

Разработка технологии обработки информационных ресурсов: прикладные аспекты

Н.Г. Журавлева,
н.с., zhurav@ipu.ru,
ИПУ РАН, г. Москва,
А.А. Мелихов,
megadelphin@mail.ru,
АО ЦНИИ ЭИСУ, г. Москва,
Е.И. Кублик,
к.т.н., доц., ewkub@mail.ru.
Финансовый университет, г. Москва

Настоящая работа посвящена прикладным аспектам формирования информационных ресурсов при организации системы жизненного цикла продукта. Рассматривается проблемный вопрос организации единого информационного пространства в рамках PDM-системы с уточнением классификации видов информационных ресурсов. Предлагается практический подход в организации информационных ресурсов на основе их семантической близости.

The former article is devoted to information resource management in product lifecycle. Main focus is on creating the homogeneous information exchange space for PDM-system. Information resources of PLM are classified. The approach for information resource management, based on semantic conformity, is proposed.

Введение

Управление жизненным циклом (ЖЦ) промышленных изделий включает в себя совокупность процессов, реализуемых от возникновения потребности в создании продукции до момента её утилизации ввиду утраты определённых потребительских свойств. Очевидно, что появление комплексного подхода к организации материального производства является результатом повышения сложности как самих технологических процессов, так и отношений между поставщиками сырья, производителями продукции и её непосредственными потребителями [1]. Учёт потребностей общества, равно как и моделирование процесса изменения свойств продукции в процессе эксплуатации позволяет значительно оптимизировать производственные и логистические затраты, что, в конечном итоге значительно снижает стоимость готовых изделий. В контексте автоматизации управления ЖЦ, как правило, рассматривается совокупность мер, направленных на создание единого информационного пространства, в рамках которого реализованы базовые процессы накопления, хранения, обработки, формализации, а также взаимной интеграции данных различного типа: производственной и эксплуатационной документации, законодательных актов и стандартов и т.п. [2]. В практике, указанная задача решается посредством PDM-подсистем (англ. Product Data Management — управление данными об изделии) систем PLM (англ. Product Lifecycle Management – управление жизненным циклом изделия).

Основную сложность разработки PLM-систем является высокий уровень гетерогенности исходной информации. В его состав входят документы различного рода и назначения – технические и технологические, юридические, нормативные и т.п. Указанные документы имеют ограниченные сроки применения (как правило – определённый этап жизненного цикла) и предназначены для специалистов в различных предметных областях, обладающих различными компетенциями. Целью настоящей работы является представление подхода к разработке информационного обеспечения, преимущественно направленного на снижение степени гетерогенности информационного обеспечения.

1. Информационные ресурсы PLM-системы

Традиционный подход к организации информационного обеспечения предполагает иерархию источников в соответствии с этапами жизненного цикла. Например, для этапа проектирования оно будет включать соответствующие стандарты, описывающую содержательную часть проектную документацию, а также дополнительные источники информации, состав которых явно не регламентируется ГОСТом (словари и тезаурусы, базы данных и знаний, классификаторы и т.п.). При этом совокупность документов для определённого этапа образует замкнутую систему (рис. 1):

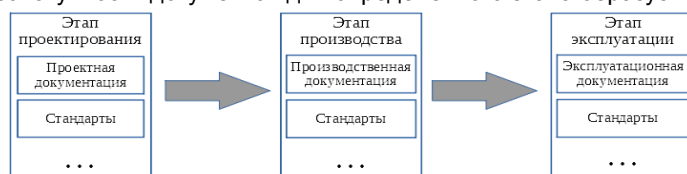


рис. 1 Информационное обеспечение этапов жизненного цикла

Однако практический опыт показывает, что многообразие информационных ресурсов PLM-системы не ограничивается непосредственно проектной документацией и совокупностью производственных стандартов. Современные технические проекты создаются большими коллективами, зачастую не связанными между собой непосредственно. На определённых этапах производства требуется документация на различные части - детали, узлы и агрегаты, программные модули. Также, стоит выделить в отдельный класс ресурсов информационно-лингвистическое обеспечение (ИЛО) средств автоматизации производства, также оказывающее существенное влияние на технологический процесс. Например, ИЛО может использовать отличные от принятых на производстве обозначений и/или определять специфическую форму графического интерфейса.

Таким образом, для каждого этапа производства необходимы следующие виды информационных ресурсов: нормативная документация, документация на составные части изделия, собственная документация на изделие, предметно-специфичные ИР (базы знаний, тезаурусы и т.п.), общие ИР.

Под общими ИР в данном случае понимаются информационные ресурсы, содержащие в себе так называемые «онтологии здравого смысла»¹. Например, в качестве общего ресурса может выступать онтология физических величин [3]. Фактически, на уровне PDM доступны все относящиеся к изделию информационные ресурсы, однако их иерархизация в соответствии с этапами жизненного цикла приводит к появлению противоречия. С одной стороны, ресурсы доступны в полном объёме, с другой стороны, данные разнообразны и разнородны, что, в свою очередь приводит к увеличению избыточности и усложнению контроля за их целостностью. Для разрешения указанного противоречия предполагается новый подход к интеграции информационных ресурсов в PLM-системе с учётом их предметной области и способа формализации.

2. Интеграция информационных ресурсов на основе их семантико-лингвистических свойств

Вся совокупность информационных ресурсов жизненного цикла продукта представляет собой сложную понятийную систему, образующую дискурс. В основе разработанного подхода лежит организация элементов дискурса на основе их семантико-лингвистических свойств. В данном контексте, семантические свойства ИР, такие как предметная область и цели формирования ресурсов рассматриваются в экстралингвистическом контексте. К семантико-лингвистическим свойствам относятся: язык формализации документа, предметная область, связи с внешними документами и ресурсами (цитирование, обобщение, уточнение), назначение документа (связь с целями создания и адресатом), условия применимости (связь с этапом ЖЦ).

Все доступные для PLM-системы документы подлежат классификации согласно значениям перечисленных свойств, а затем производится их группировка (рисунок 2):

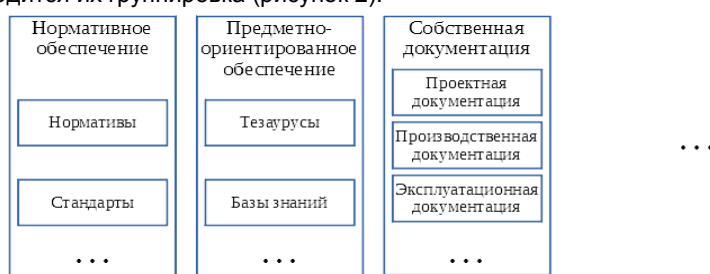


рис. 2 Организация информационных ресурсов по целевому назначению

При этом данная классификация должна основываться на результатах наблюдения за параметрами конкретных информационных ресурсов, т.е. носить дескриптивную² форму. Подробное наблюдение за значениями параметров и анализ их взаимных зависимостей позволяет создавать более компактные метаописания ресурсов [4], которые затем могут быть представлены в более компактные свёртки [5].

Заключение

Таким образом, основной проблемой, затрудняющей создание единого информационного пространства в рамках PDM-системы, является высокий уровень гетерогенности информационных источников. В рамках настоящей публикации разработан подход, основанный на объединении информационных ресурсов согласно семантико-лингвистическим особенностям. Новизна подхода заключается в переходе от группировки ИР по этапам разработки в пользу создания общего информационного ресурса, элементы которого структурированы по принципу семантического соответствия. Применение разработанного подхода позволяет снизить накладные затраты на создание, изменение и верификацию информационных ресурсов за счёт повышения степени их семантической связности.

Литература

1. Павловский И.С. Концептуальные исследования проблемы интеграции систем управления технологическими процессами / Труды 15-й Международной конференции «Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта» (CAD/CAM/PDM-2015, Москва). М.: ООО "Аналитик", 2015. С. 67-70.
2. Нечаев В.В., Гончаренко В.И., Рожнов А.В., Лычев А.В., Лобанов И.А. Интеграция компонентов виртуальной семантической среды и обобщенной модели анализа среды функционирования // Современные информационные технологии и ИТ-образование. 2016. Т. 12 (№ 3), часть 1. С. 187-194.
3. Журавлева Н.Г., Мелихов А.А., Губин А.Н., Гудов Г.Н. Синтаксический анализ и преобразование единиц измерения в предложениях естественного языка информационных ресурсов / Труды 15-й Международной конференции «Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта» (CAD/CAM/PDM-2015, Москва). М.: ООО "Аналитик", 2015. С. 341-344.
4. Melikhov A.A., Rozhnov A.V. Vectorizing textual data sources to decrease attribute space dimension // Proceedings of the 10th International Conference "Management of Large-Scale System Development" (MLSD). 2017. - URL: <http://ieeexplore.ieee.org/document/8109662/>.
5. Мелихов А.А., Лобанов И.А. Использование кривых гильберта при выборе вычислительных узлов распределённой системы обработки естественного языка // Материалы 11-й Международной конференции «Управление развитием крупномасштабных систем» (MLSD'2018, Москва). М.: ИПУ РАН, 2018 (в печати).

¹ Т.к. в большинстве доступных русскоязычных публикациях на сегодняшний день устоявшаяся терминология отсутствует, - в данном случае использован прямой перевод выражения «commonsense ontology».

² От англ. «description» - описание, противопоставление предскриптивному, от слова «prescription» - предписание.