

ИНЖЕНЕРИЯ СПРАВОЧНЫХ ДАННЫХ В СООТВЕТСТВИИ С НОТАЦИЕЙ СТАНДАРТА ISO 15926

Винников А.Н.

Научный руководитель: Мирошниченко Е.А.

Томский политехнический университет

634050, Россия, г. Томск, пр-т Ленина, 30

E-mail: franzkz@tpu.ru

Введение

Задача интеграции данных возникает в тот момент, когда какие-то данные требуются прикладной программе, не приспособленной для работы с «посторонним» для неё хранилищем. Так, ERP-системе очень непросто получить доступ к данным PLM-системы, и наоборот. Схемы баз данных разных приложений не соответствуют друг другу: что является «атрибутом» для одного хранилища – может быть «объектом» другого, что представлено текстовой строкой в одной схеме данных – может оказаться группой текстовых и числовых полей в другой. Но потребность в передаче данных между разными хранилищами становится всё больше, ибо повторный ручной ввод данных из приложения в приложение пожирает дорогое время квалифицированных сотрудников, является источником ошибок, а количество таких перенабивок в большом проекте может приближаться к десяткам и сотням. Для формализации передачи данных инженерного проекта и промышленных каталогов между хранилищами информации компьютерных приложений разработан стандарт ISO 15926, который предназначен для применения на протяжении всего жизненного цикла инженерного проекта.

Структура ISO 15926

Изначально данный стандарт разрабатывался для применения в нефтегазовой промышленности, но на данный момент он считается достаточно универсальным для применения во всех отраслях промышленности.

Стандарт состоит из нескольких частей. Первые две из них адаптированы в России в качестве ГОСТ Р ИСО 15926[1]. Ниже приведено назначение основных частей стандарта.

Часть 1. Обзор и функциональные принципы.

Часть 2. Модель данных. Эта часть содержит 201 тип, к которым должны быть отнесены все сущности, участвующие в онтологиях, построенных по принципам стандарта.

Часть 3. Справочные данные для геометрии и топологии.

Часть 4. Первоначальный набор справочных данных.

Часть 7. Шаблоны (информационные структуры, предназначенные для выражения сведений о событиях, сложных связях и т.д.

Часть 8. Реализация ISO 15926 средствами синтаксиса OWL.

Часть 9. Интеграция распределенных систем (описывает технологический принцип передачи данных между информационными системами; подразумевает использование SPARQL точек доступа для взаимодействия между приложениями. Каждая система размещает данные, которыми готова поделиться с другими, в точку доступа SPARQL, называемую «фасадом». Другие системы могут обращаться к фасаду с запросом информации, или, наоборот, помещать в него какие-либо сведения, используя SPARQL-запросы).

Часть 10. Методы тестирования.

Часть 11. Руководство по промышленному применению [2].

Инженерия справочных данных

Существует два основных направления использования стандарта ISO 15926:

- использование стандарта для передачи данных из хранилища в хранилище или «ISO 15926 outside», именно это направление используется в инженерии справочных данных;
- использование стандарта в самих хранилищах данных компьютерных приложений – «ISO 15926 inside» – это направление на данный момент нигде не реализовано в полной мере, то есть форматы данных в хранилищах не соответствуют ISO 15926.

Инженерия справочных данных в рамках ISO 15926 является одной из дисциплин работы с нормативно-справочной информацией (НСИ, основными данными, master data) в рамках работы по поднятию качества данных и обеспечению управления основными данными. Принципиальное отличие этой технологии от существующих на данный момент технологий управления корпоративной НСИ (enterprise MDM, master data management) в том, что стандарт ISO 15926 изначально предполагает использование справочных данных в независимых друг от друга организациях (а не просто декларирует «пересечение справочными данными границ организации» с последующим обсуждением проблем только одной организации). Согласно этому стандарту доступ к НСИ осуществляется в едином формате, а сами НСИ содержатся в совокупности («федерации») административно независимых библиотек справочных данных. Поэтому технология ISO 15926 пригодна для интеграции данных не только в масштабах крупной инжиниринговой компании или холдинга, но и для интеграции данных в масштабах отрасли или даже крупного межотраслевого проекта.

Виды справочных данных

Справочные данные (reference data) – это такие данные, которые используются в нескольких проектах, на нескольких стадиях жизненного цикла и/или интересуют множество пользователей.



Рис. 1. Пирамида данных жизненного цикла

Согласно ISO 15926, они состоят из единиц справочных данных (reference data items). Такими единицами справочных данных являются:

- классы индивидов (class of individuals), классы отношений (class of relationships) и классы классов (class of class);
- отдельные индивиды (individuals);
- шаблоны (templates);
- экземпляры шаблонов (template instances).

По степени общности отражаемой информации окружающего мира среди справочных данных выделяется несколько категорий, составляющих пирамиду, приведенную на рисунке 1[3]:

- «Модель данных» – это 201 тип Части 2 ISO 15926;
- основные классы (core classe);
- стандартные классы (standard classes);
- специализированные классы;
- протошаблоны (proto templates);
- базовые шаблоны (base templates);
- основные шаблоны (core templates);
- специализированные шаблоны (specialized templates).

Жизненный цикл справочных данных

В процессе работы инженерного проекта справочные данные согласно стандарту ISO 15926 проходят следующие этапы жизненного цикла [4]:

- выявление набора проектных данных, который нужно будет перевести в форму, соответствующую ISO 15926. Выходными рабочими продуктами стадии являются наборы данных, которые подлежат характеристике, и указание на то, в каком контексте будут использованы результаты всей работы: из какого хранилища будут извлекаться эти наборы данных, и в какое хранилище потом попадать;
- выявление единиц описания данных хранилища – на этой стадии определяются, какие именно данные будут передаваться из одного хра-

нилища в другое, чтобы обеспечить передачу данных для каждого хранилища информации, найденного на предыдущей стадии (как источника данных, так и места назначения);

- характеристика набора идентифицированных единиц описания данных. На этой стадии создаются справочные данные ISO 15926, которые затем используются при мэппинге. Это самая сложная и трудозатратная часть работы. Входной рабочий продукт этой стадии – выделенные на предыдущей стадии единицы описания данных;
- мэппинг – на этой стадии прописывается соответствие выбранных единиц справочных данных ISO 15926 единицам описания проектных данных хранилища;
- передача проектных данных. Осуществить передачу проектных данных между хранилищами становится возможным тогда, когда для обоих хранилищ есть мэппинги единиц описания данных к одному и тому же набору справочных данных. С помощью мэппинга для источника данных из хранилища-источника его адаптер создаёт обменные файлы, а с помощью мэппинга для приёмника данных его адаптер помещает данные в хранилище-приёмник в «родном» для него формате;
- верификация результатов – на этой стадии проверяется полнота и непротиворечивость данных, сформированных в хранилище-приёмнике. Эта верификация может проводиться с помощью прикладных программ, предназначенных для работы с данным хранилищем, или же вручную.

Заключение

В данной статье рассмотрены базовые принципы интеграции данных в соответствии с нотацией стандарта ISO 15926. На сегодня разработка библиотек справочных данных различного уровня с применением стандарта ISO 15926 становится выполнимой задачей, даже с учетом постоянной доработки стандарта и методологий его использования.

Литература

1. Официальный сайт федерального агентства по техническому регулированию и метрологии [Электронный ресурс]. – Режим доступа: <http://www.norm-load.ru/SNiP/Data1/57/57931/index.htm>, свободный
2. Wikipedia [Электронный ресурс]. – Режим доступа: http://en.wikipedia.org/wiki/ISO_15926 свободный.
3. Сообщество ISO 15926 [Электронный ресурс]. – Режим доступа: <http://dot15926.livejournal.com/>, свободный.
4. Официальный сайт компании TechInvestLab.ru [Электронный ресурс]. – Режим доступа: <http://techinvestlab.ru/ISO15926> свободный