

Систематизация данных по наноматериалам: онтологический подход

Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А.
ОИВТ РАН, Москва

Онтологическое моделирование, обеспечивая систематизацию и интеграцию данных в предметных областях со сложной и эволюционирующей структурой, компенсирует ограниченность баз данных (БД), использующих жесткую логическую структуру. Преодолевается и другой недостаток БД - многообразие моделей и форматов, затрудняющих интеграцию разнородных источников. Наметившееся в последние годы решение указанных проблем состоит в построении онтологии, то есть единой модели предметной области и единого словаря для определения смысла данных. В [1] предложены онтологии применительно к данным по свойствам веществ и материалов.

Специфика наноматериалов предъявляет дополнительные требования к систематизации [2], что привело к таксономии понятий с выделением главных классов: Object, Property, Provenance, Application. Класс Object имеет 2 субкласса, Nanoobjects и NanostrMaterials, отвечающих наноструктурам (кластеры, нанотрубки и т.п.) и объемным материалам. Второй субкласс является родительским по отношению к NanoComp для нанокомпозитов, тип которых определяет как матрица, так и наполнитель в виде наночастиц. Помимо таксономии, особенность наномира отражается в атрибутах каждого класса, обеспечивающих детализированное представление наноматериала: пространственная размерность, химическая природа, структурные особенности и т.п. Соответствующая таксономия разработана также и для класса Application с выделением двух субклассов Advice и Technology, свойства которых позволяют соотнести инновацию с рубриками классификаторов, например, принятых на Федеральном портале «Нанотехнологии и наноматериалы». Особую значимость для наноматериалов имеет класс Provenance для детализации методов синтеза, диагностики и сертификации. Наряду с главными, онтология включает совокупность вспомогательных классов для детализации типа данных, их неопределенности, формата и проч. Будучи представленной в виде формализованной записи на OWL, онтология обеспечивает возможность семантической интеграции данных, то есть связывания из множества источников с унификацией смысла терминов.

Работа выполнена при поддержке РФФИ, проекты №№ 13-07-00218, 13-08-00404.

1. Еркимбаев А.О., Жижченко А.Б., Зицерман В.Ю., Кобзев Г.А., Сон Э.Е., Сотников А.Н. Научно-техническая информация. Сер. 2. Информационные процессы и системы. ВИНТИ РАН, 2012, №8, стр. 1-8.
2. Eletsii A.V., Erkimbaev A.O., Kobzev G.A., Trachtengerts M.S., Zitserman V.Y. Data Science Journal. – 2011, V. 11. P.126-139.