

20. Two-stage sintering of nano-sized yttria stabilized zirconia process by powder injection moulding / P. C. Yu, Q. F. Li, J. Y. H. Fuh et al. // Journ. of Mater. Proc. Technology. – 2007. – 192-193. – P. 312–318.
21. Fukushima H., Matsui M. Reviews of Microwave Processing of Ceramics in Japan // Amer. Cer. Soc., Ceramic Transactions. – 2002. – 133. – P. 201–210.
22. Microwave technique applied to the hydrothermal synthesis and sintering of calcia stabilized zirconia nanoparticles / A. Rizzuti, A. Corradi, C. Leonelli et al. // Journ. of Nanoparticle Res. – 2010. – 12. – № 1. – P. 327–335.

Поступила 12.04.10

УДК 681.324.06

В. Н. Колодницкий, канд. физ.-мат. наук; **В. Н. Кулаковский**, **А. А. Лебедева**,
Е. М. Чистяков, **К. З. Гордашник**, кандидаты технических наук; **И. В. Скворцов**

Институт сверхтвердых материалов им. В. Н. Бакуля НАН Украины, г. Киев

ЭЛЕКТРОННЫЙ ИНФОРМАЦИОННО-ПОИСКОВЫЙ ТЕЗАУРУС ПРЕДМЕТНОЙ ОБЛАСТИ «СВЕРХТВЕРДЫЕ МАТЕРИАЛЫ» КАК ФОРМАЛЬНАЯ СИСТЕМА ЗНАНИЙ

An electronic multilingual information retrieval thesaurus for the system of materials science of the Superhard Materials ontology that includes the concepts of the problem-oriented knowledge, which correspond to the scientific directions, developed at the Bakul Institute for Superhard Materials, National Academy of Sciences of Ukraine has been created.

Знания наряду с финансовыми и материальными ресурсами являются одной из важнейших основ развития как отдельной организации, так и государства в целом. Организации, которые быстрее других создают и находят новые знания, обеспечивают их сохранение и оперативное использование сотрудниками, воплощают системы управления знаниями в практическую инновационно-инвестиционную деятельность, приобретают конкурентные преимущества. Интеллектуальный капитал и интеллектуальная собственность получают все возрастающее признание в качестве нового источника богатства научных учреждений.

Практика показывает, что наиболее действенным средством овладения научными сотрудниками новых знаний мирового уровня в разных проблемно-ориентированных областях являются современные программные системы класса информационно-интеллектуальных технологий.

В настоящее время в Институте сверхтвердых материалов им. В. Н. Бакуля НАН Украины (ИСМ) для всестороннего использования мировых информационных потоков осуществляется процесс интеграции и трансформации существующих возможностей информационно-поисковых систем, а также корпоративных баз данных и баз знаний в единую информационно-интеллектуальную экспертную систему (ИИЭС). Для успешного функционирования ИИЭС на основе Интернет- и Web-технологий разрабатывается новая технологическая программная инфраструктура – информационно-интеллектуальная система управления знаниями (ИИСУЗ), которая позволит преодолеть коммуникационное и пространственно-географическое препятствия на пути общего и оперативного использования знаний ИСМ и мирового уровня его сотрудниками.

С этой целью разработан унифицированный метод формирования модели знаний предметной области (ПрО) «Сверхтвердые материалы» [1] и использования экспертной системы по

работе со знаниями для решения таких базовых задач, как семантический поиск, классификация, навигация и формирование эффективных рекомендаций для пользователей знаний.

Одним из главных принципов построения данной модели является использование электронного тезауруса (словаря терминов) для описания характеристик объектов и процессов, точных и однозначных определений всех терминов этого словаря и классификации логических взаимосвязей этих терминов.

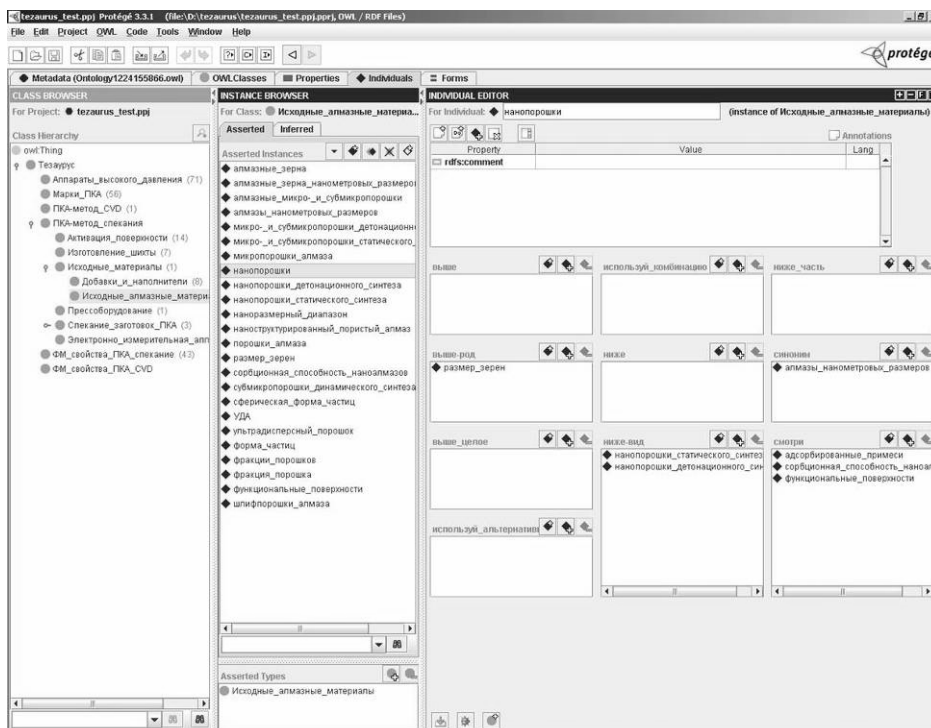
Тезаурус – это особая разновидность словаря специальной лексики, где указаны семантические отношения (синонимы, антонимы, паронимы, гипонимы, гиперонимы и пр.) между лексическими единицами. В отличие от толковых словарей, тезаурус позволяет выявить смысл не только с помощью термина, но и посредством соотнесения слова с другими понятиями и их группами. Следовательно, тезаурус, особенно в электронном формате, – один из действенных инструментов для описания и исследования научной терминологии и логико-понятийной структуры отдельной научной области, в частности «Сверхтвердые материалы».

Главный прием построения тезауруса – обеспечение однозначности терминов. Каждый термин в тезаурусе представляется на естественном языке и понимается только в каком-либо одном значении – дескрипторе. Дескриптор – это лексическая единица, выраженная информативным словом (вербально) или кодом и являющаяся именем класса синонимичных или близких по смыслу ключевых слов. Многочисленным словам естественного языка соответствуют несколько дескрипторов, а нескольким синонимичным словам и выражениям – один дескриптор. Если определенное понятие имеет два или больше толкований (синонимов), только одно из них условно выбирается в качестве дескриптора. Все другие синонимичные толкования, которые могли бы служить поисковыми элементами, имеют ссылку на этот дескриптор. Если термин полисемичный (многозначный) или омонимичный (звуковое совпадение разных по значению терминов), к нему приводится уточнение. Дескрипторы используют для координатного индексирования документов и информационных запросов в целях последующего поиска.

Информационно-поисковые тезаурусы (ИПТ), построенные для совокупности областей знания, называются политематическими (например, [2]), а отражающие ограниченную область знания или практической деятельности – специализированными [3; 4].

Повышение статуса английского языка как международного языка профессиональной коммуникации и языка науки обуславливает необходимость построения, как правило, двуязычного тезауруса. В ИСМ разработан специализированный электронный тематический русско-английский ИПТ ПрО «Сверхтвердые материалы», который является словарем дескрипторного информационно-поискового языка для координатного индексирования документов и информационных запросов посредством дескрипторов и/или ключевых слов с зафиксированными в нем парадигматическими отношениями лексических единиц. В созданном словаре дескрипторы и их синонимы находятся в общем алфавитном ряду. Термины (слова и словосочетания) ПрО «Сверхтвердые материалы» отражены в виде иерархических лексических шаблонов, для чего создан иерархический указатель – список дескрипторов высшего уровня иерархии, где для каждого из дескрипторов приводятся подчиненные дескрипторы, расположенные в порядке убывания общности. Такая иерархия понятий использована для интеграции тезауруса и онтологии понятий в единую модель ПрО. Для фиксации значимых отношений между терминами выделяли основные связи между ними, служащие основой проектирования онтологии [5; 6].

Развитие компьютерных и телекоммуникационных технологий, информатизация науки требуют создания новых форм существования научно-технического лексикографического источника по сверхтвердым материалам. Исходя из этого онтология и тезаурус «Сверхтвердые материалы» спроектированы в виде электронной версии с использованием системы Protégé [7]. Общий вид управляющего интерфейса тезауруса показан на рисунке.



Управляющий интерфейс ИПТ ПрО «Сверхтвердые материалы» в системе Protégé

Порядок построения ИПТ ПрО «Сверхтвердые материалы» выбрали следующий:
определение тематического охвата ПрО;
сбор массива лексических единиц;
формирования слóвника (алфавитный список, реестр слов в словаре) ИПТ;
построение словарных статей и указателей;
ведение ИПТ;
экспертиза ИПТ.

Тезаурус разработали на основе специальной научно-технической литературы. При отборе документов для сбора лексических единиц использовали источники, касающиеся проблем сверхтвердых материалов: шеститомную монографию «Сверхтвердые материалы. Получение и применение» [8]; научные статьи журнала «Сверхтвердые материалы», отчеты ИСМ о выполнении научно-исследовательских работ; терминологические, толковые, энциклопедические и политехнические словари; классификаторы научно-технической информации; таблицы УДК, МКИ, тематические рубрикаторы; термины и определения стандартов; описания изобретений.

При формировании слóвника ИПТ использовали следующие типы лексических единиц: единичные слова, именные словосочетания, аббревиатуры и лексически значимые компоненты сложных слов. Именные словосочетания включены в слóвник как наиболее употребляемые в области сверхтвердых материалов (СТМ) или выражающие смысловое содержание процессов, материалов, свойств, оборудования. Например, сверхтвердые материалы, аппараты высокого давления, межзеренная граница, фазовые превращения. Разбиение таких словосочетаний на компоненты приводит к потере важных для поиска парадигматических связей.

Аббревиатуры использованы в тезаурусе в качестве синонимов, которые отражают марки сверхтвердых материалов, широко применяемых в производстве тугоплавких соединений, а также аппаратов высокого давления для синтеза СТМ, и в списке идентификаторов отражены полными развернутыми формами их значений. Например, АД – аппарат высокого давления, ПКА – поликристаллический алмаз, АКТМ – алмазный композиционный термостойкий материал, АТП – алмазно-твердосплавная пластина, УДА – ультрадисперсный детонационный алмаз.

Тезаурус состоит из двух частей – лексико-семантического указателя и дополнительных указателей, включающих систематический и алфавитный, а также списка идентификаторов.

Информационно-поисковый тезаурус по СТМ подготовлен в соответствии с требованиями ДСТУ 4032–2001 (ISO 2788:1986) «Інформація та документація. Одномовний тезаурус. Методика розроблення» [9], по построению ИПТ, содержащего лексические единицы, взятые из одного языка, и ДСТУ 4031–2001 (ISO 5964:1985) «Інформація та документація. Багатомовний тезаурус. Методика розроблення» [10], согласно которому разрабатываются ИПТ, содержащие лексические единицы, взятые из нескольких естественных языков. Многоязычный ИПТ представляет эквивалентные по смыслу понятия на русском и английском языках и предназначен для обработки документов и информационного поиска в целях обмена информацией на указанных языках.

Ведение тезауруса с использованием системы Protégé [4] предусматривает следующие принципы:

- введение тезауруса, его пополнение, коррекцию терминов;
- включение, выключение и консервацию элементов тезауруса;
- включение отношений между дескрипторами и др.

Построенный на приведенных принципах электронный тезаурус представляет специальным образом организованный и контролируемый словарь. Основанием для определения места работы в классификации, а впоследствии – для формирования поискового образа информационного источника служит соответствие ключевых слов отдельного документа содержанию тезауруса. При этом особое значение для научных сотрудников приобретает такое свойство тезауруса, как отражение содержания науки и практики об СТМ в определенных границах. В то же время тезаурусная форма представления знания об СТМ предполагает анализ и при необходимости совершенствование тезауруса как в техническом, так и в семантическом отношении. Именно наличие такого тезауруса позволяет осуществлять интеллектуальный анализ текстовой информации в области создания и применения СТМ [11]. К тому же работа по созданию тезаурусов в значительной степени опирается на использование компьютерной техники, что значительно облегчает обработку первичных документов и выполнение стандартных аналитических процедур при широкой программной поддержке. Обмен информацией, аккумулированной в системе Protégé, через формат XML (Extensible Markup Language) [12] позволяет гибко присоединять тезаурус к другим анализирующим текстовые документы программным продуктам.

Разработанный ИПТ «Сверхтвердые материалы» предназначен для специалистов, использующих ИПС по СТМ, в частности, может использоваться при составлении поисковых предписаний для поиска информации в ИПС, в базах данных по СТМ, включающих информацию о классах материалов. Этот тезаурус могут использовать также ученые при написании статей (для выбора терминов-дескрипторов) и при поиске информации о новых материалах по предметным указателям к реферативным журналам и другим изданиям по материаловедению. Кроме того, он может служить терминологическим руководством при реферировании, редактировании и переводе литературы по сверхтвердым материалам.

В связи с активным развитием технической киберлексикографии как науки об использовании телекоммуникационных технологий и Интернета в разработке и создании электронных технических лексикографических источников, разработанный электронный тезаурус можно представить в формате HTML (HyperText Markup Language) для использования в режиме on-line как в корпоративной сети, так и в глобальной сети Интернет.

Таким образом, ИПТ «Сверхтвердые материалы» является не просто инструментом для информационного поиска. Его можно рассматривать как универсальную модель терминологической системы, а потому – как формальную систему знаний, содержащихся в языке научной области «Сверхтвердые материалы».

Литература

1. Интегрированная модель предметной области «Сверхтвердые материалы» / В. Н. Кулаковский, А. А. Лебедева, И. В. Скворцов и др. // Сверхтвердые материалы. – 2009. – № 5. – С. 90–91.
2. Інформаційно-пошуковий тезаурус. – Донецьк: Альфа-Прес, 2004. – 1104 с.
3. Информационно-поисковый тезаурус классов органических соединений. – М.: Наука, 1986. – 157 с.
4. Єршов В. М., Рукас Т. П. Міні-тезаурус з металознавства та термічного оброблення металів: навч. посіб. для студ. вищих навч. закл. / Донбаський гірничо-металургійний ін-т. – Алчевськ, 2002. – 144 с.
5. Лебедева А. А., Фидаров Т. З., Скворцов И. В. Подход к построению предметной онтологии для экспертной системы “Сверхтвердые материалы” // Породоразрушающий и металлообрабатывающий инструмент – техника и технология его изготовления и применения: сб. науч. тр. – К.: ИСМ им. В. Н. Бакуля НАН Украины, 2007. – Вып. 10. – С. 370–374.
6. Онтологический подход к построению базы знаний “Сверхтвердые материалы” / В. Н. Кулаковский, А. А. Лебедева, К. З. Гордашник и др. // Искусств. интеллект. – 2008. – № 4. – С. 91–102.
7. http://protege.stanford.edu/publications/ontology_development/ontology101.html.
8. Сверхтвердые материалы. Получение и применение: Моногр.: В 6 т. / Под ред. Н. В. Новикова. – К.: ИСМ им. В. Н. Бакуля НАН Украины, 2003–2007.
9. ДСТУ 4032–2001; ISO 2788:1986. Інформація та документація. Одномовний тезаурус. Методика розроблення / А. Стогній (розроб.). – Чин. від 01.01.02. – К.: Держстандарт України, 2001. – 60 с.
10. ДСТУ 4031–2001; ISO 5964:1985. Інформація та документація. Багатомовний тезаурус. Методика розроблення / А. Стогній (розроб.). – Чин. від 01.01.02. – К.: Держстандарт України, 2001. – 74 с.
11. Интеллектуальный анализ текстовой информации в системе управления базами знаний в области создания и применения сверхтвердых материалов / А. А. Лебедева, В. В. Цегельнюк, В. Н. Кулаковский и др. // Синтез, спекание и свойства сверхтвердых материалов: Сб. науч. тр. – К.: ИСМ им. В. Н. Бакуля, 2010. – С. 169–178.
12. Маршал Б. XML в действии: Пер. с англ. – М.: Триумф, 2002. – 365 с.

Поступила 03.06.10