммы, изображения и т.д.

2. Исследование многообразия форм представления одних и тех же концептов как «квантов» ментального знания. Кроме форм представления конвенциональных и стабильных концептов, предметом исследования являются формы представления личностных и коллективных (согласованных) концептов. В рамках этого направления предполагается выполнение исследований процессов формирования конвенциональных концептов на основе личностных и коллективных концептов.

3. Создание нового поколения интеллектуальных информационных систем, которые должны обеспечить семантическую интероперабельность в процессе совместной работы пользователей этих систем.

Степень новизны интеллектуальных систем, поддерживающих процессы выявления и экспликации стадий генерации и эволюции нового знания, предлагается оценивать, сравнивая

их с традиционными системами управления знаниями (Knowledge Management Systems - KMS), основанными на гипотезе стабильности ментального знания человека. Согласно этой гипотезе, в процессе создания и применения KMS можно не учитывать эволюцию во времени ментального знания, представленного в KMS.

4. Исследование принципиальных возможностей и средств влияния на процессы генерации новых или эволюции существующих систем знаний в процессе совместной деятельности коллективов специалистов. Наиболее актуальные вопросы этого направления исследований связаны с пространственно-распределенными коллективами специалистов, совместная деятельность которых обеспечивается сетевыми технологиями. В случае решения задач целенаправленной генерации ментального знания предполагается, что имеет место ситуация его неполноты, эта неполнота специфицирована, и перед коллективом специалистов поставлена задача уменьшения степени лакуарности существующих систем знаний.

Приведенный перечень из четырех направлений исследований говорит о том, что участники этого семинара придают ИКТ ключевую роль в решении проблем формирования новых и развития существующих систем знаний. При этом они существенно расширили границы предметной области представления знаний. Это расширение произошло, в основном, за счет вопросов генерации новых знаний и влияния на процессы генерации и эволюции систем знаний средствами ИКТ. Анализ материалов этого семинара позволяет сделать вывод о том, что предметная область, очерченная приглашенными экспертами, включает, в дополнение к традиционным задачам представления знаний, следующие три актуальных направления исследований:

- 1) генерация и компьютерное представление в цифровой электронной среде личностных и согласованных концептов как «квантов» ментального знания, формируемых пространственно-распределенными коллективами специалистов;
- 2) анализ и оценивание степени релевантности систем формируемых концептов социальным, экономическим, технологическим и другим общественно значимым потребностям, в интересах удовлетворения которых они формируются;
- 3) целенаправленное влияние с помощью ИКТ и когнитивных технологий на генерацию и эволюцию систем формируемых концептов, необходимых для удовлетворения общественно значимых потребностей.

В рамках первого направления исследований автором получены решения двух проблем:

- концептуальное индексирование геоизображений как форм представления личностного ментального знания специалиста о геообъектах;

- генерация систем новых экспертных знаний, формируемых пространственно-распределенными коллективами специалистов в интересах обеспечения мониторинга программно-целевой деятельности.

В процессе решения этих проблем были разработаны количественные модели процессов генерации и методы компьютерного кодирования личностных и согласованных концептов как структурных элементов систем экспертных знаний. Теоретические основания для создания этих моделей были построены в процессе развития подходов Горна и Шрейдера к интеграции информационной и компьютерной наук.

Рассматриваемые в докладе теоретические основания интеграции включают: деление предметной области информационно-компьютерной науки на среды; использование двух классов понятий (однородных по своей природе и двуединых), распределение однородных понятий по средам и двуединых понятий по границам между средами; использование двух взаимосвязанных стадий представления концептов в цифровой электронной среде в виде компьютерных кодов.

Приведем краткое описание перечисленных оснований интеграции информационной и компьютерной наук, используя систему терминов из работы [1]. В эту систему терминов включены два основных класса понятий: однородные по своей природе (знания, знаковая информация, цифровая информация, цифровые данные и компьютерные коды) и двуединые (семиотические знаки, формокоды и семокоды). В начале процесса построения теоретических оснований рассматривались следующие четыре среды (затем число сред было увеличено) и соответствующие им однородные понятия:

- *ментальная среда знаний человека*, которая включает выражаемые и невыражаемые (имплицитные) знания;
- *социально-коммуникационная среда*, к которой относятся отчужденные от человека сенсорно-воспринимаемые формы представления его знаний (знаковая информация);
- *материальная среда объектов и явлений*, на основе интерпретации которых человеком генерируются новые концепты знания, относящиеся к *ментальной среде*;
- *цифровая электронная среда*, к которой относятся цифровая информация, цифровые данные, компьютерные коды концептов и форм представления выраженных знаний.

1. Зацман И.М. Концептуальный поиск и качество информации. – М.: Наука, 2003. – 271 с.

Таким образом, каждое однородное по определению понятие было соотнесено только с одной из перечисленных четырех сред. После распределения однородных понятий по средам были описаны две последовательные стадии представления знаний в цифровой электронной среде в виде компьютерных кодов. Первая стадия представления знаний относится к границе между ментальной и социально-коммуникационной средами. Семиотические знаки как двуединые по своей природе понятия принадлежат именно этой границе, на которой осуществляется ассоциативное соотнесение концептов и форм их представления с использованием семиотических знаков.

Вторая стадия представления концептов как «квантов» ментального знания относится к границе между социально-коммуникационной и цифровой электронной средами. На этой границе используются таблицы компьютерного кодирования символов, необходимые для представления текстов в цифровой электронной среде. Эти таблицы принадлежат границе между этими средами. Для вербализуемых знаний последовательное использование сначала языковых знаковых систем в процессе генерации текстов, а затем таблиц компьютерного кодирования является традиционным способом представления концептов в цифровой электронной среде.

Важно отметить, что получились не две, а три границы между:

- ментальной средой знаний человека и социально-коммуникационной средой;
- социально-коммуникационной средой и цифровой электронной средой;
- ментальной средой знаний человека и цифровой электронной средой.

Определение двуединых понятий на третьей границе, которые в используемой системе терминов названы семокодами, дало возможность предложить методы компьютерного кодирования тех концептов, которые не имеют вербальных форм представления или имеют несколько разных форм их представления из-за асимметрии знаковых систем.

Решение двух названных нами проблем было найдено в результате использования и развития подходов Горна и Шрейдера к определению информационно-компьютерной науки как единой области знания. Отметим, что при решении первой проблемы (концептуальное индексирование геоизображений) использовались все четыре среды, а при решении второй (генерация систем экспертных знаний) – только три среды из четырех. Однако отсюда не следует, что для описания предметной области информационно-компьютерной науки всегда будет достаточно именно этих четырех сред.