

Методы построения и анализа онтологической модели и эталонной базы данных цифрового фонда интеллектуальной собственности

В. О. Сиротюк

Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова РАН, г. Москва, Россия

Аннотация. В статье рассмотрены цели, задачи и принципы цифровой трансформации традиционных систем управления интеллектуальной собственностью (ИС), требования к цифровому информационному фонду интеллектуальной собственности (ЦИФИС), играющему основополагающую роль при построении цифрового органа управления системы управления ИС, особенности построения онтологической модели предметной области системы управления ИС. Предложены методы построения и анализа онтологической модели и концептуальной структуры предметной области ЦИФИС. Разработаны методы построения и анализа канонических структур эталонных (типовых) баз данных (ЭБД) патентной и научно-технической информации ЦИФИС. Для построения онтологической модели и канонической структуры ЭБД используются объектно-ориентированные методы анализа и проектирования.

Ключевые слова: система управления интеллектуальной собственностью, цифровой информационный фонд интеллектуальной собственности, базы данных патентной информации, базы данных научно-технической информации, объекты данных, онтологическая модель предметной области, эталонная база данных, стандарты ВОИС.

DOI 10.14357/20718632220306

Введение

Становление цифровой экономики является одним из приоритетных направлений научно-технического прогресса большинства развитых стран мира. В России переход к цифровой экономике предусмотрен Стратегией развития информационного общества в России на 2017-2030 годы, утвержденной Указом Президента РФ.

Системы управления интеллектуальной собственностью (ИС) играют важную роль при проведении хозяйствующими субъектами патентных и научных исследований с целью принятия обоснованных решений в различных областях науки и техники, определения целесо-

образности проведения собственных разработок техники и технологий или закупки лицензий, патентования результатов интеллектуальной деятельности или о путях обхода патентов конкурирующих фирм и т.д.

Цифровая трансформация системы управления ИС позволяет перейти на новые бизнес-модели и методы управления ИС и должна осуществляться на принципах клиентоориентированности и омниканальности (интеграции разрозненных каналов коммуникации с пользователями в единую систему), максимизации эффективности обслуживания запросов пользователей.

Создание цифрового органа управления ИС и экосистемы в патентно-информационной деятельности позволяет повысить эффективность и качество работы патентных и научно-исследовательских организаций и оказываемых ими патентно-информационных и технологических услуг. Новые бизнес-модели должны отвечать требованиям пользователей и охватывать жизненный цикл производства и оказания услуг от формирования бизнес-предложения, направленного на решение задач пользователя, до обслуживания его запросов и оказания сопутствующих услуг.

Необходимым условием трансформации традиционной системы управления ИС и перехода к цифровому органу управления ИС является наличие и доступность информационных фондов патентной и научно-технической документации в цифровом виде, в совокупности образующих цифровой информационный фонд интеллектуальной собственности (ЦИФИС) [1].

Большие объемы разнородной информации, хранимой в базах данных (БД) ЦИФИС, обуславливают необходимость разработки методов и средств оптимизации их проектирования, сопровождения и развития, разработки информационных технологий сбора, оцифровки, обработки, хранения и использования данных, защиты данных, информационной и обеспечивающей инфраструктуры ЦИФИС, обеспечения их сохранности и восстановления в случае сбоев.

Построение эффективной экосистемы в патентно-информационной деятельности предъявляет высокие требования к качеству информации БД ЦИФИС (полноте, достоверности, доступности, своевременности и др.) со стороны как внутренних (экспертов), так и внешних пользователей, проводящих патентные и научные исследования [1-3]. Инфраструктура современного ЦИФИС должна обеспечивать доступ к локальным и внешним БД патентной (ПБД) и научно-технической (БД НТИ) информации и их совместное использование.

От эффективности используемых методов и средств проектирования, создания и эксплуатации ПБД и БД НТИ ЦИФИС, обеспечения полноты, корректности, достоверности, доступности

и защиты содержащейся в них информации во многом зависит эффективность и качество функционирования цифровых систем управления ИС и предоставляемых ими услуг.

В этой связи проектирование эффективных ПБД и БД НТИ ЦИФИС должно осуществляться на основе анализа общих системных требований, предъявляемых стандартами ВОИС, международными и национальными стандартами, требованиями и рекомендациями цифровых библиотек интеллектуальной собственности (ЦБИС) к составу, структуре и качеству патентной и научно-технической документации, функциональным и сервисным возможностям информационных систем для проведения эффективных патентно-информационных поисков [1]. Данные требования являются общими для разных тематик (областей знаний), по которым проводятся поиски, поэтому эффективные БД ЦИФИС должны строиться на основе онтологической модели и единой концептуальной структуры предметной области (ПрО) ЦИФИС.

Отсутствие формальных методов и алгоритмов построения и анализа онтологической модели предметной области научных и патентных исследований и концептуальной структуры ПрО ЦИФИС не позволяет построить оптимальные структуры ПБД/БД НТИ на всех основных этапах их проектирования (каноническом, логическом, физическом), а также эффективные эталонные БД (ЭБД), создаваемые для оценки показателей качества ПБД и БД НТИ ЦИФИС (полноты, достоверности, непротиворечивости, доступности и др.).

Настоящая работа направлена на решение данных проблем. В работе предложены формализованные методы и алгоритмы построения онтологической модели и концептуальной структуры предметной области ЦИФИС, формирования канонической структуры ЭБД и ее анализа. Для формализации используются модели и методы объектно-ориентированного анализа и проектирования, так как они наиболее адекватно и полно отражают технологию формирования и хранения БД ЦИФИС и предоставляемых ими услуг.

1. Особенности построения онтологической модели предметной области системы управления ИС

Использование онтологий для моделирования предметных областей пользователей информационно-управляющих систем получает в последнее время все более широкое распространение [4]. С помощью онтологий предпринимается попытка всеобъемлющей и подробной формализации некоторой области знаний с помощью концептуальной схемы. Обычно такая схема состоит из структуры данных, содержащей все релевантные классы объектов, связи между ними, правила и ограничения, принятые в этой области. Онтологические модели представляют собой наиболее развитый сегодня вариант семантического описания предметной области, формализующий знания о ней. Современные онтологии строятся по большей части одинаково, независимо от языка описания, и состоят из объектов, понятий (классов объектов), атрибутов и отношений [5]. В общем случае они включают словарь терминов предметной области и множество их логических взаимосвязей. В области интеллектуальной собственности в качестве словаря-справочника основных понятий, терминов и определений может использоваться, например, издание Совета Федерации РФ [6]. В таком качестве онтологии претендуют на роль универсальной модели представления знаний для различных предметных областей. В целом онтологии выполняют функцию интеграции, обеспечивая общий семантический базис в процессах принятия решений и интеллектуального анализа данных, а также единую платформу для объединения разнообразных информационных систем.

Онтологическая модель предметной области (ПрО) системы управления ИС строится на основе анализа общих системных требований, предъявляемых стандартами ВОИС к бизнес-процессам и информационным технологиям управления ИС, составу и структуре патентной и научно-технической документации ЦИФИС и рекомендациями ЦБИС к инструментально-программным средствам поиска, доступа к данным и обработки, сервисным средствам, требуемым для проведения эффективных па-

тентных поисков международного типа и выполнения бизнес-процессов. Помимо этих требований при моделировании предметной области научных исследований могут использоваться положения стандарта CERIF (Common European Research Information Format - общий европейский формат для исследовательской информации) [7]. Данный стандарт поддерживает концепцию объектов (сущностей) научных данных с атрибутами, отношения типа «многие-ко-многим» между объектами. Основными декларируемыми объектами данных в стандарте CERIF являются Person, OrganisationUnit и Project (Человек, Организация и Проект), каждый из которых рекурсивно связан сам с собой и поддерживает отношения с другими объектами. Стандарт описывает также множество специальных и дополнительных объектов, с помощью которых полностью описываются научно-исследовательские проекты, их участники, результаты их совместной работы и другие результаты, например, такие объекты как ResultPublication (Публикация), ResultPatent, (Патент), ResultProduct (Продукт), FundProg (программа финансирования), Event (событие), Prize (вознаграждение), Equip (оборудование) и др. Отношения между объектами описываются в стандарте с помощью связей, которые строятся путем наследования названий и идентификаторов объектов-родителей и дополнительно наделяются атрибутами даты начала и конца действия связи. Таким образом, все возможные взаимоотношения между объектами ПрО научных исследований (проектами, людьми, организациями, продуктами, публикациями и т.д.) формально можно представить в виде графа концептуальной структуры ПрО, вершинами которого является множество объектов данных, а дугами - отношения между ними. Семантика объектов и отношений между ними задается в таблицах, описывающих возможные роли и взаимодействия между объектами, необходимые для описания бизнес-процессов предметной области системы управления ИС (свойства объектов, отношения, ограничения, аксиомы и утверждения об объектах).

Отмеченные требования стандартов и рекомендации являются общими для разных областей знаний (тематик) ПрО системы управления ИС. Поэтому формируемая в дальнейшем на основе построенной онтологической модели

каноническая структура эталонной БД (ЭБД) является типовой и может использоваться при оценке показателей качества ПБД и БД НТИ разных тематик [2, 3].

2. Методы построения онтологической модели и концептуальной структуры предметной области ЦИФИС

В общем случае, онтологическая модель ПрО системы управления ИС формируется в результате объединения онтологических моделей отдельных объектов цифровизации и бизнес-процессов (подобластей ПрО), к которым относятся: формирование, сопровождение и развитие ЦИФИС; патентно-информационный поиск и экспертиза; выдача патентов и авторских свидетельств; публикация патентной и научно-технической информации; проведение научных исследований; внедрение инновационных разработок; коммерциализация результатов научных исследований и др. [1, 8].

В работе рассматривается процесс построения онтологической модели, охватывающей и описывающей объекты и бизнес-процессы формирования, загрузки, сопровождения и развития БД ЦИФИС. Построение онтологических моделей других объектов цифровизации и бизнес-процессов системы управления ИС выполняется аналогично рассмотренным ниже с учетом специфики их описания и особенностей реализации.

Пусть $O = \{o_\varepsilon / \varepsilon = \overline{1, \varepsilon_0}\}$ - множество объектов данных ПрО ЦИФИС. В качестве объектов данных ЦИФИС выступают произведения науки, литературы и искусства, программы для ЭВМ, изобретения, исполнения, товарные знаки и т.д.;

$T = \{t_n / n = \overline{1, N}\}$ - тезаурус понятий (терминов) ПрО ЦИФИС. Понятиями или классами объектов являются абстрактные группы, коллекции или наборы объектов. Понятия могут включать в себя помимо объектов другие классы либо их сочетания. В качестве понятий ПрО ЦИФИС используются такие как заявка, патент, публикация, автор, патентообладатель, авторское право и т.п. [6];

$H = \{h_j / j = \overline{1, J}\}$ - множество методов (процедур) обработки данных, описывающих поведенческие аспекты объектов данных и технологические отношения между ними. Процедурами обработки ЦИФИС являются электронная подача заявок на изобретения, подача материалов научного произведения, переписка с автором, проверка статьи на плагиат, публикация статьи, публикация патента и др.;

$E = \{e_p / p = \overline{1, P}\}$ - множество атрибутов (информационных элементов) предметной области ЦИФИС, описывающих свойства объектов данных. К ним относятся имя автора, наименование произведения, индексы МПК для патентной и УДК для научно-технической информации и др. [1];

$R = \{r_y / y = \overline{1, Y}\}$ - множество отношений и взаимосвязей между объектами, атрибутами и процедурами обработки данных.

Формализовано онтологическая модель предметной области ЦИФИС описывается с помощью множеств $\{O, T, H, E\}$ и отношений (взаимосвязей) между ними. Выделим пять типов отношений, характеризующих структурные, семантические и технологические свойства и характеристики между элементами ПрО ЦИФИС: $\{OR_1T\}, \{ER_2O\}, \{HR_3O\}, \{OR_4O\}, \{TR_5T\}$, где R_1 - отношение принадлежности объекта понятию (описание понимания объекта данных, семантический аспект), R_2 - отношение описания свойств объектов (структурный аспект), R_3 - отношение описания поведенческих свойств объектов (технологический аспект), R_4 - описание взаимосвязей между объектами данных (информационных и технологических), R_5 - описание взаимосвязей между понятиями ПрО ЦИФИС (семантический аспект).

На основании формализованного описания онтологической модели осуществляется построение концептуальной структуры ПрО ЦИФИС, представляемой в виде мультиграфа $G(D, U)$ с одним типом вершин $D = \{d_l / l = \overline{1, L}\}$, соответствующих множеству структурных элементов ПрО (объектам, понятиям и атрибутам, $D = O \cup T \cup E$) и двумя типами дуг: $U^{jn} = \{(d_l, d_r)\}$ - множество инфор-

мационных взаимосвязей (семантических отношений) между понятиями, объектами данных и атрибутами и $U^{np} = \{(d_r, d_p)\}$ - множество технологических (процедурных) взаимосвязей между понятиями, объектами и атрибутами, $U = U^{nl} \cup U^{np}$. Построение графа $G(D, U)$ осуществляется путем последовательного соединения структурных элементов $d_l, d_{l'} \in D$ дугой при наличии отношения (взаимосвязи) между ними, зафиксированного в множествах $\{R_1, \dots, R_5\}$ онтологической модели ПрО. Граф концептуальной структуры ПрО ЦИФИС описывается матрицей смежности $B = \|b_{ll'}\|$, элемент которой $b_{ll'} = 1$, если между элементами d_l и $d_{l'}$ существует взаимосвязь (информационная или технологическая), $b_{ll'} = 0$ - в противном случае.

3. Модели и методы построения канонической структуры эталонной БД ЦИФИС

Эталонные БД (ЭБД) используются для определения степени общности структур ПБД и БД НТИ и соответствующих ЭБД и расчета численных значений показателей качества (полноты, достоверности, доступности и др.) БД ЦИФИС. Построение ЭБД осуществляется на основе онтологической модели и концептуальной структуры ПрО ЦИФИС.

Рассмотрим формализованные модели и методы построения канонической структуры ЭБД.

Проектирование канонической структуры ЭБД осуществляется в несколько этапов. На первом этапе осуществляется нормализация графа концептуальной структуры ПрО ЦИФИС $G(D, U)$, заключающаяся в выполнении следующих процедур анализа графа $G(D, U)$ и его преобразования: выявлении избыточных взаимосвязей и дублируемых элементов и процедур и их устранение; корректировки составов объектов данных; выделении идентификаторов объектов (ключевых элементов); выявлении классов объектов; упорядочении объектов данных по уровням иерархии, и других процедур, описанных в [9]. В

результате их выполнения формируется нормализованная объектная каноническая структура ЭБД, которая представляется в виде графа $G_{kc}^{ob}(O, \Delta)$, вершинами которого $O = \{o_\varepsilon / \varepsilon = \overline{1, \varepsilon_0}\}$ являются объекты данных, а дугами $\Delta = \{\delta_{\varepsilon\varepsilon'} / \varepsilon, \varepsilon' = \overline{1, \varepsilon_0}\}$ - связи (отношения) между объектами. Каждый объект (вершина графа) описывается атрибутами из оптимизированного в результате выполнения процедур нормализации множества $D_o = \{d_l / l = \overline{1, L_{opt}}\}$. Граф $G_{kc}^{ob}(O, \Delta)$ формализовано описывается матрицей смежности $B_{kc}^{ob} = \|b_{\varepsilon\varepsilon'}\|$ между объектами. Свойства объектов (вершины графа) описываются матрицей смежности $V = \|v_{l\varepsilon}\|$, элемент которой $b_{l\varepsilon} = 1$, если атрибут $d_l \in D_o$ входит в состав объекта O_ε , $b_{l\varepsilon} = 0$ - в противном случае. Поведение объектов описывается технологической матрицей смежности $W = \|w_{j\varepsilon}\|$, элемент которой $w_{j\varepsilon} = 1$, если элемент $o_\varepsilon \in O$ (объект данных, атрибут объекта) используется при выполнении процедуры h_j , $w_{j\varepsilon} = 0$, в противном случае, $j = \overline{1, J_{opt}}$.

Характеристиками графа G_{kc}^{ob} являются количественные показатели объектов данных $\{z_\varepsilon^{nl}, z_\varepsilon^{np}\}$, используемые в дальнейшем при количественной оценке показателей качества БД ЦИФИС, где z_ε^{nl} - информационный вес ε -го объекта, z_ε^{np} - технологический вес ε -го объекта. Информационный вес объекта данных O_ε характеризует количественный и смысловой состав атрибутов объекта, необходимый и достаточный для описания его свойств в соответствии с требованиями стандартов ВОИС, международных и национальных стандартов и рекомендаций ЦБИС. Показатель z_ε^{nl} определяется на основании матрицы

$$V = \|v_{l\varepsilon}\|: z_\varepsilon^{nl} = \sum_{l=1}^{L_{opt}} v_{l\varepsilon}. \text{ Технологический вес}$$

объекта характеризует функциональный состав объекта, т.е. перечень процедур и операций

поиска, обработки, передачи и представления данных, достаточных для выполнения объектом O_ε возложенных на него функций и задач управления в соответствии с требованиями международных и национальных стандартов и рекомендаций ЦБИС по проведению патентно-информационных поисков международного типа. Технологический вес z_ε^{np} объекта O_ε определяется на основании матрицы

$$W = \|w_{j\varepsilon}\| \text{ следующим образом: } z_\varepsilon^{np} = \sum_{j=1}^{J_{opt}} w_{j\varepsilon}.$$

На заключительном этапе анализа канонической структуры ЭБД решается задача декомпозиции графа $G_{kc}^{ob}(O, \Delta)$ на ряд подграфов, которые соответствуют суперклассам, классам-наследникам и независимым классам. При этом осуществляется модификация матрицы смежности B_{kc}^{ob} путем добавления информации о выделенных классах и отношениях между ними. Декомпозиция позволяет унифицировать, систематизировать и типизировать структурные элементы ПрО ЦИФИС и тем самым сократить размерность исходного графа канонической структуры ЭБД.

Рассмотрим алгоритм выделения суперклассов и классов-наследников. Исходными данными для алгоритма являются:

- матрица типов данных атрибутов $S = \|s_{ij}\|$, проиндексированная по осям полным множеством атрибутов $d_i \in D_o$ и множеством типов данных $j = \overline{1..J}$. Значение элемента $s_{ij} = 1$, если на основании информации пользователей информационный элемент d_i относится к j -му типу данных, в противном случае $s_{ij} = 0$;

- матрица входимости (принадлежности) объектов данных понятиям, формируемой на основании отношения R_I онтологической модели ПрО ЦИФИС: $Q = \|q_{\varepsilon n}\|$, элемент которой $q_{\varepsilon n} = 1$, если объект данных O_ε описывается понятием t_n , $q_{\varepsilon n} = 0$ - в противном случае;

- матрица смежности $B_{kc}^{ob} = \|b_{\varepsilon\varepsilon'}\|$ между объектами;

- информационно-функциональные составы объектов:

$$H(O_\varepsilon) = \{D^\varepsilon, H^\varepsilon\},$$

$$D^\varepsilon = \{d_l^\varepsilon \mid l \in L_\varepsilon, L_\varepsilon \subseteq L_{opt}\},$$

$$H^\varepsilon = \{h_j^\varepsilon \mid j \in J_\varepsilon, J_\varepsilon \subseteq J_{opt}\}, \forall \varepsilon = \overline{1, \varepsilon_0};$$

- ограничение F на минимальное количество атрибутов в суперклассе.

Эвристический алгоритм выделения суперклассов и классов-наследников является рекурсивным и включает в себя выполнение следующих операций:

- 1) формирование на основании исходных данных матрицы $M = \|m_{\varepsilon\varepsilon'}\|$, проиндексированной по обеим осям множеством объектов данных, значения $m_{\varepsilon\varepsilon'}$ которой вычисляются с использованием матрицы $S = \|s_{ij}\|$ и равны количеству однотипных элементов в составах классов O_ε и $O_{\varepsilon'}$;

- 2) проверка на основании матрицы $Q = \|q_{\varepsilon n}\|$ соответствия объектов O_ε и $O_{\varepsilon'}$ одному понятию: $q_{\varepsilon n} = 1$ and $q_{\varepsilon' n} = 1$? Если «да», то переход на 3). В противном случае – на 7);

- 3) проверка: $m_{\varepsilon\varepsilon'} \geq F$? Если «да», то переход на 4). В противном случае – на 7);

- 4) создание суперкласса $O_{\varepsilon\varepsilon'}$ путем переноса в него однотипных элементов из составов объектов O_ε и $O_{\varepsilon'}$, а также методов обработки данных, использующих данные информационные элементы (атрибуты);

- 5) модификация матрицы смежности $B_{kc}^{ob} = \|b_{\varepsilon\varepsilon'}\|$ путем добавления новых строки и столбца, соответствующих новому классу $O_{\varepsilon\varepsilon'}$: $b_{\varepsilon(\varepsilon\varepsilon')} = 1, b_{\varepsilon'(\varepsilon\varepsilon')} = 1$;

- 6) проверка: все объекты $O_\varepsilon \in O$ проанализированы? Если «да», переход на 7). В противном случае – на 2).

7) Конец. Завершение алгоритма.

Структурное преобразование графа G_{kc}^{ob} сводится к следующей последовательности операций: а) вводится новая вершина $O_{\varepsilon\varepsilon'}$, соответствующая суперклассу; б) от вершины $O_{\varepsilon\varepsilon'}$ до вершин O_{ε} и $O_{\varepsilon'}$ определяются дуги нового типа, характеризующие отношение наследования между суперклассом и классами (объектами)-наследниками. На диаграмме классов языка UML таким дугам соответствуют заштрихованные стрелки, направленные от наследника к суперклассу. В результате преобразования формируется унифицированная каноническая структура G_{ks}^{uni} .

Унифицированная объектная каноническая структура ЭБД представляется в виде графа $G_{ks}^{uni}(O^*, \Delta^*)$ и матрицы смежности $B_{ks}^* = \|b_{\varepsilon\varepsilon'}^*\|$ между объектами, модифицированной с учетом информации о выявленных суперклассах и классах-наследниках и отношениях между ними. Свойства объектов данных описываются матрицей смежности $V^* = \|v_{i\varepsilon}^*\|$, а поведение - технологической матрицей смежности $W^* = \|w_{j\varepsilon}^*\|$. Характеристиками графа G_{ks}^{uni} являются количественные показатели объектов данных $\{z_{\varepsilon}^{эл}, z_{\varepsilon}^{нр}\}$.

Заключение

В работе рассмотрены проблемы и задачи цифровой трансформации системы управления интеллектуальной собственностью. Предложены формализованные методы анализа и построения онтологической модели и концептуальной структуры предметной области цифрового информационного фонда интеллектуальной собственности. Рассмотрены методы и алгоритмы формирования объектной канонической структуры эталонной базы данных ЦИФИС и ее анализа.

Разработанные модели и методы использовались при формировании и развитии ЦИФИС евразийского патентно-информационного пространства [8, 10]. Их применение позволило

повысить и качество создаваемых БД региональной и национальной патентной информации (БД патентной документации ЕАПВ, БД патентной документации России, БД патентной документации Казахстана, БД патентной документации Белоруссии и др.), а также обеспечить возможность проведения полноценных патентно-информационных поисков.

Литература

1. В.В. Кульба, В.О. Сиротюк Формализованная методология повышения эффективности и качества патентных информационных фондов и опыт ее использования при формировании и развитии евразийского патентно-информационного пространства. - М.: ИПУ РАН. Научное издание, 2019. - 236с.
2. Кульба В.В., Сиротюк В.О. Формализованная методология обеспечения полноты патентных информационных фондов. - Управление развитием крупномасштабных систем MLS-2018. Труды 11-й Международной конференции. Том 3, с.127-138.
3. В.О. Сиротюк Формализованные модели и методы анализа и оценки полноты патентных информационных фондов (на примере международной патентной организации) - ж. Проблемы управления, М.: ИПУ РАН, №6/2018, с.35-43.
4. Лапшин В. А. Онтологии в компьютерных системах. - М.: Научный мир, 2010. - 222 с.
5. Волохов В.М., Прохоров А.И., Амосова Е.С. Онтологическая модель предметной области информационной поддержки инновационных разработок институтов РАН. - Информационные ресурсы России, 2011/3. - с.27-32.
6. Интеллектуальная собственность в терминах, понятиях и определениях федерального законодательства: Словарь-справочник. - М.: издание Совета Федерации, 2008. - 104с.
7. Common European Research Information Format (CERIF)(электронный ресурс) - <https://joinup.ec.europa.eu/collection/eu-semantic-interoperability-catalogue/solution/common-european-research-information-format/about>
8. Х.Ф. Фаязов, В.О. Сиротюк, А.В. Овчинников, А.Б. Бурцев Формирование и развитие евразийского патентно-информационного пространства. М.: ИНИЦ «Патент», 2010, - 124 с.
9. Кульба В.В., Ковалевский С.С., Косяченко С.А., Сиротюк В.О. Теоретические основы проектирования оптимальных структур распределенных баз данных. Сер. «Информатизация России и на пороге XXI века». - М.: СИНТЕГ, 1999. - 660 с.
10. Материалы сайта Евразийской патентной организации: www.eapo.org/ru

Сиротюк Владимир Олегович. Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В.А. Трапезникова РАН, г. Москва. Ведущий научный сотрудник, доктор технических наук, доцент. Количество печатных работ: 120 (в т.ч. 10 монографий). Область научных интересов: информационные технологии, оптимизация структур баз данных, информационная безопасность, управление качеством информации, цифровизация систем управления интеллектуальной собственностью. E-mail: vsirotyuk@ipu.ru

Methods for Constructing and Analyzing an Ontological Model and a Master Database of a Digital Intellectual Property Fund

V. O. Sirotyuk

Federal State Budgetary Institution of Science Institute of Control Sciences of Russian Academy of Sciences, Moscow

Abstract. The goals, objectives and principles of digital transformation of traditional intellectual property (IP) management systems, the requirements for the digital information fund of intellectual property (DIFIP), which plays a fundamental role in building a digital control body of the IP management system, and the features of building an ontological model of the subject area of the IP management system are considered. Methods for constructing and analyzing the ontological model and conceptual structure of the DIFIP subject area are proposed. Methods for constructing and analyzing the canonical structures of master databases (MDB) of patent and scientific and technical information DIFIP have been developed. Object-oriented methods of analysis and design are used to build an ontological model and canonical structure of the MDB.

Keywords: intellectual property management system, intellectual property digital information fund, patent information databases, scientific and technical information databases, data objects, subject area ontology model, master database, WIPO standards/

DOI 10.14357/20718632220306

References

1. V.V. Kulba, V.O. Sirotyuk Formalizovannaya metodologiya povysheniya effektivnosti i kachestva patentnykh informacionnykh fondov i opyt ee ispolzovaniya pri formirovani i razviti evrazijskogo patentno-informacionnogo prostranstva. - M.:IPU RAN. Nauchnoe izdanie, 2019.- 236s. [Formalized methodology for improving the efficiency and quality of patent information funds and the experience of its use in the formation and development of the Eurasian patent information space-Moscow: ICS RAS. Scientific publication, 2019. – 236p.].
2. Kulba V.V., Sirotyuk, V.O. Formalizovannaya metodologiya obespecheniya polnoty patentnykh informacionnykh fondov. – Upravlenie razvitiem krupnomasshtabnykh sistem MLSD-2018. Trudy 11-j Mezhdunarodnoj konferencii. Tom 3, s.127-138. [Formalized methodology for ensuring the completeness of patent information funds. - Management of the development of large-scale systems MLSD-2018.Proceedings of the 11th International Conference. Volume 3, p.127-138].
3. V.O. Sirotyuk Formalizovannye modeli i metody analiza i ochenki polnoty patentnykh informacionnykh fondov (na primere mezhdunarodnoj patentnoj organizacii) - zh. Problemy upravleniya, M.: IPU RAN, №6/2018, s.35-43. [Formalized models and methods for analyzing and evaluating the completeness of patent information funds (on the example of the international patent organization) - Moscow: Problems of control, №6/2018, p.35-43].
4. Lapshin V. A. Ontologii v kompyuternykh sistemah. - M.: Nauchnyj mir, 2010. - 222 s. [Ontologies in computer systems- Moscow: Scientific world, 2010. – 222p.].
5. Volohov V.M., Prohorov A.I., Amosova E.S. Ontologicheskaya model predmetnoj oblasti informacionnoj podderzhki innovacionnykh razrabotok institutov RAN. – Informacionnye resursy Rossii, 2011/3. – s.27-32. [Ontological model of the subject area of information support for innovative developments of institutes of the Russian Academy of Sciences. - Information resources of Russia, 2011/3. – p.27-32].
6. Intellektualnaya sobstvennost v terminah, ponyatiyah i opredeleniyah federalnogo zakonodatelstva: Slovar-spravochnik. – M.: izdanie Soveta Federacii, 2008. – 104s.

- [Intellectual property in terms, concepts and definitions of federal legislation: Dictionary-reference book. - Moscow: publication of the Federation Council, 2008. – 104p.].
7. Common European Research Information Format (CERIF): <https://joinup.ec.europa.eu/collection/eu-semantic-interoperability-catalogue/solution/common-european-research-information-format/about>.
 8. H.F. Fayazov, V.O. Sirotyuk, A.V. Ovchinnikov, A.B. Burtcev Formirovanie i razvitie evrazijskogo patentno-informacionnogo prostranstva. M.: INIC «Patent», 2010, - 124 s. [Formation and development of the Eurasian patent information space. – Moscow: INIC «Patent», 2010, - 124p.].
 9. Kulba V.V., Kovalevski S.S., Kosyachenko S.A., Sirotyuk V.O. Teoreticheskie osnovy proektirovaniya optimalnyh struktur raspredelennyh baz dannyh. Ser. «Informatizaciya Rossi i na poroge HHI veka». – M.: SINTEG, 1999. – 660 s. [Theoretical foundations for designing optimal structures for distributed databases - Moscow: SINTEG, 1999. – 660p.].
 10. Materialy sajta Evrazijskoj patentnoj organizacii: www.eapo.org/ru [Materials of the website of the Eurasian Patent Organization: www.eapo.org/ru].

Sirotyuk V.O. Ph.D, Federal State Budgetary Institution of Science Institute of Control Sciences of Russian Academy of Sciences, Moscow, Profsoyuznaya st., 65, e-mail: vsirotyuk@ipu.ru