

Научная статья

DOI: 10.15593/2499-9873/2022.3.08

УДК 004.89

Д.В. Яруллин

Пермский национальный исследовательский политехнический университет, Пермь, Россия

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ ПОДГОТОВКОЙ ИТ-СПЕЦИАЛИСТОВ НА ОСНОВЕ ДЕНОТАТИВНОЙ АНАЛИТИКИ

В рамках исследования обозначена растущая проблема несоответствия компетенций выпускников ИТ-специальностей образовательных организаций требованиям работодателей. Поставлена цель разработки комплекса моделей и алгоритмов, а также реализующего их программного обеспечения для повышения эффективности подготовки ИТ-кадров.

Произведен анализ существующих подходов к ее решению. Выявлены перспективные направления моделирования специалиста. Рассмотрены методы анализа рынка труда в области ИТ-сферы на основании вакансий с сайтов-агрегаторов, подходы к индивидуализации образовательных программ и вовлечения работодателей в процесс обучения, а также средства автоматизации моделирования специалиста. Также изучены онтологические подходы к моделированию специалиста, опирающиеся на экспертные знания.

Разработана концепция организации подготовки ИТ-специалистов на основе открытых слабоструктурированных данных. Описан подход, основанный на обработке данных на естественном языке и денотативной аналитике. Предложенная концепция включает в себя мониторинг и сбор данных вакансий с сайтов-агрегаторов с заданной периодичностью; их последующую обработку для построения динамической модели ИТ-специалиста, описывающей ключевые понятия, необходимые для профессиональной деятельности, и удовлетворяющей требованиям работодателей; обогащение полученной модели открытыми данными о предметной области и информацией, содержащейся в нормативных документах; построение модели специалиста-выпускника, учитывающей связи между ключевыми понятиями, путем построения денотатного графа для восстановления иерархии обогащенного множества понятий; формирование на ее основе набора рекомендаций по составлению и корректировке образовательных программ.

Разработана модель специалиста с точки зрения работодателя на основе нечетких множеств, модель специалиста-выпускника, включающая алгоритм обогащения данных, а также алгоритм формирования рекомендаций по составлению и корректировке образовательных программ.

Разработано программное обеспечение системы поддержки принятия решений, реализующее указанные модели и алгоритмы. Проведена экспериментальная проверка с привлечением экспертов из ИТ-индустрии. По итогам эксперимента получены положительные результаты.

Разработанная система может использоваться для составления и корректировки рабочих программ дисциплин, программ профессиональной переподготовки, программ дополнительного образования в области направлений, связанных с информационными технологиями.

Ключевые слова: ИТ-специалист, компетенция, кластерный анализ, оценка качества, векторизация, денотат, денотатный граф, нечеткое множество, модель специалиста, рекомендательная система.

D.V. Yarullin

Perm National Research Polytechnic University, Perm, Russian Federation

INTELLIGENT CONTROL SYSTEM FOR IT SPECIALISTS TRAINING BASED ON DENOTATIVE ANALYTICS

The study identifies the growing problem of mismatch between the competencies of IT graduates of educational institutions and the requirements of employers. The goal is to develop a set of models and algorithms, as well as software that implements them to improve the efficiency of IT training.

The analysis of existing approaches to its solution is made. Prospective directions of specialist modeling are revealed. The methods of labor market analysis in the field of IT-sphere based on vacancies from aggregator sites, approaches to individualization of educational programs and involvement of employers in the training process, as well as means of automation of specialist modeling were considered. Ontological approaches to specialist modeling based on expert knowledge are also studied.

The concept of organizing IT training based on open weakly structured data has been developed. An approach based on natural language processing and denotative analytics is described.

The proposed concept includes monitoring and collecting data on vacancies from aggregator sites with a given frequency; their subsequent processing to build a dynamic model of an IT specialist, describing the key concepts necessary for professional activity and satisfying employers' requirements; enriching the resulting model with open data on the subject area and information contained in regulatory documents; building a graduate specialist model, taking into account the links between the key concepts by constructing a denotative graph to reconstruct the hierarchy of the enriched set of concepts; forming on its basis a set of recommendations for drawing up and adjusting educational programs.

The model of specialist from employer's point of view based on fuzzy sets, model of specialist-graduate including data enrichment algorithm, as well as the algorithm of recommendations formation for making and correction of educational programs were developed.

The software for decision support system that implements the mentioned models and algorithms has been developed. Experimental testing with the involvement of experts from the IT industry was conducted. Positive results were obtained as a result of the experiment.

The developed system can be used for drawing up and adjusting work programs of disciplines, professional retraining programs, programs of additional education in the field of directions related to information technology.

Keywords: IT specialist, competence, cluster analysis, quality assessment, vectorization, denotation, denotative graph, fuzzy set, specialist model, recommendation system.

Введение

Развитие отрасли информационных технологий и подготовка высококвалифицированных ИТ-кадров являются одними из приоритетных направлений в экономике Российской Федерации [1]. В настоящее время спрос на таких специалистов в значительной мере превышает предложение [2], поскольку потребность в ИТ-специалистах возрастает во всех областях, где происходят процессы автоматизации, цифровизации, аналитики; спрос на специалистов в сфере информационных технологий наблюдается во всех развитых странах [3; 4]. Одновременно с этим статистические исследования высшей школы демонстрируют тренд на снижение числа поступающих на специальности, связанные с информационными технологиями [5].

Данную тенденцию можно соотнести с тем, что компетенции, которые ожидают работодатели от специалистов, зачастую не соответствуют компетенциям, которые будущие специалисты приобретают в процессе получения высшего образования [6]. Известно, что уже в 2013 г. в Российской Федерации 15 % от всех существующих компаний сообщали, что работники недостаточно квалифицированы [7].

В результате все большее число соискателей предпочитает приобретать необходимые для трудоустройства компетенции бессистемно, путем самообразования, краткосрочных программ дополнительного образования, курсов повышения квалификации; также укрепляется тренд на переход от очного образования к электронному на основе дистанционных технологий.

Следовательно, мониторинг, обработка и систематизация требований работодателей становятся необходимыми для составления и реализации конкурентоспособных образовательных программ в высших учебных заведениях, нацеленных на выпуск специалистов, востребованных на рынке труда, что особенно актуально именно в сфере информационных технологий.

Целесообразным становится создание рекомендательной системы для составления и корректировки образовательных программ, которая использует модель востребованного специалиста в качестве основы для рекомендательных алгоритмов.

Существующие подходы

В подходах к моделированию специалиста можно обозначить два основных направления: построение модели с опорой на требования работодателей и построение модели с учетом требований профессиональных и образовательных стандартов.

Подходы к построению модели специалиста с учетом требований работодателей предложены такими исследователями, как В.В. Калевко, Д.Г. Лагерев, А.Г. Подвесовский [8], К.В. Булах, Е.Т. Бурцева [9], И.Ю. Харитоновна и С.Ф. Сергеев [10]. Изучением методов построения модели специалиста с учетом требований образовательных организаций, Федерального государственного образовательного стандарта и профессиональных стандартов занимались Р.М. Галиев [11], Л.В. Курзаева [12], А.А. Захарова [13], А.А. Воронин, А.А. Васильченко, А.В. Зе-

нович [14], Т.Ф. Лебедева, С.А. Муравьев [15]. Концепции формирования подготовки специалистов с учетом как требований работодателей, так и образовательных институтов описаны в работах М.Б. Гузаирова, И.Б. Герасимова [16], Е.В. Васильевой [17], В.В. Мартынова, Е.И. Филосовой, О.В. Ширяева [18], А.П. Бакановой [19], Х. Марвана [20].

Большинством исследователей предлагается анализировать рынок труда в области ИТ-сферы на основании вакансий с сайтов-агрегаторов, индивидуализировать образовательные программы и вовлекать работодателей в процесс обучения.

М. Бюссемакер и соавторы предлагают онтологический подход к формированию учебной программы на примере студентов в области химической инженерии [21]. Авторы предлагают опираться на онтологическую репрезентацию предметной области с привлечением экспертов для формирования отношений между классами онтологии. Такая структура модели позволяет для каждой специальности в рамках предметной области получать список необходимых знаний, умений и навыков, а также сферы их применения.

В работе Д. Смита и А. Али [22] рассматривается извлечение навыков программистов с сайта для вакансий Dice.com. Был разработан автоматизированный алгоритм сбора данных, в котором на основании запроса ключевого слова, идентифицирующего навык, формируются данные о дате запроса, ключевом слове и количестве вакансий, которые относятся к этому навыку. Далее в работе анализировалась популярность навыка на основании данных за последние три года. Данное исследование позволило выявить тенденции рынка труда, и на основании выявленной динамики популярности навыков рекомендует корректировать, актуализировать образовательные программы.

Д.Ю. Ильиным [23] описана система, которая собирает сведения из патентных данных, вакансий для поиска работы и поисковых запросов. Осуществлялся сбор, гомогенизация и анализ данных из пяти различных источников: HeadHunter, Indeed.com, GooglePatents, GoogleScholar и GoogleTrends. Результатом обработки являлись навыки в области программирования. Был разработан прототип сервиса, автоматически формирующий аналитическую информацию о востребованности навыков на рынке труда, и прогноз развития технологий – на основе анализа патентной активности. Сервис позволяет подбирать

наиболее актуальные для изучения навыки, навыки определяются по ключевым словам.

Возможность учета различных требований работодателей рассматривается в большинстве работ, посвященных построению модели специалиста, однако исследователи не всегда подходят к решению задачи с точки зрения автоматизации, вследствие чего уменьшается число источников, которые возможно одновременно использовать при построении модели специалиста.

Кроме того, не проводится унификация и анализ компетенций, которые требуются от молодых специалистов, что, учитывая слабоструктурированный характер исходных данных о требованиях работодателей, также указывает на необходимость разработки моделей и методов обработки открытых данных о требованиях предприятий к соискателям.

Концепция управления подготовкой ИТ-кадров

С учетом анализа существующих подходов нами предлагается концепция организации подготовки ИТ-специалистов с опорой на открытые данные о требованиях работодателей, которая учитывает особенности предметной области и использует ключевые понятия в качестве базовых единиц предложенных моделей.

Источниками таких данных являются:

1. Страницы вакансий, размещаемые компаниями на сайтах-агрегаторах.
2. Информационные веб-ресурсы и веб-энциклопедии, включающие статьи и рубрики, посвященные информационным технологиям; ресурсы с нормативной документацией (например, реестр профессиональных стандартов на сайте Министерства труда и социальной защиты Российской Федерации).
3. Учебные планы и рабочие программы дисциплин, публикуемые образовательными организациями на официальных сайтах.

Важно отметить, что перечисленные источники представляют собой коллекции веб-документов на естественном языке и могут быть охарактеризованы как слабоструктурированные данные [24], поэтому целесообразно обратиться к методам обработки естественного языка [25] для сбора и структурирования указанных данных.

Одной из методик является денотативный подход [26–29] к анализу текста на естественном языке, позволяющий извлечь содержание

текста и представить его в виде набора денотатов. Денотат – слово или словосочетание в тексте на естественном языке, которое соотносится с объектом или явлением предметной области и выражает этот смысл в указанном тексте [27]. В настоящей работе денотаты соотносятся с ключевыми понятиями, используемыми для описания и классификации компетенций в области информационных технологий. К таким понятиям относятся обозначения технологий, языков программирования, инструментов, концепций и т.д.

Денотативный подход позволяет использовать единицы, уже представленные в исходных данных в явном виде, в качестве базовых элементов моделей. Еще одним преимуществом подхода является возможность использования денотатного графа в качестве средства организации данных.

Денотатный граф – ориентированный граф, узлами которого являются денотаты текста, ребрами – смысловые связи между ними [29]. Денотатный граф может использоваться как формальное представление смысловой структуры текста или множества текстов на естественном языке. Денотатный граф, построенный для множества текстов одной предметной области, может рассматриваться как онтология [30; 31].

Онтологическая модель обеспечивает гибкость и системный подход к предметной области. Онтология позволяет восстановить недостающие логические связи [32], что позволяет на основе конечного множества ключевых понятий, указанных в текстах вакансий, описать полный профессиональный портрет специалиста, требующегося работодателю на вакантную должность. Использование денотатного графа для этих задач расширяет возможности автоматизированного построения иерархии ключевых понятий на основе большого количества данных без необходимости привлечения экспертов.

Используя указанные выше источники данных, разрабатываемая информационная система должна обеспечить их автоматизированный сбор, хранение, предварительную обработку. Обработанные данные становятся основой для комплекса математических моделей, описывающих специалиста с разных точек зрения и на разных этапах его подготовки и трудоустройства. Полученные модели позволяют формировать набор рекомендаций образовательным учреждениям для повышения эффективности подготовки специалистов.

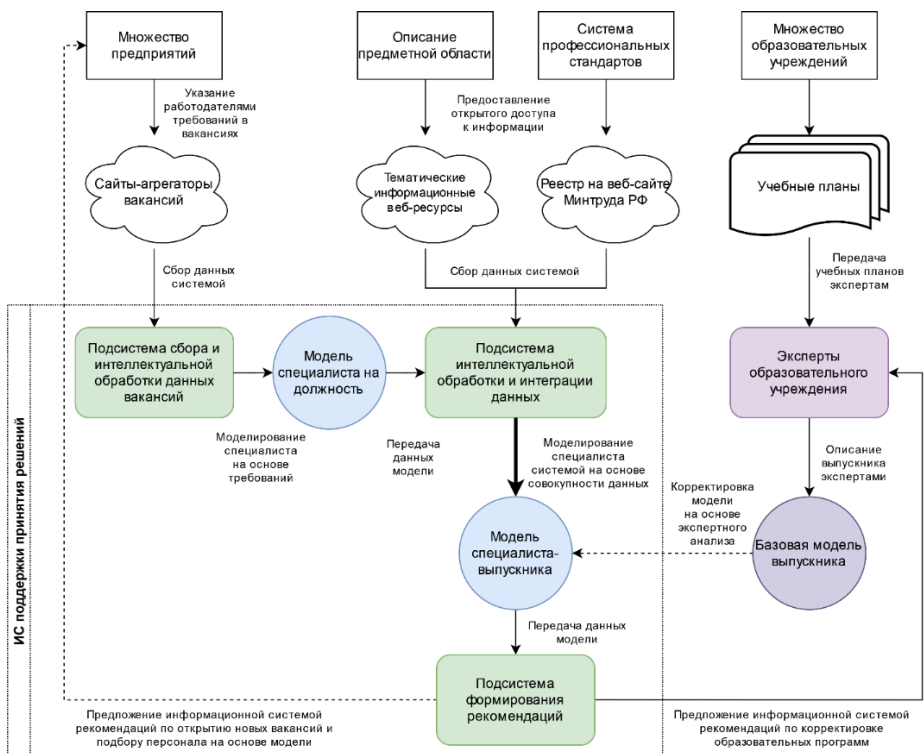


Рис. 1. Схема организации подготовки ИТ-специалистов на основе открытых данных

Схема предлагаемой концепции представлена на рис. 1 и включает в себя следующие этапы:

1. Мониторинг и сбор данных вакансий с сайтов-агрегаторов с заданной периодичностью.
2. Их последующая обработка для построения динамической модели ИТ-специалиста, описывающей ключевые понятия, необходимые для профессиональной деятельности, и удовлетворяющей требованиям работодателей.
3. Обогащение полученной модели открытыми данными о предметной области и информацией, содержащейся в нормативных документах.
4. Построение модели специалиста-выпускника, учитывающей связи между ключевыми понятиями (денотатами), путем построения денотатного графа для восстановления иерархии обогащенного множества понятий.

5. Формирование на ее основе набора рекомендаций по составлению и корректировке образовательных программ.

Модели специалиста на основе денотативного подхода

В рамках реализации концепции разработаны модели специалиста с точки зрения работодателя и востребованного специалиста-выпускника.

Модель специалиста с точки зрения работодателя представляет собой семейство частных моделей $M_1^{\text{напр}}, M_2^{\text{напр}}, \dots, M_n^{\text{напр}}$, характеризующих направления профессиональной деятельности специалиста в рамках предметной области (например, «Веб-разработка», «Разработка для мобильных устройств», «Разработка баз данных») и объединенных множеством ключевых понятий D :

$$M_{\text{пред}}^{\text{общ}} = \{M_1^{\text{напр}}, M_2^{\text{напр}}, \dots, M_n^{\text{напр}}, D\}. \quad (1)$$

Каждая частная модель $M_i^{\text{напр}}$ представляет собой нечеткое множество ключевых понятий, описывающее не только набор ключевых понятий направления профессиональной деятельности, но и величину принадлежности понятия данному направлению, что позволяет оценить значимость каждого ключевого понятия для специалиста данного направления:

$$M_i^{\text{напр}} = \left\{ \left\langle d, \mu_{M_i^{\text{напр}}}(d) \right\rangle \mid d \in D \right\}, \quad (2)$$

где D – множество ключевых понятий модели специалиста; $\mu_{M_i^{\text{напр}}}(d)$ – функция принадлежности ключевого понятия d i -му направлению профессиональной деятельности.

Направления профессиональной деятельности выявляются путем кластерного анализа корпуса вакансий, их число, в свою очередь, определяется предварительной кластеризацией множества ключевых понятий D . Кластеризация множества ключевых понятий осуществлялась методом распространения близости; для корпуса вакансий применялся алгоритм спектральной кластеризации. Качество кластеризации оценивается при помощи интегральной оценки на основе коэффициента силуэта, индекса Калински – Харабаша и индекса Дэвиса – Болдина.

Для векторного представления ключевых понятий и вакансий на естественном языке используется сглаженная метрика TF-IDF:

$$TfIdf_{\text{smooth}} = \left(\frac{f_i}{k} \right) \cdot \log \left(\left(\frac{1+n}{1+df} \right) + 1 \right),$$

где f_i – количество вхождений i -го понятия в тексте; k – общее количество понятий в тексте; n – общее количество текстов в корпусе, df – количество текстов в корпусе, в которых i -е понятие встречается минимум один раз.

Размерность вектора в случае с вектором для ключевых понятий определяется мощностью корпуса вакансий; в случае с векторным представлением вакансий – мощностью множества ключевых понятий D .

Функция принадлежности ключевого понятия $d \in D$ -му профессиональному направлению использует частотную характеристику встречаемости понятия d в исходных данных i -го кластера вакансий и рассчитывается следующим образом:

$$\mu_{M_i^{\text{напр}}}(d) = \frac{\sum_{j=1}^{|V_i|} x_{ij}}{|V_i|},$$

$$x_{ij} = \begin{cases} 1, & \text{если } d \in v_{ij} \\ 0, & \text{в противном случае} \end{cases},$$

где V_i – множество вакансий, входящих в i -й кластер; v_{ij} – текст j -й вакансии в i -м множестве вакансий; x_{ij} – переменная хотя бы единственного вхождения ключевого понятия d в текст j -й вакансию в i -м множестве вакансий.

Фрагмент автоматически построенной модели приведен на рис. 2.

И хотя модель специалиста отражает набор ключевых понятий, необходимых работодателям в тех или иных направлениях деятельности, она не учитывает связи между понятиями. Также линейная модель не включает в себя понятия компетенций, освоение которых необходимо для изучения других компетенций, но которые при этом не были включены работодателями в качестве явных требований к вакансиям (например, ключевое понятие `git` в приведенном на рис. 2 фрагменте модели требуется во всех направлениях профессиональной деятельно-

сти, но ни одно из них не содержит требования «системы контроля версий», которое является родовым понятием для git и необходимо для его изучения).

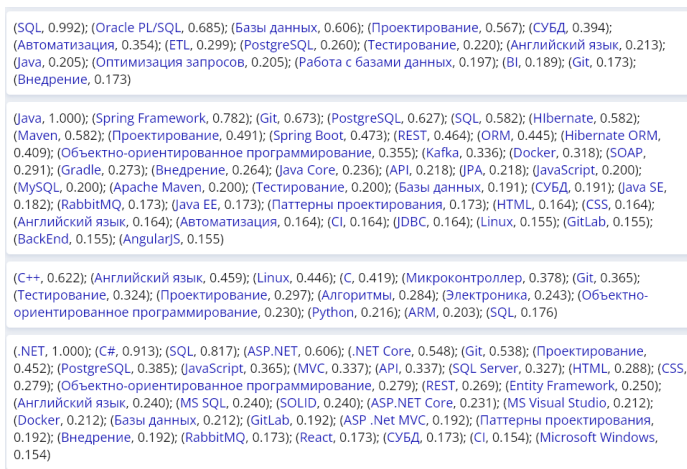


Рис. 2. Фрагмент линейной модели ИТ-специалиста на вакантную должность для предприятий Пермского края

Модель специалиста-выпускника расширяет и обогащает модель специалиста на должность работодателя, сохраняя при этом общую основу (множество ключевых понятий), что позволяет напрямую оценивать сходство моделей, в том числе после возможной экспертной корректировки модели специалиста-выпускника.

Модель специалиста-выпускника включает в себя множество частных моделей $G_i^{\text{напр}}$ и множество всех ключевых понятий S , при этом $D \subset S$, где D – множество ключевых понятий модели специалиста на должность, как указано в (1).

$$M_{\text{вып}} = \{G_1^{\text{напр}}, G_2^{\text{напр}}, \dots, G_n^{\text{напр}}, S\}.$$

Каждая из частных моделей $G_i^{\text{напр}}$ является ориентированным графом:

$$G_i^{\text{напр}} = (V, E),$$

где V – множество вершин графа, E – множество дуг графа. Количество частных моделей $G_i^{\text{напр}}$ совпадает с количеством частных моделей,

описанных в формуле (2), для каждой модели $M_i^{\text{напр}}$ производится обобщение множества ключевых понятий и строится граф.

Итоговые графы модели специалиста-выпускника сохраняются в формате DOT, фрагмент графа приведен на рис. 3.

Для формирования рекомендаций для составления и корректировки образовательных программ разработан алгоритм, принимающий на вход целевое понятие $s_{\text{цел}}$ и модель специалиста-выпускника $M_{\text{вып}}$. Разработанный алгоритм приведен на рис. 4.

На выходе алгоритма – линейно упорядоченное множество рекомендованных к изучению понятий $S_{\text{рек}}$ с расчетным количеством часов. Полученное множество передается экспертам образовательной организации и лицу, принимающему решения, как набор рекомендаций относительно составления и корректировки образовательных программ.

Указанные рекомендации не только отражают требования работодателей, но также указывают на горизонтальную и вертикальную интеграцию ключевых понятий в рамках предметной области, что дает возможность эффективного внесения изменений не только в рамках отдельных дисциплин, но также и учебных планов в целом.

```

"perl" -> "unix" [weight=10 penwidth=1.00 label=1.0]
"windows" -> "java" [weight=10 penwidth=1.00 label=1.0]
"asp net" -> "mvc" [weight=10 penwidth=1.00 label=1.0]
"json" -> "javascript" [weight=97 penwidth=3.43 label=9.7]
"go" -> "sql" [weight=10 penwidth=1.00 label=1.0]
"net" -> "java" [weight=10 penwidth=1.00 label=1.0]
"svn" -> "sqlite" [weight=10 penwidth=1.00 label=1.0]
"css" -> "xml" [weight=79 penwidth=2.98 label=7.9]
"html" -> "http" [weight=66 penwidth=2.65 label=6.6]
"php" -> "yii" [weight=10 penwidth=1.00 label=1.0]
"vue js" -> "redux" [weight=79 penwidth=2.98 label=7.9]
"svn" -> "python" [weight=10 penwidth=1.00 label=1.0]
"python" -> "java" [weight=10 penwidth=1.00 label=1.0]
"objective-c" -> "c" [weight=10 penwidth=1.00 label=1.0]
"linux" -> "embedded" [weight=10 penwidth=1.00 label=1.0]
"ajax" -> "jquery" [weight=10 penwidth=1.00 label=1.0]
"delphi" -> "linux" [weight=10 penwidth=1.00 label=1.0]
"oon" -> "java" [weight=36 penwidth=1.90 label=3.6]
"react" -> "html" [weight=135 penwidth=4.38 label=13.5]
"net" -> "powershell" [weight=10 penwidth=1.00 label=1.0]
"svn" -> "git" [weight=80 penwidth=3.00 label=8.0]
"api" -> "windows" [weight=10 penwidth=1.00 label=1.0]
"redux" -> "macos" [weight=59 penwidth=2.48 label=5.9]
"objective-c" -> "ios" [weight=10 penwidth=1.00 label=1.0]

```

Рис. 3. Фрагмент графа модели специалиста-выпускника, описанного на языке DOT

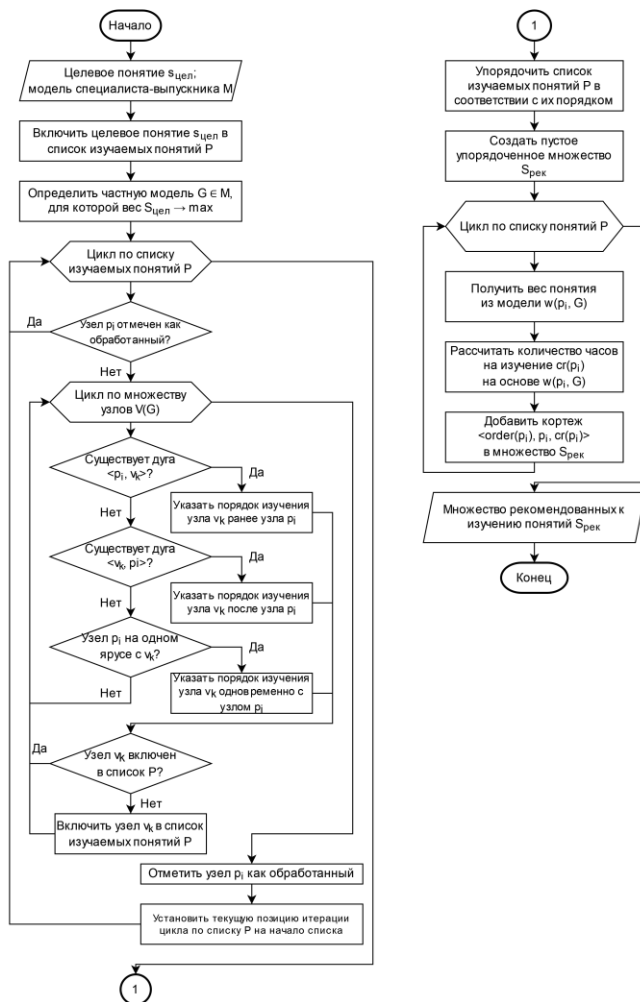


Рис. 4. Алгоритм формирования рекомендаций по включению в образовательные программы востребованных ключевых понятий

Программная реализация моделей и алгоритмов

Разработанные модели и алгоритмы реализованы в информационной системе поддержки принятия решений.

Реализация серверной части осуществлена на языке программирования Python. Она включает в себя модули, которые обеспечивают:

- запуск сервера веб-приложения;
- обработку HTTP-запросов со стороны клиента (frontend)
- формирование и вывод рекомендаций по составлению и корректировке образовательных программ;

- предоставление интерфейса обмена данных с моделями для сервера веб-приложения;

- запуск моделирования специалиста по заданным входным параметрам

- организацию ввода/вывода для базы данных

- анализ разметки документов;

- сегментацию документов;

- морфологический анализ документов;

- генерацию коллекции лемм по заданным критериям;

- фильтрацию ключевых понятий из множества лемм коллекции

- ежемесячный сбор данных о вакансиях с сайтов-агрегаторов;

- организацию и сохранение полученных данных в формате

JSON

- векторизацию ключевых понятий;

- выявление числа стабильных групп ключевых понятий при помощи кластеризации алгоритмом распространения близости

- векторизацию множества вакансий;

- группировку вакансий алгоритмом спектральной кластеризации;

- формирование нечетких множеств ключевых понятий

- сбор и кэширование текстов о предметной области с информационных веб-ресурсов;

- обогащение заданного множества ключевых понятий понятиями, извлеченными из текстов предметной области;

- построение взвешенного денотатного графа обогащенного множества понятий;

- установление иерархии узлов графа алгоритмом DOT.

Логика клиентской части написана на языке программирования JavaScript. Инструменты визуализации графов используют C++-библиотеки в формате модулей WebAssembly. Пользовательский интерфейс реализован шаблонами Jinja2 и средствами веб-разметки – HTML и CSS.

Данные о вакансиях включают в себя 495 024 текста вакансий, опубликованных в период с февраля 2020 г. Пользовательский интерфейс системы с фрагментом модели специалиста с точки зрения работодателей Пермского края на основе данных в период с ноября 2021 г. по апрель 2022 г. приведен на рис. 5.

Для взаимодействия с моделью специалиста-выпускника система предоставляет интерактивную визуализацию денотатного графа с возможностью анализа связей между денотатами. Фрагмент визуального представления исходного денотатного графа в интерфейсе системы приведен на рис. 6.

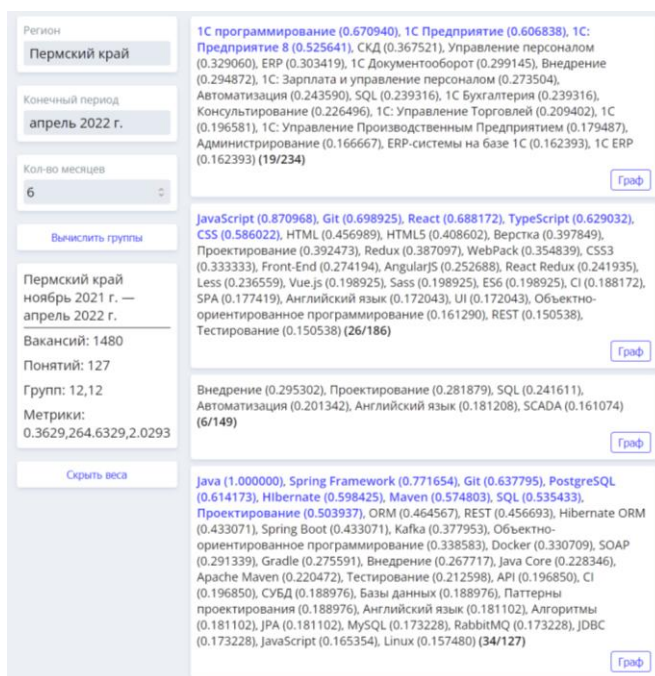


Рис. 5. Фрагмент модели специалиста на должность для Пермского края на основе данных за период с ноября 2021 г. по апрель 2022 г.

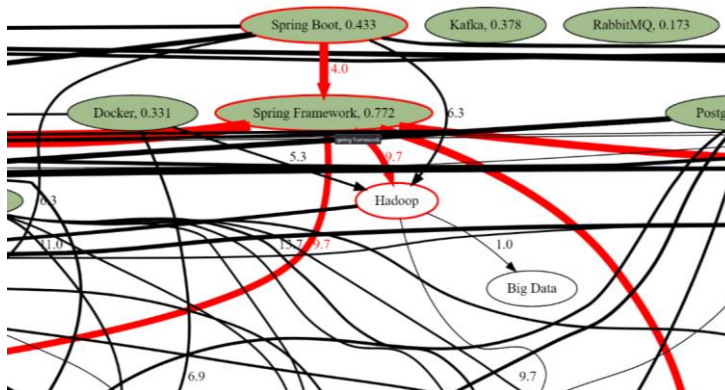


Рис. 6. Фрагмент исходного обогащенного денотатного графа, регион Пермский край, период с ноября 2021 г. по апрель 2022 г.

Апробация результатов

В рамках тестирования разработанной интеллектуальной системы проведен эксперимент с целью оценки ее эффективности. В эксперименте по проверке соответствия их компетенций требованиям работодателей приняли участие десять студентов Пермского национального исследовательского политехнического университета после окончания III курса направлений «Программная инженерия» и «Информатика и вычислительная техника».

Для оценки данного соответствия со стороны работодателей в эксперименте участвовало четыре эксперта из ООО «Датабриз» (г. Пермь) и четыре эксперта из ООО «Иннфокус» (г. Пермь). Компания «Датабриз» является разработчиком систем бизнес-аналитики с клиентскими приложениями для персональных компьютеров и мобильных устройств под управлением ОС Android и iOS. Компания «Иннфокус» специализируется на разработке математического и программного обеспечения аддитивных технологий. Обе компании заинтересованы в высокой компетенции разработчиков ПО, которые владеют стеком технологий, связанных с языком программирования Java.

Экспертами проведены стандартизированные собеседования на квалификацию стажера-разработчика с каждым из студентов, участвующих в эксперименте. Экспертами заполнены опросники с указанием оценок каждого студента по шкале от «1» (низкий уровень) до «5» (наивысший уровень) по следующим критериям:

- уровень общей подготовки – наличие у студента компетенций по основам алгоритмизации и программирования, включая парадигмы программирования; теории алгоритмов и структурам данных;
- уровень владения стеком технологий – наличие у студента компетенций по требуемым компании технологиям, включая языки программирования, фреймворки, прикладное программное обеспечение;
- уровень технического кругозора – наличие у студента понимания взаимосвязей между различными технологиями; умение подбора альтернатив;
- вероятность рекомендации к трудоустройству – оценка экспертом вероятности рекомендации соискателя к трудоустройству на позицию стажера-разработчика в компании эксперта.

Результаты первичных собеседований приведены в табл. 1, где указана средняя оценка экспертов по заданному критерию.

Таблица 1

Результаты первичных собеседований, средняя оценка экспертов, *авторские результаты*

№ участника	Общий уровень подготовки	Владение стеком технологий	Уровень технического кругозора	Вероятность рекомендации к трудоустройству
1	3,875	2,5	3,25	2
2	4,375	2,375	2,125	3,25
3	3,625	1,625	3,375	2,5
4	4,125	2,625	2,375	3,125
5	4,125	2,5	1,5	2,375
6	3,125	1,5	1,625	1,375
7	4,5	2,625	2,375	3,5
8	3,875	2,625	2,5	3
9	4	1,5	2	1,5
10	3,75	2,5	2,625	2,75
Среднее	3,9375	2,2375	2,375	2,5375

Результаты первичных собеседований показали, что, хотя студенты обладают достаточными компетенциями в области алгоритмизации и программирования (ср. балл – 3,9375), отсутствие компетенций, связанных с требуемыми технологиями, и понимания их взаимосвязи (ср. оценка владения стеком технологий – 2,24; оценка технического кругозора – 2,375) сильно снижают вероятность рекомендации к трудоустройству (средний балл – 2,54).

После первичных собеседований студенты прошли краткосрочную подготовку по программе интенсива, составленной на основе денотатного графа для направления «Java-разработка» для Пермского края на основе данных за период с ноября 2021 г. по апрель 2022 г. (фрагмент указанного графа представлен на рис. 6). По итогам двухнедельного интенсива студенты заново прошли собеседование, при этом эксперты задавали дополнительные вопросы соискателям и заново заполняли опросники.

Результаты повторных собеседований приведены в табл. 2.

Повторные собеседования показали улучшение показателей по всем критериям (средний общий уровень подготовки вырос на 9 %; владение стеком технологий – на 60 %; технический кругозор – на 46 %). Особенно важно отметить, что вероятность рекомендации к трудоустройству выросла на 40 %.

Таблица 2

Результаты повторных собеседований, средняя оценка экспертов, *авторские результаты*

№ участника	Общий уровень подготовки	Владение стеком технологий	Уровень технического кругозора	Вероятность рекомендации к трудоустройству
1	4,375	3,625	4,375	3,625
2	4,875	3,625	3,375	3,875
3	4,375	3,375	3,875	3,5
4	4,5	3,625	3,5	3,75
5	4,375	3,625	3,25	3,125
6	3,5	3,125	2,75	3
7	5	4,125	3,375	4,125
8	4,375	3,875	3,5	4
9	4,25	3,125	2,875	2,75
10	4	3,625	3,75	4
Среднее	4,3625	3,575	3,4625	3,575

Таким образом, разработанная система сокращает разрыв между компетенциями, которые приобретают соискатели в образовательных организациях, и компетенциями, которые требуются работодателям, что подтверждает ее эффективность.

Заключение

В рамках исследования обозначена растущая проблема несоответствия компетенций выпускников ИТ-специальностей образовательных организаций и требований работодателей. Произведен анализ существующих подходов к ее решению. Выявлены перспективные направления моделирования специалиста.

Предложена концепция организации подготовки ИТ-специалистов на основе открытых слабоструктурированных данных. Описан подход, основанный на обработке данных на естественном языке и денотативной аналитике. Разработаны модели специалиста с точки зрения работодателя и специалиста-выпускника, а также алгоритм формирования рекомендаций по составлению и корректировке образовательных программ.

Разработано программное обеспечение системы поддержки принятия решений, реализующее указанные модели и алгоритмы. Проведена экспериментальная проверка с привлечением экспертов из ИТ-индустрии. По итогам эксперимента получены положительные результаты.

Разработанная система может использоваться для составления и корректировки рабочих программ дисциплин, программ профессиональной переподготовки, программ дополнительного образования в области направлений, связанных с информационными технологиями.

Список литературы

1. План мероприятий по направлению «Кадры и образование» [Электронный ресурс] / Национальная программа «Цифровая экономика Российской Федерации». – URL: <http://static.government.ru/media/files/-k87YsCABuiyuLAjcWDFILEh6itAirUX0.pdf> (дата обращения: 20.03.2021).
2. ИТ: обзор рынка вакансий и топ-15 специальностей [Электронный ресурс] / Группа компаний HeadHunter. – URL: <https://perm.hh.ru/article/24562> (дата обращения: 20.02.2021).
3. Земнухова Л.В. Ит-работники на рынке труда // Социология науки и технологий. – 2013. – Т. 4. – № 2. – С. 77–90.
4. ICT specialists in employment [Электронный ресурс] / Eurostat Statistic Explained. – URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/-ICT_specialists_in_employment#Number_of_ICT_specialists (дата обращения: 10.02.2021).
5. Образование в цифрах [Электронный ресурс] / НИУ ВШЭ. – URL: <https://www.hse.ru/primarydata/os2018> (дата обращения: 20.04.2021).
6. Попова Т.Н. Структурный дисбаланс системы занятости в регионе // Современная экономика: проблемы, тенденции, перспективы. – 2011. – № 5. – С. 1–6.
7. Гимпельсон В. Дефицит квалификации и навыков на рынке труда (недостаток предложения, ограничения спроса или ложные сигналы работодателей?) // Вопросы экономики. – 2004. – № 3. – С. 76–94.
8. Калевко В.В., Лагерев Д.Г., Подвесовский А.Г. Управление образовательной программой вузов в контексте подготовки конкурентоспособных разработчиков программного обеспечения // Современные информационные технологии и ИТ-образование. – 2018. – Т. 14. – № 4. – С. 803–814.
9. Булах К.В., Бурцева Е.Т. Модель профессиональной подготовки специалиста с учетом требований современного рынка труда // Вестник Адыгейского государственного университета. – 2018. – № 3(223). – С. 39–44.
10. Харитоновна И.Ю., Сергеев С.Ф. Роль рамок квалификаций/компетенций в формировании образовательных программ ИКТ-направлений подготовки высшего профессионального образования // Вестник университета. – 2014. – № 19. – С. 87–96.
11. Галиев Р.М. Автоматизированная система для разработки профессиональных образовательных программ // Современные образовательные

технологии в мировом учебно-воспитательном пространстве. – 2016. – № 4. – С. 24–29.

12. Курзаева Л.В. Компетентностный подход в подготовке ИТ-специалистов: обзор в поисках точки опоры // Современные информационные технологии и ИТ-образование. – 2012. – № 8. – С. 208–214.

13. Информационная система оценки образовательных программ на основе требований работодателей / А.А. Захарова, А.Н. Лазарева, О.Ю. Зорина, В.В. Останин // Современные проблемы науки и образования. – 2015. – № 2 (1). – С. 136–143.

14. Воронин А.А., Васильченко А.А., Зенович А.В. Анализ содержания университетской подготовки ИТ-специалистов на основе федеральных государственных образовательных стандартов с позиций профессиональных ИТ-стандартов // *Artium Magister*. – 2013. – № 14. – С. 50–61.

15. Лебедева Т.Ф., Муравьев С.А. Профессиональные стандарты в сфере информационных технологий как руководство к действию // Профессиональное образование в России и за рубежом. – 2016. – № 4 (24). – С. 92–97.

16. Гузаиров М.Б., Герасимова И.Б. Системный подход к оценке качества образовательного процесса // Вестник Уфимского государственного авиационного технического университета. – 2007. – № 5(9). – С. 29–33.

17. Васильева Е.В. Система квалификации как инструмент соответствия интересов сфер труда и образования // Вестник университета. – 2014. – № 19. – С. 10–16.

18. Мартынов В.В., Филосова Е.И., Ширяев О.В. Организация подготовки и информационная поддержка реализации динамических образовательных программ, учитывающих требования работодателя // Управление экономикой: методы, модели, технологии. – 2014. – С. 110–113.

19. Bakanova A. The use of Ontologies in the Development of a Mobile E-Learning Application in the Process of Staff Adaptation // *International Journal of Recent Technology and Engineering*. – 2019. – № 2S10 (8). – С. 780–789.

20. Ali A., Marwan H. Exploring career management competencies in work-based learning (WBL) implementation // *J. Tech. Educ. Train.* – 2019. – № 11. – С. 159–166.

21. An ontological approach to chemical engineering curriculum development / M. Bussemaker, N. Trokanas, F. Cecelja // *Computers and Chemical Engineering*. – 2017. – № 106. – С. 927–941.

22. Smith D., Ali A. Analyzing computer programming job trend using web data mining // *Issues in Informing Science and Information Technology*. – 2014. – № 11. – С. 203–214.

23. Информационно-аналитический сервис формирования актуальных профессиональных компетенций на основе патентного анализа технологий и выделения профессиональных навыков в вакансиях работодателей /

Д.Ю. Ильин, Е.В. Никульчев, Г.Г. Бубнов, Е.О. Матешук // Прикаспийский журнал: управление и высокие технологии. – 2017. – № 2(38). – С. 71–88.

24. Яруллин Д.В. Извлечение слабоструктурированных данных при помощи N-мерных векторов (на примере рабочих программ дисциплин ВО) // Естественные научные методы в цифровой гуманитарной среде: материалы Всерос. науч. конф. с междунар. участием. – 2018. – С. 120–123.

25. Jurafsky D., James H.M. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* [Электронный ресурс]. – URL: <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf> (дата обращения: 10.03.2022).

26. Новиков А.И., Нестерова Н.М. Реферативный перевод научно-технических текстов. – М., 1991. – 148 с.

27. Новиков А.И. Семантика текста и ее формализация. – М.: 1983. – 216 с.

28. Новиков А.И. Текст и «контртекст»: две стороны процесса понимания // Вопросы психолингвистики. – 2003. – № 1. – С. 64–76.

29. Новиков А.И. Текст и его смысловые доминанты. – М.: Институт языкознания РАН, 2007. – 224 с.

30. Курушин Д.С., Леонов Е.Р., Соболева О.В. О возможном подходе к автоматическому построению денотатного графа гипертекста // Информационная структура текста. – 2018. – С. 113–118.

31. Курушин Д.С., Соболева О.В., Вяткин Д.С. Денотативная модель основного содержания текста в лингвистическом обеспечении робототехнических исследований // Информационная структура текста. – 2018. – С. 104–112.

32. Стенников В.А., Барахтенко Е.А., Соколов Д.В. Применение онтологий при реализации концепции модельно-управляемой разработки программного обеспечения для проектирования теплоснабжающих систем // Онтология проектирования. – 2014. – № 4(14). – С. 54–68.

References

1. Plan meroprijatij po napravleniju «Kadry i obrazovanie» [Action plan for human resources and education]. Available at: <http://static.government.ru/media/files/k87YsCABuiyuLAjcwDFILEh6itAirUX0.pdf> (Accessed 20 March 2021).

2. IT: obzor rynka vakansij i top-15 special'nostej [Overview of the job market and the top-15 qualifications]. Available at: <https://perm.hh.ru/article/24562> (Accessed 20 February 2022).

3. Zemnukhova L.V. IT-workers at the labour market. *Sociology of science and technology*, 2013, vol. 4, no. 2, pp. 77–90.

4. ICT specialists in employment. Available at: <https://ec.europa.eu/eurostat/statistics-explained/index.php/>

ICT_specialists_in_employment#Number_of_ICT_specialists (Accessed 10 February 2022).

5. *Образование в цифрах* [Education in numbers]. Available at: <https://www.hse.ru/primarydata/oc2018> (Accessed 20 April 2021).

6. Popova T.N. Strukturnyj disbalans sistemy zanjatosti v regione [Structural imbalances in the employment system in the region]. *Sovremennaja jekonomika: problemy, tendencii, perspektivy*, 2011, no. 5, pp. 1–6.

7. Gimpelson V. Shortage of Skills in the Labor Market: Limited Supply, Lack of Demand, or False Signals from Employers? *Voprosy Ekonomiki*, 2004, no. 3, pp. 76–94.

8. Kaleyko V.V., Lagerev D.G., Podvesovskii A.G. Educational program management of universities in the context of training competitive software developers. *IT education: methodology, methodological support*, 2018, vol. 14, no. 4, pp. 803–814.

9. Bulakh K.V., Burtseva E.T. Model' professional'noi podgotovki spetsialista s uchetom trebovanii sovremennogo rynka truda [The model of specialist training considering the requirements of the modern labor market]. *Vestnik Adygeiskogo gosudarstvennogo universiteta*, 2018, no. 3 (223), pp. 39–44.

10. Kharitonova I.Iu., Sergeev S.F. The role of the qualifications/competences framework in developing of educational programs for ict-related training areas of higher education. *Vestnik universiteta*, 2014, no.19, pp. 87–96.

11. Galiev R.M. Avtomatizirovannaja sistema dlja razrabotki professional'nyh obrazovatel'nyh programm [Automated system for the development of vocational training programmes]. *Sovremennye obrazovatel'nye tehnologii v mirovom uchebno-vospitatel'nom prostranstve*, 2016, no. 4, pp. 24–29.

12. Kurzaeva L.V. Kompetentnostnyj podhod v podgotovke IT-specialistov: obzor v poiskah točki opory [Competency-based approach in IT training: a review in search of a reference point]. *Sovremennye informacionnye tehnologii i IT-obrazovanie*, 2012, no. 8, pp. 208–214.

13. Zakharova A.A., Lazareva A.N., Zorina O. Iu, Ostanin V.V. Informacionnaja sistema ocenki obrazovatel'nyh programm na osnove trebovanij rabotodatelej [Information system for evaluating educational programmes based on employers' requirements]. *Sovremennye problemy nauki i obrazovanija*, 2015, no. 2 (1), pp. 136–143.

14. Voronin A.A., Vasilchenko A.A., Zenovich A.V. The analysis of IT specialists' preparation at Russian universities on the ground of the federal state educational standards from the perspective of the professional IT standards. *Artium Magister*, 2013, no. 14, pp. 50–61.

15. Lebedeva T.F., Muravyev S.A. Professional standards in the IT sector as a guide to action. *Contemporary tendencies in professional education development*, 2016, no. 4 (24), pp. 92–97.

16. Guzairov M.B., Gerasimova I.B. Sistemnyj podhod k ocenke kachestva obrazovatel'nogo processa. *Vestnik Ufimskogo gosudarstvennogo aviacionnogo tehničeskogo universiteta*, 2007, no. 5 (9), pp. 29–33.

17. Vasileva E.V. Qualification system as a tool for alignment of the interests of the spheres of labour and education. *Vestnik universiteta*, 2014, no. 19, pp. 10–16.

18. Martynov V.V., Filosova E.I., Shiryayev O.V. Organizacija podgotovki i informacionnaja podderzhka realizacii dinamičeskikh obrazovatel'nyh programm, uchityvajushhih trebovanija rabotodatelja [Organisation of training and information support for the implementation of dynamic educational programmes tailored to the requirements of the employer]. *Upravlenie jekonomikoj: metody, modeli, tehnologii*, 2014, pp. 110–113.

19. Bakanova A. The use of Ontologies in the Development of a Mobile E-Learning Application in the Process of Staff Adaptation. *International Journal of Recent Technology and Engineering*, 2019, no. 2S10 (8), pp. 780–789.

20. Ali A., Marwan H. Exploring career management competencies in work-based learning (WBL) implementation. *J. Tech. Educ. Train*, 2019, no. 11, pp. 159–166.

21. Bussemaker M., Trokanas N., Cecelja F. An ontological approach to chemical engineering curriculum development. *Computers and Chemical Engineering*, 2017, no. 106, pp. 927–941.

22. Smith D., Ali A. Analyzing computer programming job trend using web data mining. *Issues in Informing Science and Information Technology*, 2014, no. 11, pp. 203–214.

23. Ilin D.Yu., Nikulchev Y.V., Bubnov G.G., Mateshuk Y.O. Information analysis service to identify current professional competencies based on patent analysis of the technology market and required skills in the vacancies of employers. *CASPIAN JOURNAL: Control and High Technologies*, 2017, no. 2 (38), pp. 71–88.

24. Yarullin D.V. Extraction of semi-structured data using n-measured vectors (on the example of higher education syllabuses). *Natural Science Methods in the Digital Humanitarian Environment*, 2018, pp. 120–123.

25. Jurafsky D., James H.M. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Available at: <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf> (Accessed 10 March 2022).

26. Novikov A.I., Nesterova N.M. Referativnyj perevod nauchno-tehničeskikh tekstov [Abstract translation of scientific and technical texts]. Moscow, 1991, 148 p.

27. Novikov A.I. Semantika teksta i ee formalizacija [Text semantics and its formalisation]. Moscow, 1983, 216 p.

28. Novikov A.I. Tekst i «kontrtekst»: dve storony processa ponimaniya [Text and "counter-text": two sides of the comprehension process]. *Voprosy psicholingvistiki*, 2003, no. 1, pp. 64–76.

29. Novikov A.I. Tekst i ego smyslovye dominanty [Text and its semantic dominants]. Moscow: Institut jazykoznanija RAN, 2007, 224 p.

30. Kurushin D.S., Leonov E.R., Soboleva O.V. On a possible approach to the automatic generation of denotative graph of a hypertext. *Informacionnaja struktura teksta*, 2018, pp. 113–118.

31. Kurushin D.S., Soboleva O.V., Viatkin D.S. Denotative content model as a mean of linguistic support of robotics research. *Informacionnaja struktura teksta*, 2018, pp. 104–112.

32. Stennikov V.A., Barakhtenko E.A., Sokolov D.V. Usage of ontologies in the implementation of the concept of model-driven engineering for the design of heat supply systems. *Ontologija proektirovanija*, 2014, no. 4 (14), pp. 54–68.

Сведения об авторе

Яруллин Денис Владимирович (Пермь, Россия) – старший преподаватель, кафедра «Информационные технологии и автоматизированные системы», Пермский национальный исследовательских политехнический университет (614990, Пермь, Комсомольский проспект, 29, e-mail: d.v.yarullin@ya.ru)

About the author

Denis V. Yarullin (Perm, Russia) – assistant professor, Department of Information Technologies and Computer-Aided Systems, Perm National Research Polytechnic University (29 Komsomolsky Ave., Perm, 614990, e-mail: d.v.yarullin@ya.ru)

Библиографическое описание статьи согласно ГОСТ Р 7.0.100–2018:

Яруллин, Д.В. Интеллектуальная система управления подготовкой ИТ специалистов на основе денотативной аналитики / Д. В. Яруллин. – текст : непосредственный. – DOI: 10.15593/2499-9873/2022.3.08 // Прикладная математика и вопросы управления / Applied Mathematics and Control Sciences. – 2022. – № 3. – С. 141–164.

Цитирование статьи в изданиях РИНЦ:

Яруллин, Д. В. Интеллектуальная система управления подготовкой ИТ специалистов на основе денотативной аналитики / Д. В. Яруллин // Прикладная математика и вопросы управления. – 2022. – № 3. – С. 141–164. – DOI: 10.15593/2499-9873/2022.3.08

Цитирование статьи в references и международных изданиях

Cite this article as:

Yarullin D.V. Intelligent control system for IT specialists training based on denotative analytics. *Applied Mathematics and Control Sciences*, 2022, no. 3, pp. 141–164. DOI: 10.15593/2499-9873/2022.3.08 (*in Russian*)

Финансирование. Исследование не имело спонсорской поддержки.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Вклад автора. 100 %.

Поступила: 10.04.2022

Одобрена: 20.04.2022

Принята к публикации: 01.09.2022