

В.В. Диковицкий
**МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ И
ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ В МУЛЬТИПРЕДМЕТНЫХ
ИНФОРМАЦИОННЫХ СИСТЕМАХ ПРОМЫШЛЕННЫХ
ПРЕДПРИЯТИЙ**

Диковицкий В.В. Методы интеллектуальной обработки и представления информации в мультипредметных информационных системах промышленных предприятий.

Аннотация. В работе рассматривается задача формирования единого информационного пространства промышленного предприятия в виде мультипредметной информационной системы промышленного предприятия (МИСПП). Предложена архитектура мультипредметной информационной системы промышленного предприятия, метод автоматизированного формирования семантической модели предметной области информационной системы на основе принципа «пользователь как эксперт», метод формирования когнитивных пользовательских интерфейсов, адаптированных для различных категорий пользователей, и метод поиска, обеспечивающий автоматизированное расширение запроса и оценку релевантности результатов поиска на основе совместного анализа формализованной ментальной модели и семантической модели предметной области с учетом субтрактивных отношений.

Ключевые слова: мультипредметная ИС, навигационная структура, когнитивный пользовательский интерфейс.

Dikovitsky V.V. Methods of Intellectual Data Processing and Presentation in Multi-Subject Information Systems of Industrial Enterprises.

Abstract. In this paper, the problem of improving the technology of formation and functioning of multi-subject intellectualized information systems of industry is considered. As a solution to this problem, we propose an architecture of multi-subject IS information system of an industrial enterprise; an automated method of formation of a semantic domain model based on the principle of "user as an expert."; a method of formation of cognitive user interfaces adapted for different categories of users; and search method, providing automated query expansion and evaluation of the relevance of search results based on a joint analysis of the formal mental models and semantic models of a domain with the subtractive relationship.

Keywords: multisubject IS, navigation structure, cognitive user interface.

1. Введение. Деятельность современных промышленных предприятий требует аккумуляции и обновления знаний в различных предметных областях. Диверсификация современных производств влечет рост количества процессов, отражающихся в нормативно-справочной информации (НСИ) промышленного предприятия, являющейся объектом рассмотрения в данной работе. НСИ современного производственного предприятия претерпевает быстрые изменения, одни и те же процессы и объекты описываются в различных документах и могут рассматриваться специалистами с различных точек зрения [1], что усугубляет проблему разнородности и рассогласованности информации.

Существующие корпоративные информационные системы обладают рядом недостатков, делающих их неприменимыми в существ-

вующих условиях производства. К ним относятся привлечение экспертов по знаниям для поддержания актуальности знаний, реорганизация управления, ориентированность на четко структурированную иерархическую систему процессов, комплексный подход к интеграции затрудняет применение уже существующих на производстве информационных сервисов[2].

Это обуславливает необходимость формирования единого информационного пространства промышленного предприятия в виде мультипредметной информационной системы [3] промышленного предприятия (МИСПП). Предлагаемые в данной работе методы основаны на использовании формализованных знаний [4-7] в интеллектуализированной замкнутой информационной системе в условиях динамических изменений предметной области и отсутствии специалиста по знаниям. Интеллектуализированная замкнутая информационная система – это информационная система, основанная на знаниях, обладающая обратной связью, позволяющей пользователю влиять на процесс обработки информации с целью осуществления эффективного поиска. Повышение эффективности достигается за счет динамического обновления и адаптации к различным ментальным стереотипам пользователей семантической модели предметной области (СМПО) на основе пользовательского опыта (концепция “user as an expert” [8]).

Далее представлены архитектура МИСПП, принципы построения и формальное описание ментальных моделей категорий пользователей, метод автоматизированного формирования СМПО, формальная постановка задачи и метод формирования адаптивных пользовательских интерфейсов, метод поиска информации с учетом субтрактивных отношений СМПО - отношений между концептами СМПО, имеющих отрицательный весовой коэффициент. Использование субтрактивных отношений СМПО позволяет автоматизировать процесс добавления ограничений в расширенный запрос. Данные методы направлены на решение задачи организации эффективного доступа средствами поиска и навигации к ресурсам информационной системы промышленного предприятия.

2. Модель мультипредметной информационной системы промышленного предприятия. Предметная область современных производственных процессов по составу и структуре динамична. Для поддержки актуальности знаний требуется верифицируемая семантическая модель предметной области. Реализации обратной связи с целью адаптации представления информации и верификации СМПО требует хранения и учета пользовательского опыта в виде формализованных ментальных моделей (ФММ). Наличие ФММ и СПМО в

структуре ИС вносит изменения в процесс обработки данных. Логическая структура МИСПП представлена на рисунке 1.

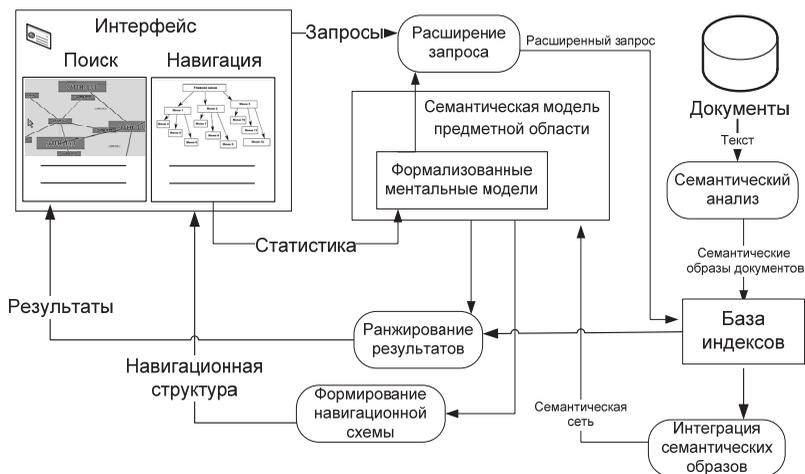


Рис. 1. Логическая структура мультипредметной информационной системы промышленного предприятия

Основу МИСПП составляют семантическая модель предметной области (СМПО), формализованная ментальная модель пользователей (ФММ), база индексов документов НСИ. Входными данными, обеспечивающими функционирование МИСПП, являются документы предприятия и статистика работы пользователя. Выходными данными являются результаты обработки запросов и навигационная структура интерфейса. СМПО формируется в результате интеграции семантических образов, вновь появляющихся в информационных хранилищах документов, и последующей верификации структуры модели на основе формализованных ментальных моделей пользователей. Семантический образ документа – семантическая сеть, множество вершин которой составляют понятия СМПО, присутствующие в документе, множество ребер – множество двухместных отношений над понятиями. Запрос – множество понятий предметной области, представленных множеством ключевых слов или множеством элементов навигационного интерфейса. Информационный элемент – документ информационной системы предприятия. Формализованная ментальная модель - семантическая сеть, множество вершин которой составляют понятия СМПО, которыми оперирует пользователь, множество ребер – множество взвешенных двухместных отношений над понятиями. ФММ служит «эталоном» в процессе формирования пользовательского интерфейса и оцен-

ки его когнитивности как степени соответствия структуры интерфейса формализованной ментальной модели пользователя. ФММ формируется на основе обработки запросов пользователя и статистики его работы с информационными элементами. Схема взаимодействия пользователя и мультимедийной информационной системой может быть представлена следующим алгоритмом (рисунок 2):

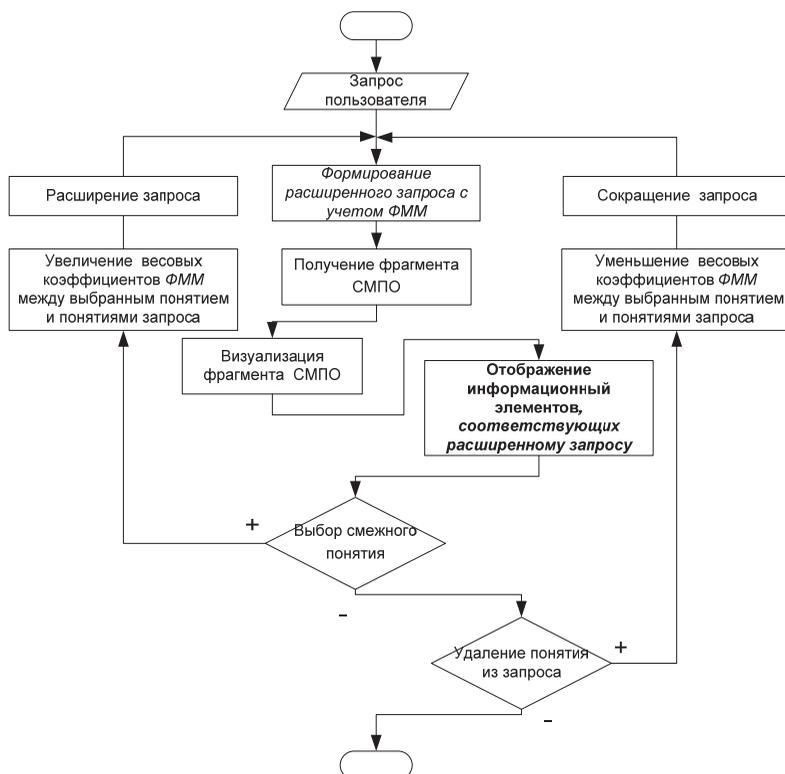


Рис.2. Алгоритм взаимодействия пользователя и ИС

Взаимодействие заключается в итеративном расширении запросов пользователя на основе ФММ, обеспечении возможности коррекции запроса средствами адаптивного интерфейса, а также учета пользовательских предпочтений путем коррекции весовых коэффициентов между понятиями ФММ. При совместном употреблении в запросе пар понятий, входящих в состав ФММ, увеличиваются весовые коэффициенты отношений между данными понятиями. Системные информаци-

онные базы используются в дальнейшем для формирования навигационно-поискового интерфейса, обеспечивающего эффективный, в смысле скорости и релевантности, доступ пользователя к требуемым данным. Последовательность действий, описывающих взаимодействие компонентов МИСПП, представлена следующими диаграммами:

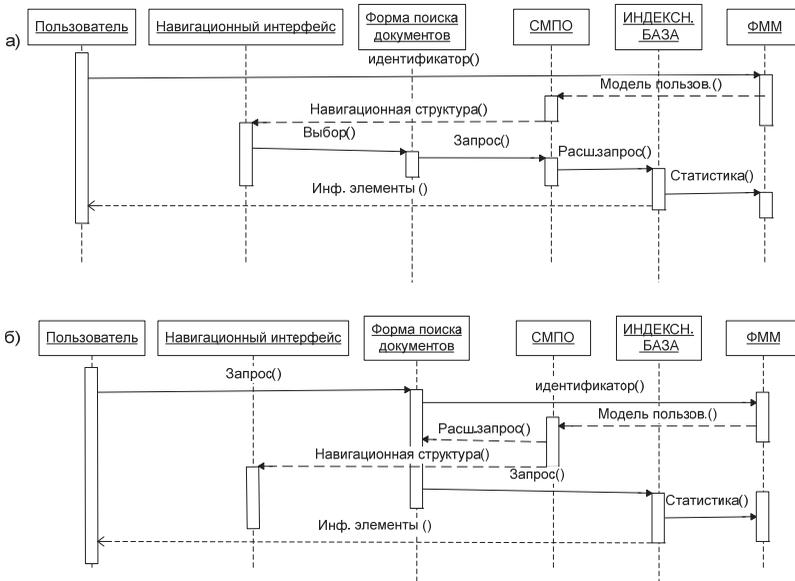


Рис.3. Диаграмма последовательности действий, а) навигация б) поиск

С целью снижения объемов хранимых системных данных модель строится не для каждого пользователя в отдельности, а для категорий пользователей. Формально категории пользователей определяются следующим образом: C - некоторое множество понятий (концептов), U - множество пользователей. Каждый концепт c имеет множество атрибутов:

$$A(c) = \{a(c)_i\}, a(c)_i \in C, i = \overline{1, N_c},$$

Упорядочив множество атрибутов по убыванию степени их значимости для пользователя u , получим последовательность, характеризующую его представление о данном концепте:

$$A^u(c) = \{a^u(c)_i\}, i = \overline{1, N_c} : a^u(c)_i \varphi^u a^u(c)_j, \forall i \leq j,$$

где φ^u - отношение, задающее значимость атрибутов для пользователя u ; $a \varphi^u b$ означает, что «для пользователя u a не менее значим, чем b ».

Определим группу пользователей, имеющих схожие представления о понятиях из некоторого множества C . Назовем подобную группу пользовательской *категорией* k -го порядка на множестве концептов C , и определим ее следующим образом:

$$U_C^k = \{u \mid \{a^u(c)_i\} = \{a^{u'}(c)_i\}, i = \overline{1, k}, \forall c \in C, \forall u' \in U_C^k\},$$

Ментальная модель некоторой k -й категории пользователей представлена взвешенным мультиграфом:

$$UM_k = \{C, L_k\}, L_k = \{l_k^{jm}\},$$

где C – множество вершин графа, представляющих понятия СМПО, L_k – множество взвешенных дуг, вес которых характеризует значимость семантического отношения m -го типа между i -м и j -м понятиями для k -й категории пользователей.

В совокупности, ФММ всех пользователей образуют фрагмент базы знаний МИСПП, представляющий собой мультиграф с векторными весами дуг. Матрица инцидентности мультиграфа имеет размерность 3:

$$M_l : C \times L \times U_k \rightarrow w_k^{ij},$$

где C – множество вершин графа, представляющих понятия СМПО, L – множество дуг, задающих отношения над C , U_k – множество категорий пользователей. Элементами матрицы являются весовые коэффициенты w_k^{ij} , задающие вес связи между концептами c_i и c_j для k -й категории пользователей. На рисунке 4 представлены фрагменты ментальных моделей пользователей различных категорий.



Рис.4. Атрибуты понятия «Карьерный экскаватор» для различных категорий пользователей

3. Метод автоматизированного формирования семантической модели предметной области промышленного предприятия информационной системы на основе принципа «пользователь как эксперт».

Формально СМПО представлена n-арной семантической сетью:

$$KB = \{C, L, Tr\}, Tr = \{synonymOf, HyponymOf, associateWith, subStruct\}$$

где C - множество концептов, L - множество отношений над концептами. Tr – множество типов отношений. Основой СМПО выступает тезаурус русского языка, расширяемый семантическими образами документов информационной системы. Процесс формирования семантической модели предметной области предприятия на основе коллекции документов информационной системы и расширяемого тезауруса состоит из следующих этапов:

1. Формирование семантического образа документа D статистическими и лексико-грамматическими методами обработки текста:

$$D = \{C_D, L_D\}, C_D \subset C, L_D = C_D \times C_D, L_D \subset L,$$

где D - семантическая сеть, C_D - множество концептов, выделенных в

документе, L_D - множество отношений, выделенных в документе, C - множество всех концептов СМПО, L - множество всех отношений СМПО.

2. Интеграция семантических образов D в СМПО на основе модифицированной составной семантической метрики [9], заключающейся в оценке близости двух концептов путем вычисления следующих оценок:

I. Вычисление оценки сходства контекста концептов документа C_D с контекстом СМПО:

$$\forall c_i^D \in C : \text{Syn}(c_i^D, c_j^{KB}) = |C_s^D \cap C_s^{KB}|, i = \overline{1..n}, j = \overline{1..m},$$

где C_s^D, C_s^{KB} - множество синонимов концепта c_i^D и c_j^{KB} соответственно.

II. Вычисление оценки сходства структурного положения концептов документа C_D и структурного положения концептов СМПО как количества общих гипонимов:

$$\forall c_i^D \in C : \text{Poseq}(c_i^D, c_j^{KB}) = |C_H^D \cap C_H^{KB}|, i = \overline{1..n}, j = \overline{1..m}$$

где C_H^D, C_H^{KB} - множество гипонимов концепта c_i^D и c_j^{KB} соответственно.

III. Вычисление оценки сходства имен понятий c документа D с именами понятий СМПО:

$$\forall c_i^D \in C : \text{Eq}(c_i^D, c_j^{KB}) = x/l, i = \overline{1..n}, j = \overline{1..m}$$

где x – длина эквивалентной цепочки символов в именах понятий c_i^D и c_j^{KB} , l – длина имени понятий.

IV. Добавление отсутствующих в СМПО концептов и отношений на основании результатов вычисления пороговой функции от среднего трех оценок:

$$f(c_i^D) = \frac{a \cdot \text{Eq}(c_i^D, c_j^{KB}) + b \cdot \text{Poseq}(c_i^D, c_j^{KB}) + \text{Syn}(c_i^D, c_j^{KB})}{3} > z, \\ i = \overline{1..n}, j = \overline{1..m},$$

где a, b, c – некоторые коэффициенты, z – значение пороговой функции, которые определяются экспертно, исходя из объема и разнородности

предметной области.

3. Верификация СМПО осуществляется изменением весовых коэффициентов существующих отношений между понятиями. Данный процесс инициируется при совместном использовании двух понятий в одном пользовательском запросе. Величина изменения весового коэффициента определяется следующим образом:

$$dWU'(t, n_i) = \frac{\sum_{\langle t, n_j \rangle \in (ACW)} (WU_k(t, n_j))}{|ACW|} * (WU(t, n_i) + dWU(t, n))$$

– для весового коэффициента отношения между совместно использованными понятиями, и

$$dWU'(t, n_i) = - \frac{\sum_{\langle t, n_j \rangle \in (ACW)} (WU_k(t, n_j))}{|CW / ACW|} * (WU(t, n_i) + dWU(t, n))$$

– для весового коэффициента отношения между остальными (неиспользованными) понятиями, где $dWU(t, n)$ - предыдущее значение изменения коэффициента, ACW и CW – множества отношений с совместно использованными и неиспользованными понятиями.

Предложенный метод предлагается использовать для автоматизированного формирования и обеспечения актуального состояния семантической модели предметной области производственного предприятия в условиях динамических изменений предметной области без привлечения эксперта по знаниям. На основе полученной модели решаются задачи обеспечения эффективного доступа пользователей к НСИ предприятия, а именно реализуются методы поиска и формирования адаптивных пользовательских интерфейсов.

5. Метод формирования адаптивных пользовательских интерфейсов обеспечивает повышение когнитивности МИС за счет автоматизированного формирования интерфейсов, адекватных моделям пользовательских интересов.

Пользовательский интерфейс представляет собой пару:

$$UI = \langle I, s \rangle,$$

где I – множество информационных элементов; s – навигационная структура. Навигационная структура определяет иерархию групп информационных элементов (ИЭ). На каждом уровне иерархии исходное множество информационных элементов делится на

подмножества в соответствии с одним или несколькими классификационными признаками. В качестве классификационных признаков используются атрибуты понятий предметной области. Очевидно, что при использовании на одном уровне навигационной структуры нескольких атрибутов, полученные множества ИЭ могут пересекаться. Введем следующие обозначения:

$\Gamma^l = \{G_i^l\}$ – множество разделов l -го уровня навигационной структуры;

G_i^l – i -я группа информационных элементов l -го уровня навигационной структуры;

$P^l = \{p_i^l\}$ – множество классификационных атрибутов, используемых для формирования групп ИЭ на l -м уровне навигационной структуры.

Взаимодействия пользователя и МИСПП заключается в поиске некоторых информационных элементов по имеющемуся у пользователя образу. При этом образ, чаще всего, неточный: в нем специфицируются лишь некоторая часть идентифицирующих атрибутов. Вследствие этого, пользователь с разной степенью уверенности может предполагать в какой из групп ИЭ на некотором уровне навигационной структуры находится искомый элемент. Эта уверенность тем выше, чем более точно представляет пользователь информационные элементы, составляющие потенциальное содержимое группы. Введем следующую функцию, задающую числовую оценку степени уверенности пользователя u (чем выше значение, тем выше степень уверенности):

$$p^u : \Gamma^l \rightarrow (0,1] .$$

Оценка времени, требуемого для доступа к искомому информационному элементу в рамках навигационной структуры на l -м уровне, будет равна:

$$O\left(\frac{\max_i |G_i^l|}{p^u(G_i^l)}\right).$$

Таким образом, при прочих равных условиях, степень уверенности пользователя в принадлежности информационного элемента к той или иной группе определяет качество интерфейса в смысле скорости доступа к требуемой информации.

Пусть навигационная структура интерфейса имеет глубину \hat{l} уровней. Тогда в качестве количественной оценки когнитивности

интерфейса для пользователя u может использоваться сумма:

$$\sum_{l=1}^{\hat{l}} p^u(l),$$

где \hat{l} – количество уровней навигационной структуры интерфейса, $p^u(l)$ – функция, задающая числовую оценку степени уверенности пользователя u в какой из групп ИЭ на уровне l навигационной структуры находится искомый элемент.

Данная мера может использоваться для оценки уже существующих интерфейсов, когда известно значение \hat{l} . Для решения же прямой задачи, то есть структуризации исходного множества информационных элементов в рамках навигационной структуры, требуется учитывать дополнительные ограничения. Эти ограничения обусловлены психологией восприятия человека, ограничивающей максимальное количество одновременно эффективно воспринимаемых объектов. Вследствие этого необходимо ограничивать размер группы информационных элементов, а также глубину навигационной структуры.

С учетом сказанного, оптимальная для пользователя u структура интерфейса есть решение следующей задачи с ограничениями [10]:

$$\begin{aligned} \max_s \sum_{l=1}^{\hat{l}(s)} p^u(l) \\ g(s) \leq K, \\ \hat{l}(s) \leq K'. \end{aligned}$$

Здесь $\hat{l}(s)$ – количество уровней в навигационной структуре s ; $g(s)$ – максимальный размер группы информационных элементов $\hat{l}(s)$ -го в навигационной структуре s ;

K – когнитивная константа, определяющая максимальное число одновременно предъявляемых пользователю информационных элементов для их эффективного восприятия;

K' – когнитивная константа, определяющая максимальное число уровней навигационной структуры, в рамках которых поиск информации для пользователя остается комфортным. Константы K и K' определяются экспертно исходя из соображений психологии восприятия [11–14].

Для формирования навигационного интерфейса, удовлетворяющего приведенной формулировке задачи синтеза, предлагается процедура, основанная на формализованной ментальной модели. Процедура содержит несколько этапов:

1. Определение информационной потребности специалиста на основе модели пользовательских интересов и текущего запроса:

$$UQ = f_{tr}(Q, UM), UM : \exists l \exists L : t_i \in Q, t_j \in UM,$$

где UM - модель пользовательских интересов; Q - запрос специалиста, f_{tr} - функция трансляции запроса в модель пользовательских интересов;

2. Определение множества информационных элементов навигационной структуры, соответствующих информационной потребности специалиста.

$$G = \{g \mid \exists a^u = g\}, a^u \in UQ,$$

где G - множество информационных элементов навигационной структуры; a^u - концепты текущей информационной потребности UQ .

3. Разбиение множества информационных элементов навигационной структуры на подмножества и их ранжирование в соответствии с весовыми коэффициентами модели пользовательских интересов.

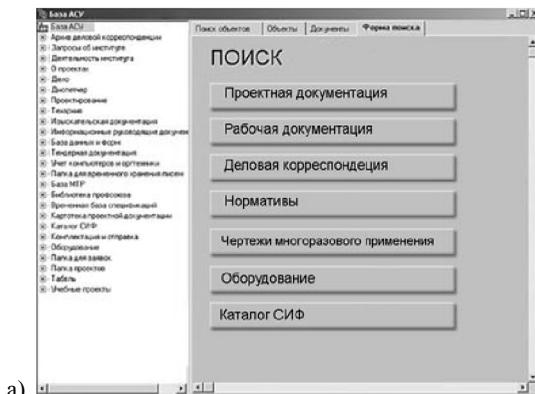
$$G_i^d = \{g_k \mid (\forall g_k, g_m \exists l \in L : w(g_k, g_m) > x) \wedge (\exists g_z : \exists l_{sup}(g_k, g_z), l_{sup}(g_m, g_z) \in L)\}$$

$$k, m = \overline{1..n}, z = \overline{1..d}, G_i^d \subset G,$$

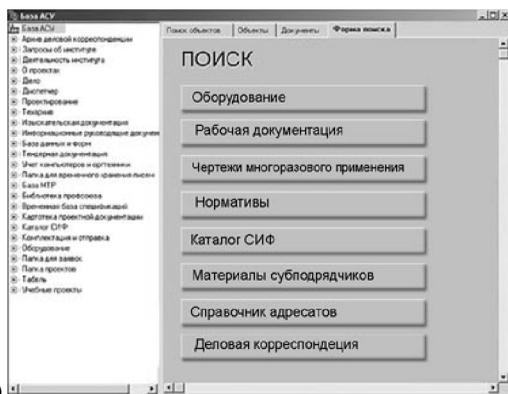
где G_i^d - i -я группа информационных элементов d -го уровня навигационной структуры; $l_{sup}(g_k, g_z)$ - отношение классификации в модели интересов пользователя;

$w(g_k, g_m)$ - весовой коэффициент отношения l над информационными элементами g_k, g_m ; d - количество уровней навигационной структуры; x - порог вхождения информационного элемента в навигационную структуру, задается на основе ограничений на максимальное число одновременно предъявляемых пользователю информационных элементов K .

На основе групп множеств G формируется интерфейс мультипредметной информационной системы в виде множества меток («tags»), отражающих информационное содержание ресурсов в терминах СМПО с учетом модели категории пользователей. На рисунке 5 представлен адаптивный пользовательский интерфейс, динамически изменяющий навигационную структуру в зависимости от ментальной модели пользователя.



а)



б)

Рис.5. Вид пользовательского интерфейса для категорий пользователей: а) отделов «Управление производством»; б) «Техническое планирование»

6. Метод поиска информации с учетом субтрактивных отношений позволяет увеличить эффективность поиска в смысле релевантности результатов за счет автоматического включения в запрос ограничений. Ограничения формируются на основе субтрактивных отношений искомого концепта, заданных в СМПО. Использование субтрактивных отношений СМПО позволяет автоматизировать процесс добавления ограничений в расширенный запрос.

Метод поиска включает 3 составляющих: Способ представления документа, способ представления запроса и функцию соответствия запроса результатам. Документ представлен фрагментом СМПО и множеством ключевых слов (меток) в базе индексов D , выделенных семантическим анализатором на этапе индексации. Запрос представлен

в виде множества ключевых слов Q . Процесс поиска документов, соответствующих запросу, состоит из следующих этапов:

1. Формирование запроса в терминах СМПО:

1.1. Формирование расширенного запроса EQ , содержащего отношения и соответствующие запросу $Q = \{c_i\}$ концепты СМПО:

$$EQ = f_{tr}(KB, Q) = \{T^q, L^q \mid Eq(c_i^Q, t_j^{KB}) > 1 - \varepsilon\}, L = \{l\} = \{< t_i, t_j, tp, \bar{w}_k >\},$$

$$i = \overline{1..n}, j = \overline{1..m}, tp \in Tp, k = \overline{1..r},$$

где KB – семантическая модель предметной области, T^q – множество концептов СМПО, L^q – множество отношений над концептами, $f_{tr}()$ – функция трансляции запроса в модель пользовательских интересов, устанавливающая соответствие между словами запроса и концептами базы знаний, $Eq()$ – оценка сходства концептов, t_i, t_j – концепты, tp – тип отношения, \bar{W} – вектор весовых коэффициентов отношений, компоненты которого соответствуют уровню значимости отношений для различных групп пользователей, r – количество групп пользователей, ε – погрешность оценки сходства концептов.

1.2. Расширение запроса с учетом весовых коэффициентов отношений и субтрактивных отношений, ограничивающих контекст запроса:

$$EQ = \{T^q, L^q\} \cup \{T', L' \mid \forall l \in L' : t_i \in T^q, t_j \in T', f_u(\bar{w}_k, n) > x\},$$

$$\{T', L'\} \in KB, \bar{w}_k = \langle w_1, \dots, w_k, \dots, w_r \rangle,$$

где T' – множество концептов СМПО, связанных с концептами множества T^q отношениями из множества L' вида, $f_u(\bar{w}_k, r)$ – функция получения n -ой компоненты вектора весовых коэффициентов отношения l , x – порог включения отношения в расширенный запрос, определяемый экспертно исходя из точности и полноты поиска по коллекции документов.

2. Получение множества документов D , соответствующих расширенному запросу EQ :

$$D = \{d_i \mid T^{d_i} \cap T^q \neq \emptyset\},$$

где T^{d_i} – множество терминов СМПО, присутствующих в документе d_i , T^q – множество терминов СМПО, присутствующих в запросе EQ .

3. Ранжирование множества документов с учетом весовых коэффициентов отношений:

$$R(d_k) = \sum_{L_{d_k}} (f_u(\bar{w}_k, r)) - \sum_{L'_{d_k}} (f_u(\bar{w}_k, r)),$$

$$L_{d_k} = \{l^d \mid (t_i, t_j \in d_k) \wedge (tp \in \{\text{synonymOf}, \text{HyponymOf}, \text{associateWith}\})\},$$

$$L'_{d_k} = \{l^d \mid (t_i, t_j \in d_k) \wedge (tp \in \{\text{subStract}\})\}, \quad i, j = \overline{1..n}, k = \overline{1..m}, tp \in Tp,$$

где $f_u(\bar{w}_k, r)$ – функция получения компоненты вектора весовых коэффициентов отношений из множества L_{d_k} между концептами t_i и t_j , присутствующими в документе d_k , для категории пользователей r . Tp – множество типов отношений. Результатом ранжирования является упорядоченное по убыванию оценки $R(d_k)$ множество документов $D = \{d_k\}$, представляющих результаты поиска. Таким образом, на ранжирование результатов оказывают влияние наличие понятий формализованной ментальной модели в документе, весовые коэффициенты между понятиями формализованной ментальной модели, коррелирующие с частотой использования понятий пользователем, а также наличие субтрактивных отношений.

7. Апробация. Предложенные методы были апробированы в информационной системе крупного промышленного предприятия региона – ОАО «Оленегорский горно-обогатительный комбинат». Семантическую модель предметной области мультипредметной информационной системы составляет русскоязычный тезаурус WordNet 3.0, расширенный результатами работы семантического анализатора над коллекцией нормативно-справочных документов предприятия. Эффективность предложенных методов была проверена путем натурального эксперимента с привлечением тестовой выборки пользователей, которым предлагалось решить схожие информационно-поисковые задачи с помощью существующих систем поиска с использованием интерфейса строки поиска и с помощью адаптивного интерфейса мультипредметной информационной системы.

Для оценки характеристик адаптивного пользовательского интерфейса был использован метод GOMS [15], заключающиеся в сравнении временных характеристик выполнения пользователями производственных задач в рамках существующего и разрабатываемого интерфейса. Целью пользователей являлось получение доступа к требуе-

тому документу. Среднее время выполнения операции представлены в таблице 1.

Таблица 1. Оценка показателей использования методов поиска

Оцениваемые характеристики	Навигация с использованием существующего интерфейса	Навигация с использованием адаптивного интерфейса
Среднее время, затраченное на навигационный поиск, с.	31,8	20,85

Для оценки метода поиска осуществлялся поиск по заранее проиндексированной коллекции объемом 14 тыс. документов. Экспертами оценивались процесс поиска результаты выполнения 10 запросов. В качестве критериев оценки выступали скорость поиска – время, затраченное на удовлетворение информационной потребности, выраженной одним запросом; точность – соответствие результатов запросу; и полнота результатов – полнота охвата документов с упоминанием об объекте поиска:

$$Precision = \frac{|D_{rel} \cap D_{retr}|}{|D_{retr}|}, \quad Recall = \frac{|D_{rel} \cap D_{retr}|}{|D_{rel}|},$$

где D_{rel} – множество релевантных документов в базе, D_{retr} – множество документов, найденных системой.

Для оценки альтернатив экспертам была предложена лингвистическая шкала измерений.

Оценка i -й альтернативы производилась j -м экспертом по формуле:

$$v_{ij} = 1 - \frac{(l-1)}{k},$$

где l – индекс значения лингвистической шкалы; k – количество значений этой шкалы.

Для оценки i -й альтернативы n экспертами используется формула:

$$s_i = \sum_{j=1}^n v_{ij}.$$

Результаты оценки характеристик приведены в таблице 2.

Таблица 2. Оценка показателей использования методов поиска

Оцениваемые характеристики	Используемый метод поиска документов	
	Поиск по ключевым словам с использованием интерфейса строки ввода	Поиск с использованием метода поиска информации с учетом субтрактивных отношений
Оценка скорости выполнения поиска	0,45±0,09	0,82±0,15
Точность результатов	0,96±0,04	0,91±0,08
Полнота результатов	0,73±0,09	0,94±0,06
Среднее значение оценок	0,71	0,89

Высокая точность поиска с использованием интерфейса строки ввода обусловлена знакомством пользователей с предметной областью и, как следствие, малой неопределенностью при формировании запроса, а также относительно малым объемом коллекции документов. Тем не менее, результаты экспериментов позволяют сделать вывод о корректности и обоснованности использования динамической автоматически формируемой и верифицируемой СМПО для реализации адаптивного интерфейса и метода поиска информации с учетом субтрактивных отношений в мультипредметных информационных системах промышленных предприятий.

8. Заключение. Рост объема, разнородность и динамика изменения информации современных промышленных предприятий обуславливают необходимость развития методов оперативного доступа к ней. Одним из путей решения данной проблемы является построение мультипредметных информационных систем, способных предоставить пользователю необходимые функциональные возможности для оперирования большими массивами данных. В данной работе представлены архитектура МИСПП, принципы построения и формальное описание ментальных моделей категорий пользователей, метод автоматизированного формирования СМПО, формальная постановка задачи и метод формирования адаптивных пользовательских интерфейсов, метод поиска информации с учетом субтрактивных отношений СМПО. В рамках документооборота промышленного предприятия проведена программная реализация и оценка эффективности разработанных методов. Результаты проведенного анализа эффективности разработанных методов мультипредметных информационных систем с использованием разработанного программного обеспечения показали, что при осуществлении навигации обеспечивается сокращение времени доступа к необходимым документам в 1,5 раза, при осуществлении информационного поиска обеспечивается сокращение времени поиска в 1,9 раза, и увеличение полноты результатов в 1,2 раза.

Литература

1. *Acker L., Porter B.* Extracting viewpoints from knowledge bases // Proceedings of The 12th National Conference on Artificial Intelligence. 1994. pp. 547–552.
2. *Трубейко К.Н., Кулешов А.А., Клебанов А.Ф., Владимиров Д.Я.* Современные системы управления горнотранспортными комплексами // СПб.: Наука. 2007. 344 с.
3. *Диковицкий В.В., Ломов П.А., Сепеда-Эррера Р.Р., Шишаев М.Г.* Современные методы создания мультипредметных веб-ресурсов на базе визуализации и обработки формализованной семантики // Вестник Кольского научного центра РАН. 2011. №3. С.63–73.
4. *Смирнов А.В., Пацкин М.П., Шилов Н.Г., Левашова Т.В.* Онтологии в системах искусственного интеллекта: способы построения и организации // Новости искусственного интеллекта. 2002. №1. С.3–13.
5. *Гаврилова Т.А., Гулякина Н.В., Голенков В.В.* Технология проектирования интеллектуальных систем // Информационные системы и технологии (IST'2009): Мат. V Междунар. конф.-форума. Минск: А.Н.Вараксин. 2009. С.93–96.
6. *Коголовский М.Р.* Перспективные технологии информационных систем // М.: Компания АйТи, 2003. 288 с.
7. *Кучуганов В.Н.* Элементы теории ассоциативной семантики // Управление большими системами: Сб. научн. тр. М.: ИПУ РАН. 2012. Вып. 40. С. 30–48.
8. *Шишаев М.Г., Ломов П.А., Диковицкий В.В.* Использование концепции «user as an expert» в разработке мультипредметных веб-ресурсов, основанных на онтологиях // Труды Института системного анализа РАН: Прикладные проблемы управления макросистемами. 2012. Т. 62. С.40–47.
9. *Шишаев М.Г., Ломов П. А.* Интеграция семантически связанных информационных ресурсов на основе онтологий // Информационные технологии в региональном развитии: Сборник научных трудов ИИММ КНЦ РАН. Апатиты: Изд-во КНЦ РАН. 2007. Вып. VII. С. 61–68.
10. *Шишаев М.Г., Ломов П.А., Диковицкий В.В.* Формализация задачи построения когнитивных пользовательских интерфейсов мультипредметных информационных ресурсов // Труды Кольского научного центра РАН. Информационные технологии. 2013. С.90–97.
11. *Koffka K.* Principles of Gestalt psychology // N.Y.: Routledge. 1935. 720 p.
12. *Miller G.* The Magical Number Seven, Plus or Minus Two // The Psychological Review. 1956. pp. 81–97.
13. *Brooke J., Jordan P.W., Thomas B., Weerdmeester B.A., McClelland I.L.* SUS: A 'quick and dirty' usability scale // Usability evaluation in industry. London: Taylor & Francis. 1996. pp.189–194.
14. *Салвенди Г.* Человеческий фактор // М: Мир. 1991.
15. *John, B. E., Kieras, D. E.* The GOMS family of user interface analysis techniques: Comparison and contrast // ACM Transactions on Computer-Human Interaction. 1996. pp. 320–351.

References

1. *Acker L., Porter B.* Extracting viewpoints from knowledge bases. Proceedings of The 12th National Conference on Artificial Intelligence. 1994. pp. 547–552.
2. *Trubeckoj K.N., Kuleshov A.A., Klebanov A.F., Vladimirov D.Ja.* *Sovremennyye sistemy upravleniya gorno-transportnymi kompleksami* [Modern control systems of mining and transport]. St. Petersburg. 2007. 344 p. (In Russ.).
3. *Dikovickij V.V., Lomov P.A., Sepeda-Jerrera R.R., Shishaev M.G.* [Modern methods of creating multi-subject web resources on the basis of visualization and processing of

- formal semantics]. *Vestnik Kol'skogo nauchnogo centra RAN – Herald of Kola Science Centre RAS*. 2011. vol. 3. pp. 63–73. (In Russ.).
4. Smirnov A.V., Pashkin M. P., Shilov N. G., Le-vashova T. V. [Ontologies in artificial intelligence systems: methods of construction and organization] *Novosti iskusstvennogo intellekta – AI NEWS Newsletter*. Moscow. 2002. vol. 1. pp.3–13. (In Russ.).
 5. GavriloVA T.A., Guljakina N.V., Golenkov V.V. [Technology of designing intelligent systems]. *Informacionnyye sistemy i tehnologii (IST'2009): materialy V Mezhdunar. konf.-forumu* [Information Systems and Technology (IST'2009): Collected papers]. Minsk: A.N.Varaksin. 2009. pp. 93–96. (In Russ.).
 6. Kogalovskij M.R. *Perspektivnyye tehnologii informacionnyh sistem* [Promising Information Systems Technology]. Moscow. 2003. 288 p. (In Russ.).
 7. Kuchuganov V.N. [Elements of the theory of associative semantics]. *Upravlenie bol'shimi sistemami: Sbornik trudov* [Managing large systems: Collected papers]. Moscow: IPU RAN. 2012. pp. 30–48. (In Russ.).
 8. Shishaev M.G., Lomov P.A., Dikovickij V.V. [Using the concept of «user as an expert» in the development of multi-subject web resources, based on ontologies]. *Trudy Instituta sistemnogo analiza RAN: Prikladnye problemy upravlenija makrosistemami – Proceedings of the Institute for Systems Analysis: Applied management problems of macrosystems*. Moscow. 2012. pp.40–47. (In Russ.).
 9. Shishaev M.G., Lomov P.A. [The integration of semantically related information resources on the basis of ontologies]. *Informacionnyye tehnologii v regional'nom razviii: Sbornik nauchnyh trudov IIMM KNC RAN – Information technologies in regional development: Proceedings of the IIMM KSC*. Apatity: KNC RAN. 2007. vol. VII. pp. 61–68. (In Russ.).
 10. Shishaev M.G., Lomov P.A., Dikovickij V.V [The formalization of the task of building cognitive user interfaces for multi-subject Information Resources]. *Trudy Kol'skogo nauchnogo centra RAN. Informacionnyye tehnologii – Proceedings of the Kola Scientific Center, Russian Academy of Sciences. Information technology*. 2013. pp. 90–97. (In Russ.).
 11. Koffka K. Principles of Gestalt psychology. N.Y.: Routledge. 1935. 720 p.
 12. Miller G. The Magical Number Seven, Plus or Minus Two. *The Psychological Review*. 1956. pp. 81–97.
 13. Brooke J., Jordan P.W., Thomas B., Weerdmeester B.A., McClelland I.L. SUS: A 'quick and dirty' usability scale. *Usability evaluation in industry*. London: Taylor & Francis. 1996. pp. 189–194.
 14. Salvendi G. *Chelovecheskij faktor* [Human factor]. Moscow: Mir, 1991. (In Russ.).
 15. John B.E., Kieras D.E. The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction*. 1996. pp. 320–351.

Диковицкий Владимир Витальевич — младший научный сотрудник, Федеральное государственное бюджетное учреждение науки Институт информатики и математического моделирования технологических процессов Кольского научного центра Российской академии наук. Область научных интересов: интеллектуализированные системы, адаптивный интерфейс, структурирование знаний, семантика. Число научных публикаций — 22. dikovitsky@gmail.com; ул. Ферсмана, д.30, кв.57, Апатиты, Мурманская обл., 184209; р.т.: +79021364888.

Dikovitsky Vladimir Vital'evich — junior researcher, Establishment of Russian Academy of Sciences Institute for Informatics and Mathematical Modelling of Technological Processes of the Kola Science Center RAS. Research interests: intellectualized systems, adaptive interface, structured knowledge, semantics. The number of publications — 22. dikovitsky@gmail.com; 30-57, Fersman St., Apatity, Murmansk Region, 184209; office phone: +79021364888.

РЕФЕРАТ

Диковицкий В.В. **Методы интеллектуальной обработки и представления информации в мультипредметных информационных системах промышленных предприятий.**

В работе рассмотрена задача формирования единого информационного пространства промышленного предприятия в виде мультипредметной информационной системы промышленного предприятия (МИСПП).

Во введении перечислены проблемы предметной области, в качестве решения предлагается формирование единого информационного пространства промышленного предприятия в виде мультипредметной информационной системы на основе использования формализованных знаний.

Рассмотрена логическая структура МИСПП, перечислены основные компоненты, обеспечивающие функционирование ИС, а так же принципы формирования семантической модели предметной области (СМПО), формализованной ментальной модели пользователей (ФММ), базы индексов документов.

Представлен метод автоматизированного формирования семантической модели предметной области промышленного предприятия на основе принципа «пользователь как эксперт». Формально описана СМПО и этапы процесса ее формирования.

Представлен метод формирования адаптивных пользовательских интерфейсов, обеспечивающий повышение когнитивности мультипредметной информационной системы за счет автоматизированного формирования интерфейсов, адекватных моделям пользовательских интересов. На основе ФММ и ограничений задачи синтеза предложена процедура формирования структуры навигационного интерфейса.

Представлен метод поиска, обеспечивающий автоматизированное расширение запроса и оценку релевантности результатов поиска на основе совместного анализа формализованной ментальной модели и семантической модели предметной области с учетом субтрактивных отношений. Рассмотрены способ представления документа, способ представления запроса и функция соответствия между ними, составляющие соответствующие этапы процедуры поиска, отличающейся автоматической выработкой ограничений на основе СМПО и адаптивным ранжированием результатов поиска на основе ФММ.

В заключении приводятся результаты апробации предложенных методов.

SUMMARY

Dikovitsky V.V. **Methods of Intellectual Data Processing and Presentation in Multi-Subject Information Systems of Industrial Enterprises.**

The paper considers the problem of forming a unified information space of an industrial enterprise in the form of a multi-subject information system of industrial enterprise (MISIE).

In introduction, problems of subject domain are listed. As a solution to these problems, formation of a unified information space of the industrial enterprise in the form of multi-subject information system based on the use of formalized knowledge is proposed.

The logical structure of MISIE is considered. The main components providing functioning of the IS as well as principles of the formation of a semantic domain model (SDM), a formalized mental model of users (FMM), and index databases are listed.

The method of automated formation of a semantic domain model of an industrial enterprise based on the principle "user as an expert" is presented. SDM and the steps in the process of its formation have been formally described.

The method of the formation of an adaptive user interface, providing increased cognition in multi-subject information system by automated formation of adequate for models of user interests interfaces. On the basis of the FMM and the constraints of the problem of synthesis a procedure for forming the structure of navigation interface is proposed.

The search method provides an automated query expansion and evaluation of the relevance of search results through a sharing of analysis of the formal mental models and semantic domain model with the subtractive relations. The method of presentation of the document, the method of the query presentation and the function of correspondence between them, constituting the corresponding steps in the procedure of search, which provides an automatic generation of restrictions based on SDM and adaptive ranking of search results based on the FMM are examined.

The results of the proposed technologies testing are presented in conclusion.