

Структура информационного контейнера третьего уровня зрелости и организация работы с нормативно-технической документацией в ходе жизненного цикла сложного инженерного объекта

В.Е. Мельников, М.Г. Жабицкий, В.П. Куприяновский, Р.Р. Джаппаров

Аннотация — Рассмотрены исторически сложившиеся подходы к созданию информационной модели на основе BIM-проектирования как основы для создания цифрового двойника сложного инженерного объекта. Приведены нормативные документы, формирующие современный уровень информационного моделирования в нашей стране. Рассмотрены методологические подходы к формированию информационных контейнеров различных уровней зрелости. Рассмотрено содержание информационного контейнера на каждом уровне зрелости – от «0» до «3» на примере энергоблока АЭС. Показана роль нормативной документации в процессе жизнедеятельности объекта на всех этапах жизненного цикла и её связь с объектом регулирования. Обозначены проблемы использования этой документации как «документоцентричной» и показан её неструктурированный информационный характер. Рассмотрены примеры подходов к управлению документацией и делается вывод о необходимости применения технологии Text Mining при формировании информационного контейнера нормативной документации в информационной модели сложного инженерного объекта (на примере энергоблока АЭС). Разработана концепция перехода модельного информационного контейнера на следующий уровень зрелости – создание информационного контейнера информационной модели третьего уровня зрелости. Рассмотрены варианты создания информационного контейнера уровня «3+» -автономная цифровая обработка базы данных механизмами Text Mining и экспертно-аналитическая система, в которой объединяется обработка баз данных экспертами с использованием механизмов Text-&-Data Mining. Приведён пример архитектуры системы обработки нормативной документации.

Ключевые слова— цифровой двойник, жизненный цикл, нормативно-техническая документация, цифровизация, интеллектуальный анализ текстов, Text Mining, информационный контейнер.

Статья получена 20 октября 2020.

Валерий Евгеньевич Мельников, Национальный исследовательский ядерный университет «МИФИ» (email: valeriy_melnikov_65@mail.ru). Михаил Георгиевич Жабицкий, Национальный исследовательский ядерный университет «МИФИ» (email: jabitsky@mail.ru). Куприяновский Василий Павлович, Российский университет транспорта (МИИТ) (email: v.kupriyanovsky@rut.digital).

Джаппаров Руфион Рахимович, Национальный исследовательский ядерный университет «МИФИ» (email: rufion@mail.ru).

I. ВВЕДЕНИЕ

Четвертая промышленная революция (4ПР) привела к широкому применению цифровых методов обработки и управления информацией. Одним из её результатов является широкое внедрение такого понятия как Цифровой двойник (ЦД), одной из основ которого является Информационная модель (ИМ). В первом десятилетии 21 века основные зарубежные игроки приняли у себя на национальном уровне ряд документов, определяющие принципы информационного моделирования при строительстве (BIM) [1-4].

Для успешного решения современных задач в строительстве были опубликованы первые две части международного стандарта ISO 19650: ISO 19650-1:2018 «Организация и оцифровка информации о зданиях и строительных работах, включая информационное моделирование зданий (BIM). Концепции и принципы» и ISO 19650-2:2018 «Организация и оцифровка информации о зданиях и строительных работах, включая информационное моделирование зданий (BIM). Фаза доставки активов» [5]. Стандарты были разработаны Техническим комитетом ISO / TC 59 «Здания и строительные работы» и подкомитетом SC 13 «Организация и оцифровка информации о зданиях и строительных работах, включая информационное моделирование зданий (BIM)».

В Российской Федерации аналогичная норма об информационной модели введена в Градостроительный кодекс Федеральным законом от 27.06.2019 N 151-ФЗ (статья 57.5) - Правила формирования и ведения информационной модели, состав сведений, документов и материалов, включаемых в информационную модель и представляемых в форме электронных документов, требования к форматам указанных электронных документов устанавливаются Правительством Российской Федерации.

При этом требования к нормам и подходам к процессам, инструментам и результатам информационного моделирования упреждающе прорабатывались в ряде документов, В частности:

-СП 331.1325800.2017 «Информационное моделирование в строительстве. Правила обмена между информационными моделями объектов и моделями, используемыми в программных комплексах»,

-СП 328.1325800.2017 «Правила описания компонентов информационной модели»,

-СП 333.1325800.2017 «Правила формирования информационной модели объектов на различных стадиях жизненного цикла»,

-СП 494.1325800.2018. «Правила разработки планов проектов, реализуемых с применением технологии информационного моделирования». Эти разработки закладывали основы понимания информационного моделирования на современном уровне в национальном масштабе, в ходе их создания, обсуждения и использования формируется широкое профессиональное сообщество.

Примером развития внесённой в Градостроительный кодекс нормы является принятый (и затем отменённый) ГОСТ Р 58349.1-2019 «ОРГАНИЗАЦИЯ ИНФОРМАЦИИ ОБ ОБЪЕКТАХ КАПИТАЛЬНОГО СТРОИТЕЛЬСТВА. Информационный менеджмент в строительстве с использованием технологии информационного моделирования. Часть 1», устанавливающий понятия информационного контейнера и уровней зрелости информационного менеджмента. В этом документе большой интерес представляет анализ и классификация понятия уровней зрелости информационного менеджмента [6,7].

отражением реально существующего объекта и виртуально выполнять все функции прототипа [8,9]. Но на поведение прототипа влияют не только внутренние, присущие ему свойства, обусловленные физической природой объекта, но и ограничения, накладываемые внешней «социальной» средой – правила, нормы, регламенты и т.д., принимающие форму Федеральных законов, регламентов и стандартов международных организаций, Технических регламентов РФ и СНГ, ГОСТов, СНИПов, СТП и так далее. Следовательно, и ЦД должен подчиняться этим требованиям. Это рассуждение позволяет сделать вывод о том, что информационная модель, как основа для реализации ЦД, должна содержать информационный контейнер, включающий в себя корпус нормативно-технической документации, определяющий рамки функционирования объекта во внешнем мире и соответствующего поведения его ЦД. Интересным представляется рассмотреть развитие этапов формирования информационного контейнера ИМ в существующих условиях:

Естественным представляется рассматривать развитие этой деятельности подробнее. Обсуждаемая ниже концепция проиллюстрирована Схемой 1.

В реальности практикуется широкий спектр уровней работы с нормативно-технической документацией:

- «0» уровень, аналоговый, бумажный – полностью неструктурированная информация. Это представляло из себя огромный набор текстовых и графических документов. Если рассматривать на примере энергоблока АЭС – то такое количество документов составит и составляет 2-3 железнодорожных вагона, т.е. их объём можно измерять в кубометрах. Из этого количества объём нормативной документации составит

Тип	Аналоговый		Цифровые		Ценность информационного взаимодействия
	Бумажный	2D/3D	Информационный контейнер на основе файлов	Информационный контейнер на основе базы данных	
Уровень	0	1	2	3	Уровни зрелости информационного менеджмента
Состав	ВМ				
	Документы	Набор образов документов и файлов	Файлы и базы данных	Связанные базы данных	
Слой информации	Неструктурированная информация	Структурированная информация	Сводные информационные модели	Информационные модели на базах данных	
					Среды общих данных на основе файлов/моделей/ информационных контейнеров
Слой технологий					
Слой стандартов			ГОСТ Р 58349.1-2019	Перспектива разработки стандартов ИСО и национальных стандартов	

Схема1 Спектр методологических подходов к работе с нормативно-технической документацией

II. МЕТОДОЛОГИЧЕСКИЕ ПОДХОДЫ К РАБОТЕ С НОРМАТИВНО-ТЕХНИЧЕСКОЙ ДОКУМЕНТАЦИЕЙ

Цифровой двойник должен являться полным

5-10 кубометров и фактически хранилась она в технических библиотеках, архивах, редко на столах специалистов (кстати, иногда покупалась за собственные деньги). При подготовке практически любого

документа, содержащего ссылки на первоисточники, их проверка могла занять не один день. Т.е. мы имеем полностью документоцентричную ситуацию, со всеми присущими ей недостатками – огромный объём, структурированность на уровне библиотечного каталога и полка в библиотеке, распределённость в физическом плане, неопределённость в статусе полученного документа (действует – не действует), низкая скорость внесения изменений и их отслеживаемость в зависимых документах, большое время на поиск необходимой информации в нужном документе и т.д.

- «1» уровень, цифровой. Характеризуется переводом документов в электронный вид – сканирование (2D). Ситуация изменилась – появились электронные каталоги, и в случае наличия доступа к электронному архиву, у вас появилась возможность получить скан необходимого документа, что не отменило необходимости проверки его статуса и поиска необходимой информации в нём. Можно отметить некие зачатки «датацентричности» - но только по форме. Нормативно-технические документы изменили форму представления, но не поменялся порядок работы с ними. Методы работы не изменились – анализ и выбор необходимой документации, поиск связей.

- «2» уровень, цифровой – информационный контейнер на основе файлов (сводные информационные модели), среда общих данных на основе файлов/моделей информационных контейнеров. Формирование наборов файлов – с течением времени и развитием информационных технологий формируются наборы файлов (кластеры) по стадиям жизненного цикла. Находятся в разных руках (проектанты, производители, эксплуатанты, регулирующие органы и т.д.), имеют разные, как правило, нестыкующиеся форматы – т.е. возникает проблема интероперабельности и т.д. Можно говорить о формировании информационных контейнеров у каждого отдельно взятого стейкхолдера в соответствии с принятыми внутренними нормами. Начинается разработка и создание на цифровом уровне систем управления документацией – АСУ ТД, электронных перечней, систем поиска информации, порталов знаний. Ситуация с нормативно-технической информацией не изменилась по сравнению с «1» уровнем – использование сканов (pdf – файлов).

- «3» уровень, цифровой – характеризуется как информационные контейнеры на основе баз данных, информационная модель на основе баз данных, среда общих данных на основе моделей/баз данных. Под информационным контейнером (information container) понимается именованный неизменяемый набор информации, извлекаемый из иерархии файлов, систем или программных приложений. Примером может служить подкаталог, файл с информацией (включая модель, документ, таблицу, календарный график) или некоторое подмножество такого файла, например, глава или раздел, слой или символ. Информационные контейнеры со структурированной информацией включают в себя цифровые информационные модели, спецификации, базы данных. Контейнеры с неструктурированной информацией включают в себя

текстовую и графическую документацию, видеозаписи, звукозаписи.

Приведённый анализ показывает, что ситуация с нормативной документацией за последние 20 лет практически не изменилась, кроме скорости доступа. Поэтому вопросы формирования информационного контейнера нормативно-технической документации 3-го уровня зрелости становятся всё более актуальными.

С юридической точки зрения нормативность означает обязательность требований, установленных документами (нормативными актами) для соответствующих субъектов права, круг которых является неопределённым. Вопрос о юридическом статусе нормативных документов по стандартизации имеет принципиальное значение, так как введение новых или отмена действующих документов, изменение их юридической силы связано со значительными материальными затратами и оказывает непосредственное влияние на работу предприятий, объем прав, обязанностей и юридической ответственности государственных органов и предприятий, эффективность защиты прав потребителей.

По своей сути НТД всегда будет являться текстовыми документами. Поэтому информационный контейнер нормативной документации ещё долгое время останется «документоцентричным» (как и массивы другой документированной информации – чертежи, съёмка, а также результаты технологий VR и AR в понимании «документа»), т.е. с точки зрения теории баз данных «неструктурированной» информацией. Текстовые документы практически невозможно преобразовать в табличное представление без потери семантики текста и отношений между сущностями. По этой причине такие документы хранятся в БД без преобразований, как текстовые поля (BLOB-поля). В то же время в тексте скрыто огромное количество информации, но ее неструктурированность не позволяет использовать алгоритмы Data Mining. Для решения этой проблемы предназначены методы анализа неструктурированного текста (Text Mining).

В качестве примера была рассмотрена АСУ ТД АО «Концерн Росэнергоатом», включающей в себя работу с эксплуатационной и управленческой документацией АС, проектной, учебно-методической, регламентной и методической документацией, рабочей, ремонтной, исполнительной и заводской документациями. Как видно из перечня, сюда не включены ряд блоков документов. Например, не представлена документация, регламентирующая процессы закупок, управления персоналом, экономическая и финансовая документация, документация по конфигурации и т.д., следовательно, говорить о полноценной информационной модели рано. При этом функции Text Mining задекларированы в функционале АСУ ТД, но находятся в «опытно-промышленной эксплуатации», т.е. в настоящее время практически не эксплуатируются.

По данным их открытых источников, аналогичная

ситуация наблюдается и в ряде других крупных корпораций [10-13] – АО «Камаз», АО «РЖД», «Газпром» и т.д.

Интересен опыт МАГАТЭ, выпустившего в 2019 году стандарт SPESS E, версия 3.0 «Система управления знаниями и контентом, механизм обратной связи, ИТ-платформа NSS-OUI для процесса проверки / пересмотра и публикации для Серии стандартов безопасности, Серии изданий по физической ядерной безопасности и других публикаций NS», описывающий систему управления контентом для характеристики содержания публикаций путем определения того, какие тематические области затрагиваются в этих публикациях, а также к разработке системы управления знаниями, основанной на том же наборе тематических областей, для характеристики обратной связи, таким образом устанавливая связь между отзывами и содержанием публикаций [14].

Авторы считают, что подходы, обозначенные в этом документе – создание объектно-ориентированной онтологии (на примере блока АЭС), применение глоссариев, явные и неявные отношения, общий подход к тегам, концепции и инструменты для последовательного и сопутствующего пересмотра набора публикаций с применением алгоритмов семантического анализа Text Mining могут внести большой вклад в создание информационного контейнера нормативной документации.

III. ПОДХОДЫ К ПРОГРАММНОЙ РЕАЛИЗАЦИИ ИНФОРМАЦИОННОГО КОНТЕЙНЕРА ТРЕТЬЕГО УРОВНЯ

Таким образом, задача построения информационного контейнера третьего уровня, в том числе, и для нормативной документации является актуальной задачей именно сегодня. Авторами разработана концепция создания модельного информационного контейнера на следующий уровень зрелости – создание ИК ИМ уровня 3+. Для этого в режиме прототипирования апробируется эксплуатация БД документации в двух вариантах. Первый (целевой) подход – автономная цифровая обработка базы данных механизмами Text Mining с элементами интеллектуального анализа текстов. Вторым, более традиционным, является вариант экспертно-аналитической системы, в которой объединяется обработка баз данных экспертами с использованием механизмов Text-&Data Mining для предобработки информационных массивов. На сегодняшнем уровне проработки второй подход дает более адекватные результаты. Однако он требователен к квалификации экспертов, имеет невысокую скорость обработки и, следовательно, высокую стоимость эксплуатации. По ходу работ в технологии первого подхода вносятся корректировки на основании опыта работы экспертов. Такой подход апробируется на модельном, относительно малом массиве документов в настоящее время. Специалистами отраслевых организаций ГК «Росатом» проявлена активная заинтересованность к поставленной задаче.

Архитектура системы является достаточно традиционной и представлена на рисунке 1, где (1) – АРМ Оператора; (2) – АРМ и КПК Пользователей; (3) – Front-end Сервер (WEB+API); (4) – Станция загрузки сканов документов; (5) – Сервер распознавания документации; (6) – Back-end Сервер; (7) – Репликация и контроль версий БД; (8) – База данных (СХД).

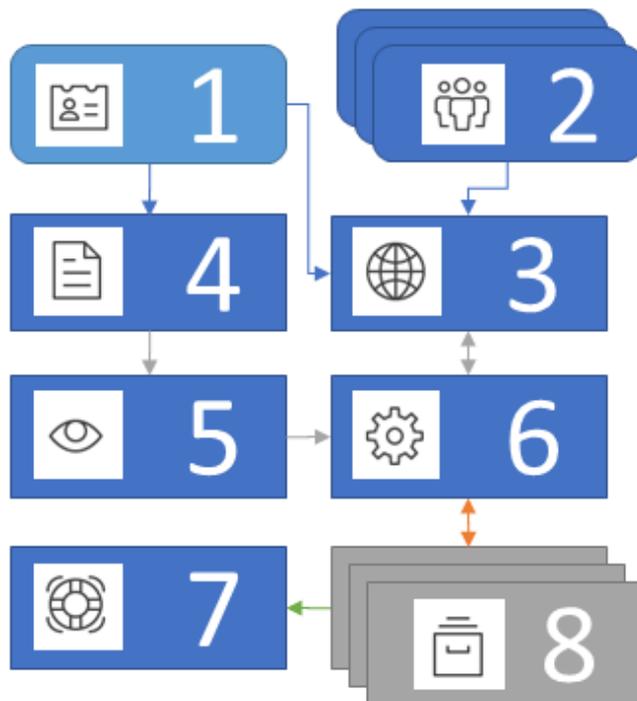


Рис 1. Архитектура системы

Краткая аннотация по функционалу узлов в приведённой схеме:

АРМ Оператора:

Операторы будут оцифровывать, проверять и вносить правки в документы архива;

АРМ Операторов будут находиться на территории архива предприятия.

АРМ и КПК пользователей:

Данные будут передаваться по зашифрованному каналу связи с применением алгоритмов, прописанных в ГОСТ Р 34.10-2012 с длиной ключа 256 или 512 бит;

Для КПК планируется разработка специального программного обеспечения (приложения) с поддержкой авторизации методами, прописываемыми в ТЗ (биометрия, смарт-карта, сертификат ЭЦП).

Back-end Сервер:

Основная нагрузка по обработке данных будет полагаться на данный узел (индексирование документации, аудит безопасности, внесение правок в БД и т.д.);

По ходу разработки платформы данный узел будет разделён на функциональные блоки – при полномасштабной эксплуатации потребуется создание кластера серверов для распределения нагрузки и обеспечения отказоустойчивости.

Узел репликации и контроля версий БД:

Будет распределён географически и логически, основной приоритет – отказоустойчивость и быстрое

восстановление данных в случае аварии на основной БД. Сохранённые копии БД будут зашифрованы с применением алгоритмов, прописанных в ГОСТ Р 34.10-2012 с длиной ключа 512 бит.

Вся сеть платформы при необходимости может быть полностью закрыта от доступа к сети Интернет, детали степени защищённости будут проработаны в ТЗ.

Вся платформа (кроме узлов репликации БД) будет находиться в ЦОД предприятия-заказчика.

Однако, более важным и интересным является логическая архитектура обработки документов. Принципиальный подход к ней представлен на рисунке 2, где (1) – файл отсканированного документа; (2) – опознавание текста (OCR); (3) – парсеры ввода данных; (4) – разбор документа: дата, ведомство, ссылки; (5) – обработка согласно матрице терминов, физических мер и формул; (6) – отработка соответствий зависимостям и иерархии ведомств, понятий в других документах; (7) – база данных; (8) – поиск конфликтов, несоответствий, ошибок и сомнительных элементов; (9) – отслеживание актуальности документов по дате и иным условиям; (10) – составление аннотаций к документам; (11) – парсеры обработки БД и вывода данных; (12) – Front-end (пользовательский интерфейс).

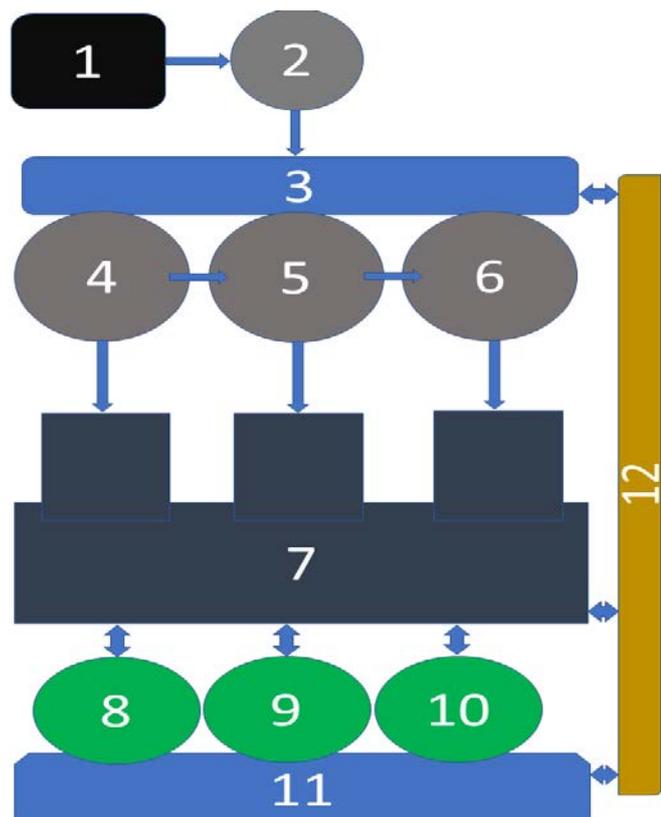


Рис. 2. Логическая архитектура обработки документов

Из схемы видно, что в системе есть два блока обработки массивов текстов. Первый из них, парсеры ввода данных, являются традиционными, и на их разборе мы останавливаться не будем. По сути, их функционал – перевод массива внутренне неструктурированных или слабо структурированных текстов любого формата в базу распознанной текстовой информации с

классификацией по формальным признакам. Это – предобработка, по сути включающая закрытый конечный перечень операций, хотя и допускающая определенные перенастройки.

Гораздо более интересными и нетривиальными являются парсеры обработки БД и вывода. Этот блок является открытым и расширяемым, в том числе модулями внешних разработчиков. На рисунке приведены лишь стартовый набор парсеров, они расширяются и дополняются в систему в зависимости от востребованности. Целью модулей обработки базы данных является структурирование и смысловой анализ корпуса текстов в выбранной области, например, отраслевой нормативно-справочной документации. Базовыми технологиями являются методы цифрового анализа текстов и онтологический подход к базе знаний. На этом уровне задачи имеют дифференциацию, как по сложности, так и по востребованности. Например, в составе типовых задач:

- построение типового дерева связей корпуса нормативной документации и выявление несоответствий (неактуальных документов и версий документов);
- построение системы терминологии, тезаурусов в предметной области и выявление противоречий и несоответствий в них;
- кластеризация корпуса документов, выявление неявных классификаций и связей;
- поиск схожих документов и смысловых связей между документами на базе анализа осмысленных тематических фрагментов;
- извлечение и подбор комплексных полноценных разносторонних выборок по проблематике тематических фрагментов;
- автоматическое аннотирование текстов;
- а также другие задачи работы со знаниями и смыслами в наборах текстов.

IV. ЗАКЛЮЧЕНИЕ. ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ ПРОДУКТА

В перспективе, целевым результатом работы является повышение производительности и качества работы предметных специалистов стартового и среднего уровня при многофакторном анализе новых для них проблем, а также обеспечение актуальности и обоснованности их выводов на основании работой с широким корпусом предметной нормативно-технической документации.

При решении рассматриваемых задач используются методы анализа текстов (Text mining), включая семантический и лингвистический анализ, статистическую обработку, онтологические подходы к формированию смыслов в предметной области.

Прототипирование и отработка методик проводится на массиве текстов рекомендаций МАГАТЭ и открытой отраслевой нормативно-технической документации атомной отрасли РФ. Перспективной задачей является включение в сферу анализа разноязычных текстов (русских и английских). Это связано с тем, что заметная часть рекомендаций и отчетов МАГАТЭ не имеет

официальных переводов на русский язык.

БИБЛИОГРАФИЯ

- [1] Volk, R., Stengel, J., Schultmann, F. (2014). Erratum: Building Information Modeling (BIM) for existing buildings - Literature review and future needs (Automation in Construction (2014) 38 (109-127)). 43. 10.1016/j.autcon.2014.02.010.
- [2] ERDC SR-12-2. The US Army Corps of Engineers Roadmap for Life-Cycle Building Information Modeling (BIM). 2012. 45 p
- [3] Скворцов А.В. Обзор международной нормативной базы в сфере BIM // САПР и ГИС автомобильных дорог. 2016. № 2(7). С. 4–48. DOI: 10.17273/CADGIS.2016.2.1.
- [4] ERDC SR-12-2. The US Army Corps of Engineers Roadmap for Life-Cycle Building Information Modeling (BIM). 2012. 45 p
- [5] ISO 19650-1:2018 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles
- [6] ГОСТ Р 58349.1-2019 «Организация информации об объектах капитального строительства. Информационный менеджмент в строительстве с использованием технологии информационного моделирования. Часть 1»
- [7] Кузин Д.В. Проблемы цифровой зрелости в современном бизнесе. Мир новой экономики. 2019;13(3):89-99. DOI: 10.26794/2220-6469-2019-13-3-89-99
- [8] Королев А.С., Александров В.С. Инструментальная поддержка процесса управления жизненным циклом сложных технических систем // Системы управления и информационные технологии. Москва-Воронеж: Научная книга. 2013. № 2.1 (52). С. 137-144.
- [9] Романовская В.Е. Цифровой двойник как основа цифрового проектирования и моделирования / Цифровые технологии в экономике и промышленности. Сборник национальной научно-практической конференции с международным участием 22-23 ноября 2019 года- СПб.:ПОЛИТЕХ- ПРЕСС, 2019.- С. 207-213.
- [10] Цыганов В.В., Савушкин С.А. Методика каталогизации, стандартизации и контроля услуг холдинга ОАО «РЖД»//Транспорт России: проблемы и перспективы - 2016. Материалы международной научно-практической конференции. 29-30 ноября 2016 г. СПб. ИПТ РАН. Санкт-Петербург. 2016 . Том 1. 400 с. С.102–106.
- [11] Годовой отчет ПАО "Камаз" 2019 [электронный ресурс]: Официальный сайт ПАО «Камаз». – Режим доступа <https://kamaz.ru/upload/iblock/0b1/0b1d45d1e678f73bdfc32e85b5997022.pdf> (дата обращения 21.11.2020)
- [12] Техническое регулирование [электронный ресурс]: Официальный сайт ПАО «Газпром». – Режим доступа <https://www.gazprom.ru/about/strategy/innovation/tech-regulation> (дата обращения 21.11.2020)
- [13] Распоряжение ОАО "РЖД" от 25.01.2018 N 120/р "Об утверждении СТО РЖД 05.010-2018 "Объекты железнодорожной инфраструктуры и подвижной состав. Система управления нормативной и технической документацией в ОАО "РЖД" [электронный ресурс]: КонсультантПлюс – Режим доступа <http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=714010#08044271986086144> (дата обращения 21.11.2020)
- [14] Knowledge and Content Management System, Feedback mechanism, NSS-OUI IT Platform for the Review/Revision and Publication Process for the Safety Standards Series, the Nuclear Security Series and other NS publications [электронный ресурс]: International Atomic Energy Agency. – Режим доступа <https://www-ns.iaea.org/committees/files/CSScomments/1708/SPESE27February2019cleanwithontology.pdf> (дата обращения 21.11.2020)

Structure of the third-level experience information container and technical-regulatory documentation work organization during the life cycle of a complicated engineering object

Valerii Melnikov, Mikhail Zhabitskii, Vasily Kupriyanovsky, Rufion Djaparov

Abstract—Historically developed approaches to creating an information model based on BIM-design as a basis for creating a digital twin of a complex engineering object are considered. The normative documents forming the modern level of information modeling in our country are given. Methodological approaches to the formation of information containers of various levels of maturity are considered. The content of the information container at each maturity level – from "0" to "3" is considered on the example of a nuclear power plant unit. The role of regulatory documentation in the process of life of the object at all stages of the life cycle and its relationship with the object of regulation is shown. Marked problem use this documentation as "documentational" and shows its unstructured information purposes only. Examples of approaches to document management are considered and a conclusion is made about the need to use Text Mining technology in the formation of an information container of regulatory documentation in the information model of a complex engineering object (for example, a nuclear power plant unit). The concept of transition of the model information container to the next level of maturity is developed – creation of the information container of the information model of the third level of maturity. Options for creating an information container of the "3+" level are considered: Autonomous digital database processing by Text Mining mechanisms and an expert-analytical system that combines database processing by experts using Text - & - Data Mining mechanisms. An example of the architecture of the regulatory documentation processing system is given.

Keywords – digital twin, life cycle, regulatory and technical documentation, digitalization, text mining, information container

REFERENCES

- [1] Volk, R., Stengel, J., Schultmann, F. (2014). Erratum: Building Information Modeling (BIM) for existing buildings - Literature review and future needs (Automation in Construction (2014) 38 (109-127)). 43. 10.1016/j.autcon.2014.02.010.
- [2] ERDC SR-12-2. The US Army Corps of Engineers Roadmap for Life-Cycle Building Information Modeling (BIM). 2012. 45 p
- [3] Skvorcov A.V. Obzor mezhdunarodnoj normativnoj bazy v sfere BIM // SAPR i GIS avtomobil'nyh dorog. 2016. # 2(7). S. 4–48. DOI: 10.17273/CADGIS.2016.2.1.
- [4] ERDC SR-12-2. The US Army Corps of Engineers Roadmap for Life-Cycle Building Information Modeling (BIM). 2012. 45 p
- [5] ISO 19650-1:2018 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles
- [6] GOST R 58349.1-2019 «Organizacija informacii ob ob'ektah kapital'nogo stroitel'stva. Informacionnyj menedzhment v stroitel'stve s ispol'zovaniem tehnologii informacionnogo modelirovanija. Chast' 1»
- [7] Kuzin D.V. Problemy cifrovoj zrelosti v sovremennom biznese. Mir novoj jekonomiki. 2019;13(3):89-99. DOI: 10.26794/2220-6469-2019-13-3-89-99
- [8] Korolev A.S., Aleksandrov V.S. Instrumental'naja podderzhka processa upravlenija zhiznennym ciklom slozhnyh tehniceskikh sistem// Sistemy upravlenija i informacionnye tehnologii. Moskva-Voronezh: Nauchnaja kniga. 2013. # 2.1 (52). S. 137-144.
- [9] Romanovskaja V.E. Cifrovoj dvojniki kak osnova cifrovogo proektirovanija i modelirovanija / Cifrovyje tehnologii v jekonomike i promyshlennosti. Sbornik nacional'noj nauchno-prakticheskoj konferencii s mezhdunarodnym uchastiem 22-23 nojabrja 2019 goda- SPb.:POLITEH- PRESS, 2019.- S. 207-213.
- [10] Cyganov V.V., Savushkin S.A. Metodika katalogizacii, standartizacii i kontrolja uslug holdinga OAO «RZhD»//Transport Rossii: problemy i perspektivy - 2016. Materialy mezhdunarodnoj nauchno-prakticheskoj konferencii. 29-30 nojabrja 2016 g. SPb. IPT RAN. Sankt-Peterburg. 2016 . Tom 1. 400 s. S.102–106.
- [11] Godovoj otchet PAO "Kamaz" 2019 [jelektronnyj resurs]: Oficial'nyj sajt PAO «Kamaz». – Rezhim dostupa <https://kamaz.ru/upload/iblock/0b1/0b1d45d1e678f73bdfc32e85b5997022.pdf> (data obrashhenija 21.11.2020)
- [12] Tehniceskoe regulirovanie [jelektronnyj resurs]: Oficial'nyj sajt PAO «Gazprom». – Rezhim dostupa <https://www.gazprom.ru/about/strategy/innovation/tech-regulation> (data obrashhenija 21.11.2020)
- [13] Rasporjazhenie OAO "RZhD" ot 25.01.2018 N 120/r "Ob utverzhdenii STO RZhD 05.010-2018 "Ob"ekty zheleznodorozhnoj infrastruktury i podvizhnoj sostav. Sistema upravlenija normativnoj i tehniceskaj dokumentaciej v OAO "RZhD" [jelektronnyj resurs]: Konsul'tantPljus – Rezhim dostupa <http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=714010#08044271986086144> (data obrashhenija 21.11.2020)
- [14] Knowledge and Content Management System, Feedback mechanism, NSS-OUI IT Platform for the Review/Revision and Publication Process for the Safety Standards Series, the Nuclear Security Series and other NS publications [jelektronnyj resurs]: International Atomic Energy Agency. – Rezhim dostupa <https://www-ns.iaea.org/committees/files/CSScomments/1708/SPESSE27February2019cleanwithontology.pdf> (data obrashhenija 21.11.2020)