

TO THE QUESTION OF THE CONCEPT OF KNOWLEDGE BASE FOR THE FORMATION OF BIG DATA

Utepbergenova G.H.¹, Baltabaeva R.B.² (Republic of Uzbekistan)

Email: Utepbergenova348@scientifictext.ru

¹Utepbergenova Gulziyra Habibullaevna - Assistant Teacher;

²Baltabaeva Rano Bekbaulievna - Assistant Teacher,

DEPARTMENT OF APPLIED MATHEMATICS,

KARAKALPAK STATE UNIVERSITY,

NUKUS, REPUBLIC OF UZBEKISTAN

Abstract: the current period of time is characterized by the total introduction of computer technology management and information processing in almost all spheres of human activity. In practice, these technologies are implemented in the form of information and communication systems (ICS), a typical representative of which is the State Automated Systems of a higher educational institution of the Republic of Uzbekistan. An integral part of this type of system is the knowledge base (KB). Currently, to build them, ready-made software products such as ORACLE, MSSQL, DB2 and others are used, which are supplemented by various software modules in which the knowledge of the problem area is laid according to the "natural language computer program" scheme.

Keywords: knowledge, data, declarative knowledge, subject data, procedural knowledge, linguistic knowledge.

К ВОПРОСУ ПОНЯТИЯ БАЗЫ ЗНАНИЙ ДЛЯ ФОРМИРОВАНИЯ ОБЪЕМНЫХ ДАННЫХ

Утепбергенова Г.Х.¹, Балтабаева Р.Б.² (Республика Узбекистан)

¹Утепбергенова Гульзира Хабибуллаевна - ассистент преподавателя;

²Балтабаева Рано Бекбаулиевна - ассистент преподавателя,

кафедра прикладной математики,

Каракалпакский государственный университет,

г. Нукус, Республика Узбекистан

Аннотация: текущий период характеризуется тотальным внедрением компьютерных технологий управления и обработки информации практически во все сферы человеческой деятельности. На практике эти технологии реализуются в виде информационно-коммуникационных систем (ИКС), типичным представителем которых является Государственная автоматизированная система высшего учебного заведения Республики Узбекистан. Неотъемлемой составной частью систем этого типа является база знаний (БЗ). В настоящее время для их построения используются готовые программные продукты типа ORACLE, MSSQL, DB2 и другие, дополняемые различными программными модулями в которые по схеме «естественный язык - компьютерная программа» закладываются знания о проблемной области.

Ключевые слова: знания, данные, декларативные знания, предметные данные, процедурные знания, лингвистические знания.

При изучении интеллектуальных систем традиционно возникает вопрос — что же такое знания и чем они отличаются от обычных данных, десятилетиями обрабатываемых ЭВМ. Можно предложить несколько рабочих определений, в рамках которых это становится очевидным [1, 19-20].

Данные — это отдельные факты, характеризующие объекты, процессы и явления предметной области, а также их свойства. При обработке на ЭВМ данные трансформируются, условно проходя следующие этапы:

1. D1 — данные как результат измерений и наблюдений;
2. D2 — данные на материальных носителях информации (таблицы, протоколы, справочники);
3. D3 — модели (структуры) данных в виде диаграмм, графиков, функций;
4. D4 — данные в компьютере на языке описания данных;
5. D5 — базы данных на машинных носителях информации.

Знания основаны на данных, полученных эмпирическим путем. Они представляют собой результат мыслительной деятельности человека, направленной на обобщение его опыта, полученного в результате практической деятельности.

Знания — это закономерности предметной области (принципы, связи, законы), полученные в результате практической деятельности и профессионального опыта, позволяющие специалистам ставить и решать задачи в этой области.

При обработке на ЭВМ знания трансформируются аналогично данным.

1. Z1 — знания в памяти человека как результат мышления;
2. Z2 — материальные носители знаний (учебники, методические пособия);
3. Z3 — *поле знаний* — условное описание основных объектов предметной области, их атрибутов и закономерностей, их связывающих;
4. Z4 — знания, описанные на языках представления знаний (продукционные языки, семантические сети, фреймы);
5. Z5 — база знаний на машинных носителях информации.

Часто используется такое определение знаний.

Знания — это хорошо структурированные данные, или данные о данных, или метаданные.

Существует множество способов определять понятия. Один из широко применяемых способов основан на идее интенционала. *Интенционал* понятия — это определение его через соотнесение с понятием более высокого уровня абстракции с указанием специфических свойств. Интенционалы формулируют знания об объектах. Другой способ определяет понятие через соотнесение с понятиями более низкого уровня абстракции или перечисление фактов, относящихся к определяемому объекту. Это есть определение через данные, или *экстенционал понятия*.

Знания могут быть классифицированы по следующим категориям:

- *Поверхностные* — знания о видимых взаимосвязях между отдельными событиями и фактами в предметной области.

- *Глубинные* — абстракции, аналогии, схемы, отображающие структуру и природу процессов, протекающих в предметной области. Эти знания объясняют явления и могут использоваться для прогнозирования поведения объектов.

Неотъемлемой составной частью современных и перспективных ИКС являются базы знаний, под которыми понимается управляемый комплекс языковых, алгоритмических, программных и технических средств, предназначенных для восприятия, обработки, хранения и отображения знаний о предметной области, включающий четыре компонента [2, 104-105]:

- a) упорядоченные каким-либо способом факты и данные, отражающие модель профессиональной сферы (предметные данные);

- b) правила, модели, алгоритмы и программы, позволяющие рассчитывать определенные показатели объектов профессиональной сферы, строить цепочки логических выводов и на этой основе делать обобщения и заключения, а также вызывать определенные ассоциации (декларативные знания);

- c) управляющая и интерпретирующая структура, определяющая порядок и способы применения моделей и правил логического вывода для получения или трансформации информации (процедурные знания);

- d) правила морфологического, синтаксического и семантического анализа входных и выходных текстов, а также списки основ слов, которые используются для организации диалога между пользователем и БЗ (лингвистические знания).

На основе этих теорий решаются следующие задачи:

- определяются принципы построения БЗ в составе ИКС типа ВУЗ КГУ в части организации общения между пользователем и БЗ, модели машинного понимания входных текстов.

- исследуются языковые средства с учетом особенностей крупномасштабных ИКС. Рекомендуется перечень языковых средств для реализации БЗ в составе ИКС.

- разрабатываются алгоритмы интеллектуальной поддержки решений по созданию баз знаний, которые позволяют формализовать знания о предметной области с использованием методов задания единиц знаний.

- решается проблема оценки качества проектных решений по созданию БЗ путем оценки по показателям, учитывающим социальную, функциональную, эргономическую, прагматическую, техническую и технологическую составляющие.

Список литературы / References

1. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. Учебник. Санкт-Петербург, 2000. 382 с.
2. Самков Е.Ю. Методы интегральной оценки качества крупномасштабных ИКС / В.В. Ключков, Е.Ю. Самков / Сб. научн. тр. Десятой науч.-практ. конф «Современные информационные технологии в управлении и образовании». М: ФГУП НИИ «Восход», 2011. С. 104-107.