

<https://doi.org/10.25559/SITITO.020.202403.782-790>
УДК 004.891.2

Оригинальная статья

Тематическое наполнение образовательных программ в сфере ИТ на основе оценки перспективности технологий по патентным данным

Е. В. Никульчев*, Я. С. Грюкан, Д. Ю. Ильин

ФГБОУ ВО «МИРЭА – Российский технологический университет», г. Москва, Российской Федерации

Адрес: 119452, Российской федерации, г. Москва, пр. Вернадского, д. 78

*nikulchev@mail.ru

Аннотация

Введение. Вопросы тематического наполнения и актуализации образовательных программ и дисциплин стоят во всех сферах обучения. Особенно часто изменения происходят в сфере ИТ, что определяется скоростью развития технологической базы. Тематическое наполнение образовательной программы определяется мнением профессионального экспертного сообщества и советами работодателей, вместе с тем желательно иметь численные оценки различных тем для принятия обоснованных решений о содержании программы или курса. В настоящее время значительное количество исследований посвящено анализу перспектив развития технологий и прогнозированию востребованности профессиональных навыков на основе патентного анализа.

Материалы и методы. В статье предложена методика анализа патентных данных, в заданной сфере ИТ, на основе динамики изменения частоты упоминания словосочетаний в патентах. Предполагается, что каждое словосочетание характеризует определенную тему или технологию, в динамике которой определено три варианта состояния в последние рассматриваемые годы – стабильное состояние, устойчивый рост, устойчивое падение.

Результаты исследования. Предложены численные параметры, характеризующие перспективность технологий: среднее количество упоминания в патентах на участках стабильности и скорость изменения количества упоминаний. Для примера рассмотрена область «Цифровой маркетинг», выбран набор словосочетаний, претендующих на включение в качестве тем в план образовательной программы. Проведен патентный анализ по словосочетаниям по патентной базе 2000-2021 годов, в результате которого построены временные ряды, характеризующие частоту упоминания словосочетаний в текстах патентов.

Обсуждение и заключение. На основе разработанной методики сформирован комплекс численных оценок для тем, и обоснованно сформировано тематическое наполнение образовательной программы по данному направлению.

Ключевые слова: образовательные программы, тематическое наполнение, анализ патентов, прогнозная аналитика, цифровой маркетинг

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Для цитирования: Никульчев Е. В., Грюкан Я. С., Ильин Д. Ю. Тематическое наполнение образовательных программ в сфере ИТ на основе оценки перспективности технологий по патентным данным // Современные информационные технологии и ИТ-образование. 2024. Т. 20, № 3. С. 782-790. <https://doi.org/10.25559/SITITO.020.202403.782-790>

© Никульчев Е. В., Грюкан Я. С., Ильин Д. Ю., 2024



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



Content Planning for Educational Programs in the IT Field Using Estimation of the Prospects of Technologies Based on Patent Data

E. V. Nikulchev*, Y. S. Gryukan, D. Yu. Ilin

MIREA – Russian Technological University, Moscow, Russian Federation

Address: 78 Vernadsky Ave., Moscow 119454, Russian Federation

*nikulchev@mail.ru

Abstract

Introduction. Problems of content planning for educational programs and disciplines, as well as ensuring their relevance, arise in all areas of education. Changes in the field of IT education occur frequently, which is determined by the speed of development of the technological base. The thematic content of the educational program is commonly determined by the opinion of the professional expert community and advice from employers, however, it is desirable to have numerical estimates of various topics in order to make well-founded decisions about the content of the program or course. Currently, a significant amount of research is devoted to analyzing the prospects for technology development and forecasting the demand for professional skills based on patent analysis.

Materials and Methods. The article proposes a procedure for analyzing patent data in a given IT field, based on the dynamics of changes in the frequency of mention of key phrases in patents. It is assumed that each key phrase reflects a specific topic or technology, the dynamics of which in the last years under study may have one of three state options – stable state, stable growth, steady decline.

Results. Numerical parameters have been proposed that characterize the prospects of technologies: the mean number of mentions in patents in last segment of stability and the rate of change in the number of mentions. As an example, the area of "Digital Marketing" is considered. A set of key phrases has been selected that are eligible for inclusion into the educational program as topics. A patent analysis of key phrases was carried out based on patent database data for the period from 2000 to 2021. As a result, time series were formed that show the frequency of mention of key phrases in patent texts.

Discussion and Conclusion. Based on the developed procedure, a set of numerical estimates for the topics was formed. As a result, a well-founded thematic content of the educational program in this area has been formed.

Keywords: educational programs, content planning, patent analysis, predictive analysis, digital marketing

Conflict of interests: The authors declares no conflict of interest.

For citation: Nikulchev E.V., Gryukan Y.S., Ilin D.Yu. Content Planning for Educational Programs in the IT Field Using Estimation of the Prospects of Technologies Based on Patent Data. *Modern Information Technologies and IT-Education.* 2024;20(3):782-790. <https://doi.org/10.25559/SITITO.020.202403.782-790>



1. Введение

Спектр образовательных услуг интенсивно расширяется, создавая условия для постоянного совершенствования учебных программ и индивидуальных траекторий обучения [1, 2]. Благодаря развитию и унификации информационных технологий и сетей, базирующиеся на них общемировые новые научные, технологические, общественные и гуманитарные знания становятся частью современного мировоззрения, достаточно быстро становясь доступным для преподавания. Курс классической высшей математики и математического анализа не претерпел существенного изменения за 150 лет, в то время как прикладные дисциплины изменяются стремительно. В сфере интересов авторов, например, использование облачных технологий превратилось всего лишь через 10 лет от недоуменного непонимания аудитории о предмете до довольно рядовой дисциплины для студентов с доступным для выполнения набором практических занятий без специализированных серверов. По мнению авторов, наблюдается тенденция во взаимосвязи

в цепочке наука – технология – навык – процесс обучения на выку. Научные знания накапливаются, в результате формируется набор специализированных терминов, далее этот набор терминов, обозначающий новые научные знания или подходы, переходит на технологический уровень (который можно оценить по объему патентов в соответствии с методикой этой статьи), достигнув определенного насыщения технологий, они начинают внедряться компаниями как лидерами, так и инновационными группами. Это формирует запрос к рынку труда о специалистах, имеющих требуемые навыки, которые соответствуют ранее определенным терминам [3, 4]. Наблюдая потребность рынка в соответствующих специалистах образовательные учреждения отзываются введением или изменением дисциплин с целью формирования новых востребованных навыков (рис. 1). На рисунке показано, что каждому этапу найдено соответствующее измеримое множество, количество элементов которого можно взять за характеристику, при этом подсчет предлагается осуществлять на основе некоторых ключевых слов.



Рис. 1. Процесс разработки новых образовательных программ

Fig. 1. The process of developing new educational programs

Источник: здесь и далее в статье все таблицы и рисунки составлены авторами.

Source: Hereinafter in this article all tables and figures were made by the authors.

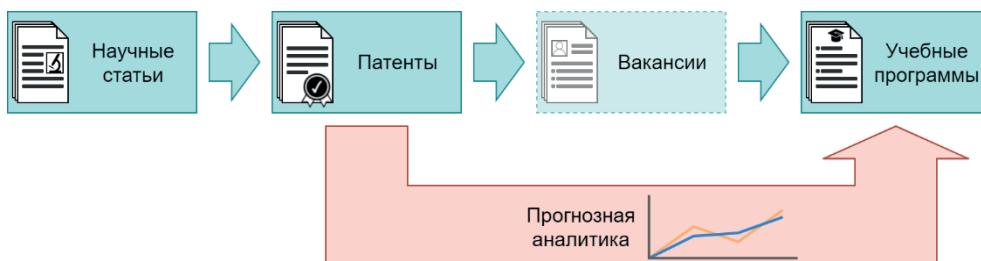


Рис. 2. Процесс разработки новых образовательных программ с применением прогнозной аналитики

Fig. 2. The process of developing new educational programs using predictive analytics

Этот процесс для разных областей разный: самые быстрые изменения происходят в ИТ-сфере, сильно медленней в медицине, поскольку требуется пройти путь от научной гипотезы к экспериментально подтвержденному результату и, в конечном итоге, к массовому распространению технологии в форме лечебных протоколов (как вида операционного вмешательства, так и лекарственных препаратов). Однако имея в настоящее время базу статей и базу патентов можем ретроспективно проследить то, как современная технология развивалась. Имея данные наблюдений за прошедшими процессами, можно сделать предположение, что технология, востребованная в будущем, может быть спро-

гнозирована сейчас на основе методов прогнозной аналитики (рис. 2).

Предлагается, наблюдая за тенденциями увеличения количества патентов (как показателем прошедшего этапа преобразования статей в патенты) прогнозировать востребованность и рынка труда, и образовательной услуги.

В работе рассматривается сфера цифровых технологий, поскольку результаты могут быть легко интерпретированы как самими авторами, так и широким кругом читателей. Однако предполагается, что разрабатываемая методика может быть расширена и перенесена на любую другую научную область, которая связана с технологиями.



Целью исследования является разработка числовых характеристик для выявления и оценки передовых технологий на основе патентного анализа. Объектом анализа являются временные ряды, построенные по каждому ключевому слову по количеству упоминаний в текстах патентов за каждый год (с 2000 по 2021).

Вклад настоящей статьи заключается в следующем:

- предложено получить интерпретируемые числовые характеристики, дающие возможность оценки перспективности технологий и применения алгоритмов выбора тем для наполнения образовательных программ;
- рассмотрены примеры вычислений предложенных характеристик для предметной области «цифровой маркетинг»;
- рассмотрены возможности применения для тематического наполнения образовательных программ.

2. Анализ связанных работ

Ключевые слова являются частью метаданных научных, технологических текстовых источников – статей, запросов и патентов [3], [5]. Патенты, как источник для анализа перспективных технологий, являются достаточно признанными, а внимание к ним привлечено в значительном количестве исследований [6]. Патентный анализ является ценным подходом для получения информации об отрасли или технологии для прогнозирования как на основе библиометрических данных, так и на основе анализа патентной информации. Например, с помощью сети цитирования и кластеризации топологии показан процесс конвергенции научных знаний для прогнозирования новых технологий [7].

Можно разделить несколько направлений анализа патентов для выявления новых технологий: работы анализирующие библиометрические и патентные метаданные для анализа ценности технологий, работы по анализу и кластеризации текста патентов, использованию методов контролируемого обучения на основе размеченных образцов, глубокого обучения [8-11]. Учитывая сложность патентных текстов, библиометрический анализ более распространен в исследованиях, поскольку он включает анализ структурированных данных. Однако анализ метаданных не может охватить подробное техническое содержание патентов. Интеллектуальный анализ текста усиливает традиционные меры, основанные на библиометрических данных в прогнозировании технологических изменений [12].

Таким образом, установлено, что патенты являются одними из лучших источников информации для выбора перспективных технологий, однако автоматический анализ текстовых источников связан с большим количеством сложностей из-за сочетания технического и юридического языка терминов [13].

Задаче идентификации новых технологий из патентных текстов посвящены недавние работы. В [14] тестируются подходы на 1600 патентах из четырех тематических исследований, система позволила выявить более 4500 направлений технологий. В [15] проанализированы кластеры терминов и были определены потенциально перспективные направления развития для различных длин горизонта прогнозирования; количественный анализ, проведенный в работе, показал, что система может успешно определять возникающие и затухающие технологии. В [16] определяется, может ли новая технология или инновация,

если она может быть точно определена заранее, пройти или потерпеть неудачу. Также отмечается информационная перегрузка и сложность технологических знаний, которые оказывают негативное влияние на точность патентных поисковых систем. Текстовые сходства патентов обсуждаются в [17]. Широко используются методы машинного обучения [11], [18].

Патентный анализ используется во многих высокотехнологичных областях, таких как автоматическое вождение, в области промышленной инженерии, а также отслеживаются пути некоторых развитых технологий наногенераторов, конструкции электромобилей [19-22].

Существуют попытки использования и отличных от патентов источников. В [23] изложен масштабируемый метод для автоматического библиометрического анализа систематическим извлечением текста из репозитория arXiv. В [1], [24] разрабатывался сервис, собирающий данный вакансий рынка труда. Однако, прошедшие 5 лет после публикации статьи [1] показали, что большая точность при прогнозировании в 2017 году была получена с использованием данных, основанных на патентах. Несмотря на широкое распространение и впечатляющие результаты применения методов машинного обучения, надо признать, что наиболее вычислительно-реализуемые и хорошо интерпретируемые результаты в области поиска перспективных технологий основаны на линейном анализе трендов [25] и статистических методах [26], что позволяет осуществлять обоснованный выбор технологий без дополнительного экспериментального заключения. Далее в этой статье будем использовать статистические методы и линейную аппроксимацию.

Важным вопросом при выборе перспективности технологий является количественный показатель, который позволяет сравнивать технологии. В работе [27], в которой анализировали Web of Science, указано, что обширном обзоре литературы авторы обнаружили пробел в методологических исследованиях, так как нет числовых метрик для новых технологий. Наличие числовых показателей для новых технологий будет стимулировать исследования, позволит оценивать конкурентоспособность компаний, технологическую компетентность компаний и отдельных личностей, будет стимулировать обучение и институты. В [27] сформулированы задачи исследования, две из которых можно отнести и к настоящему исследованию, основанному на патентах, а именно – «каковы текущие темы исследований в области новых технологий?», «каково направление будущих исследований в области новых технологий?». В статье в ответ на первый вопрос мы оставляем за системой ChatGPT, ответ на второй состоит в цели статьи.

3. Данные для анализа

Была выбрана предметная область «цифровой маркетинг» (*Digital Marketing*). Исходный набор ключевых слов был получен с использованием ChatGPT (модель GPT-3.5). Полученный ответ на запрос “Give me the list of thousand keywords representing Digital Marketing domain” был оценен как удовлетворительный и содержал 300 позиций. Затем для этих ключевых слов была совершена предобработка: были оставлены только буквы, дефисы и знаки пробелов, а также убраны получившиеся дубликаты. После предобработки осталось 296 ключевых слов.



В качестве источника патентных данных был выбран сервис Espacenet. Данные о числе патентов в год были получены за период с 2000 по 2021 год включительно автоматизированным способом. Для этого был разработан программный скрипт. С помощью технологии Puppeteer производилось последовательное обращение к страницам сервиса Espacenet в браузере Chromium. Для каждого ключевого слова формировались URL вида [https://worldwide.espacenet.com/patent/search?q=\(nftxt%3D «keywords»\) AND pd within «2000-01-01%2C2000-12-31»](https://worldwide.espacenet.com/patent/search?q=(nftxt%3D «keywords») AND pd within «2000-01-01%2C2000-12-31»), где *«keywords»* – это направление, для которого производится извлечение количественных данных. Таким образом осуществлялся поиск патентных документов по годам (в примере – за 2000 год) и содержащих упоминание всей фразы, отражающей технологию или навык в рамках предметной области. Собранные данные были сохранены в формате CSV для последующей обработки. Данные за 2022 и 2023 год не использовались, так как не все патенты сразу попадают в открытый доступ, а, следовательно, за последние 2 года получаются неполные данные.

4. Методика численной оценки перспективности технологий

Каждому *i*-му ключевому слову ($i = \overline{1, n}$) соответствует временной ряд $T_i = \{t_{i,1}, \dots, t_{i,m}\}$, отражающий динамику патентной активности за выбранный период, где *m* – длина временного ряда. Путем анализа этих временных рядов можно определить, соответствуют ли ключевые слова перспективным направлениям. Задано пороговое значение \tilde{L} – минимальное число патентов в год, которое должно встретиться хотя бы один раз за выбранный период. При $\tilde{L} = 20$ для каждого временного ряда проверим, присутствует ли хоть один элемент, превышающий заданный порог, т. е. выполняется ли условие $\max_{j=1, m} t_{i,j} > \tilde{L}$. Таким образом можно исключить из дальнейшего рассмотрения ключевые слова, не имеющие достаточного числа упоминаний в патентах, чтобы их можно было расценить как перспективные. По итогу проверки соответствия условию из 296 ключевых слов 110 были определены как подходящие для последующего анализа, а 186 были исключены.

В исследованиях, как правило, рассматривается жизненный цикл технологий, который состоит из роста, плато и последующего спада, либо он представляется в виде S-функции [11]. На практике можно встретить и другие варианты развития технологий. Например, технология может активно развиваться, затем достичь этапа стабильного использования и затем, под влиянием внешних обстоятельств, снова войти в стадию развития. По этой причине гипотеза о схожести патентной динамики с моделью жизненного цикла из роста, плато и спада была отклонена.

Так как основной интерес представляет краткосрочный прогноз на 5-6 лет, наиболее значимым является последний тренд. Для этого предлагается провести сегментирование временных рядов динамики ежегодного количества патентов и далее анализировать последний сегмент. Каждый из сегментов предлагается называть участком стабильности (стабильного роста, стабильного спада, плато).

Чтобы провести сегментирование получим конечную разность первого порядка для временного ряда T_i ($i = \overline{1, n}$) по

формуле $T'_i = \{t_{i,2} - t_{i,1}, t_{i,3} - t_{i,2}, \dots, t_{i,m} - t_{i,m-1}\}$. Затем, приведем временной ряд T'_i к диапазону $[-1, 1]$ и получим таким образом ряд T''_i . Полученный временной ряд T''_i отражает динамику относительных изменений числа патентов в год по сравнению с предыдущим годом. Если элемент ряда выше 0 – это означает рост числа патентов; если ниже 0 – спад. При значении элемента равном 0 число патентов остается прежним, что соответствует стабильному участку плато.

К каждому временному ряду T''_i применим алгоритм сегментирования, представленный на рис. 3. Он основывается на скользящем окне и классифицирует отрезки на временном ряде по трем классам: стабильный рост, стабильный спад, плато. Минимальный размер участка стабильности при этом равен размеру скользящего окна. Для настоящей задачи зададим размер окна равным 4, т. е. будем оценивать динамику изменений числа патентов за 4 года.

```

algorithm ts-segmentation is
    input: Временный ряд  $X = \{x_1, x_2, \dots, x_n\}$ ,
    Пороговое значение плато  $pt$ ,
    Размер окна  $ws$ .
    output: Список сегментов  $SEG$ , где каждый сегмент – это словарь из пар
    ключ-значение.

     $SEG \leftarrow []$ 
     $pos \leftarrow 0$ 
     $previousClass \leftarrow 'unknown'$ 
     $currentClass \leftarrow 'unknown'$ 

    while  $pos + ws \leq n$  do
         $slice \leftarrow \{x_i : x_i \in X, pos \leq i < pos + ws\}$ 

        if  $|E[slice]| \leq pt$  then
             $currentClass \leftarrow 'plateau'$ 
        if  $E[slice] > pt$  then
             $currentClass \leftarrow 'ascent'$ 
        if  $E[slice] < -pt$  then
             $currentClass \leftarrow 'descent'$ 

        if  $currentClass \neq previousClass$  then
            append  $SEG$  with dictionary {'from': pos, 'class': currentClass}
             $previousClass \leftarrow currentClass$ 
             $pos \leftarrow pos + ws$ 
        else
             $pos \leftarrow pos + 1$ 

    return  $SEG$ 

```

Рис. 3. Алгоритм сегментирования временного ряда

Fig. 3. Time Series Segmentation Algorithm

Исходя из предположения, что разница в числе патентов за выбранный период не будет равна 0, зададим пороговое значение 0.05 для определения плато. Так, если рост или спад патентов за 4 года не будет превышать 0.05, алгоритм классифицирует участок стабильности как плато. Если рост превышает 0.05, то алгоритм классифицирует сегмент как участок стабильного роста. Если спад больше 0.05, алгоритм отметит сегмент как участок стабильного спада. Если класс не изменяется по сравнению с предыдущим, то новое окно рассматривается как продолжение предыдущего участка стабильности. На рис. 4-6 представлены примеры временных рядов, на которые перенесены результаты сегментирования. Так, на рис. 4 показан пример ключевого слова, для которого наблюдается стабильный рост. На рис. 5 приведено ключевое слово, для



которого наблюдается плато на последнем участке стабильности. На рис. 6 показан пример ключевого слова, у которого идет стабильный спад числа патентов.

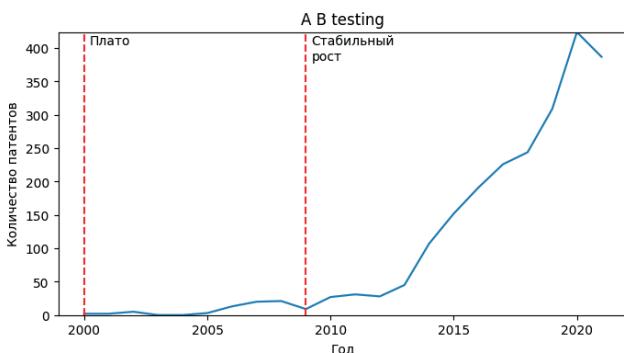


Рис. 4. Пример ключевого слова, характеризуемого ростом на последнем участке стабильности

F i g. 4. An example of a keyword characterized by growth in the last section of stability

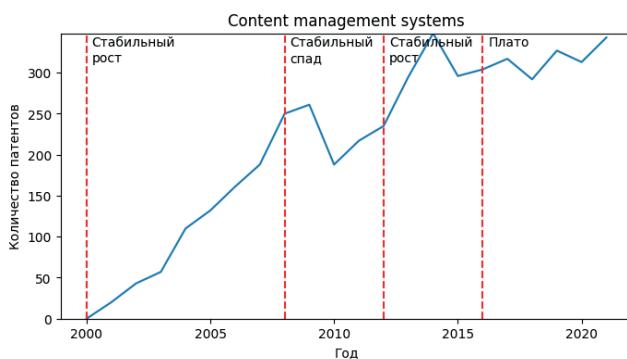


Рис. 5. Пример ключевого слова, характеризуемого плато на последнем участке стабильности

F i g. 5. An example of a keyword characterized by a plateau in the final section of stability

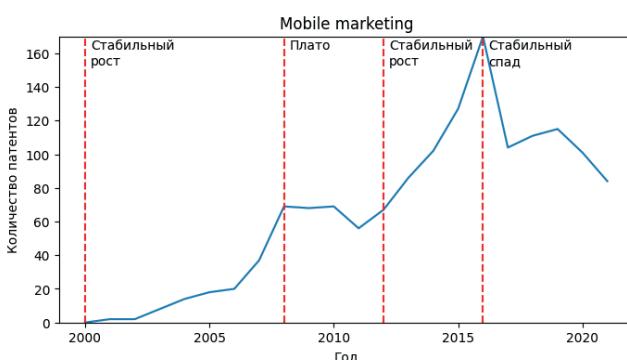


Рис. 6. Пример ключевого слова, характеризуемого спадом на последнем участке стабильности

F i g. 6. An example of a keyword characterized by a decline in the last section of stability

Из полученных с помощью алгоритма документов SEG извлечен последний. Он состоит из обозначения класса и свойства "from" – отметки начала сегмента на временном ряде. Затем из временных рядов T_i ($i = 1, n$) можно получить последний участок стабильности $\hat{T}_i = \{t_{i,j} : t_{i,j} \in T_i, q \leq j \leq m\}$, где T_i – исходный временной ряд, q – это отметка начала сегмента, m – длина временного ряда.

На основе последнего участка стабильности можно получить численные оценки, характеризующие технологию в данный момент времени:

- 1) Среднее ежегодное число патентов на последнем участке стабильности $E[\hat{T}_i]$;
- 2) Динамику числа патентов на последнем участке стабильности, которую можно получить как коэффициент линейной регрессии $S_i = \Delta y_i / \Delta x_i$.

Полученные численные оценки, характеризующие перспективность технологии, можно отобразить как координаты на плоскости параметров. Плоскость параметров позволяют интуитивно понятным образом на основе численных оценок классифицировать технологии.

Таким образом, можно сформировать методику получения численных оценок перспективности технологий, состоящую из 7 этапов.

1. Формирование с помощью эксперта или с помощью искусственного интеллекта набора ключевых слов, потенциально претендующих на роль тем образовательной программы;
2. Выбор из патентной базы патентов, в которых встречаются ключевые слова;
3. Формирование временных рядов;
4. Сокращение количества ключевых слов на основе проверки пороговых значений;
5. Классификация рядов (стабильные, растущие, падающие) на основе алгоритма поиска стабильных участков в последние годы анализа;
6. Вычисление параметров и построение плоскости характеристик;
7. Принятие решений о тематическом планировании на основе анализа плоскости параметров.

5. Результаты исследования

В качестве примера осуществим тематическое наполнение по программе «цифровой маркетинг». В соответствии с пп. 1-6 методики, представленной выше, получим плоскость параметров, показанную на рис. 7. Возможно использовать различные, в том числе и машинные методы классификации, здесь для иллюстрации используем метод разделения плоскости параметров в зависимости от значений на классы. Разделим технологии на 5 классов.

Класс 1 (синие точки) характеризуется большим количеством патентов в год (800 и более). Этот класс включает в себя основные технологические тренды, которые не ограничиваются конкретными реализациями или прикладными областями. Это определяющие направления развития технологий в выбранной предметной области и их изучение является фундаментальным для конкурентоспособности.

Класс 2 (зеленые точки) состоит из технологий, характеризуе-



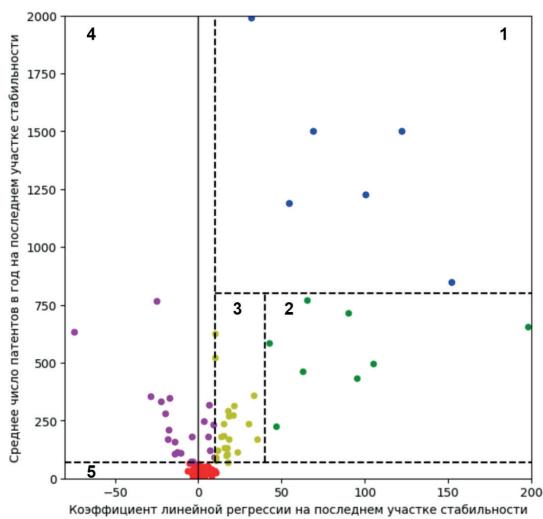


Рис. 7. Плоскость численных параметров в области «цифровой маркетинг»

Fig. 7. The plane of numerical parameters in the field of “digital marketing”

мых значительным ежегодным числом патентов (от 70 до 800) и значительным ростом. Ключевые слова, вошедшие в этот класс – это основные перспективные технологии.

Класс 3 (желтые точки) содержит перспективные технологии, у которых наблюдается снижение роста числа патентов, либо стабильный период, после которого рост может возобновиться. Эти технологии являются важными для специалистов в выбранной области. Стоит заметить, что сама предметная область «цифровой маркетинг» тоже попала в этот класс как ключевое слово. Из этого можно сделать вывод о перспективности направления в целом и проводить сравнение с другими областями используя численные оценки.

Класс 4 (точки цвета маджента) включает технологии, для которых наблюдается спад, но при этом абсолютные значения числа патентов достаточно велики, поэтому их изучение имеет смысл.

Класс 5 (красные точки) состоит из ключевых слов с небольшим числом патентов (менее 70 в год), что делает прогнозирование затруднительным.

Перечень ключевых слов, распределенных по перечисленным классам, приведен в таблице 1.

Таблица 1. Перечень ключевых слов в области «цифровой маркетинг»

Table 1. List of keywords in the field of “digital marketing”

Класс	Ключевые слова
1	Branding; Content creation; Customer relationship management; Data analysis; Data segmentation; Market research; Return on investment.
2	Brand loyalty; E-commerce platforms; Interactive content; Predictive analytics; Public relations; Retargeting; Social media platforms; User-generated content.
3	A/B testing; Behavioral analytics; Brand awareness; CRM software; Call-to-action; Content distribution networks; Customer acquisition; Customer engagement; Customer retention; Data visualization tools; Digital advertising; Digital marketing; Display advertising; Email marketing; Lead generation; Logo design; Marketing automation; Marketing campaigns; Marketing strategy; Online marketing; Online surveys; Search engine optimization; User interface design.
4	Ad targeting; Banner advertising; Blogging; Brand identity; Competitive analysis; Content management systems; Cost-per-click; Cost-per-impression; Crisis management; Customer segmentation; Internet marketing; Loyalty rewards; Mobile advertising; Mobile marketing; Online advertising; Online communities; Online presence; Search engine marketing; Sponsored content; Viral marketing; Web analytics.
5	Affiliate marketing; Affiliate networks; Analytics dashboards; Augmented reality marketing; Behavioral targeting; Blogging platforms; CRM platforms; Cart abandonment; Competitor analysis; Content calendar; Content marketing; Content syndication; Contextual advertising; Conversion funnel; Conversion optimization; Conversion tracking; Customer journey analytics; Customer lifetime value; Customer loyalty programs; Customer onboarding; Customer satisfaction surveys; E-commerce marketing; Email newsletter; Google AdWords; Heatmap analysis; In-app advertising; Influencer marketing; Lead magnet; Loyalty program management; Native advertising; Pay-per-click advertising; Programmatic advertising; Project management tools; Remarketing; SMS marketing; Search engine rankings; Social media advertising; Social media analytics; Social media engagement; Social media influencers; Social media listening; Social media management; Social media marketing; Social media monitoring; Thought leadership; User experience design; User-generated content platforms; Video content creation; Video marketing; Website optimization; YouTube advertising.

Список тем для образовательной программы по направлению «цифровой маркетинг» может быть составлен из ключевых слов, вошедших в классы 1–4. Содержание класса 1 должно быть либо обязательным к изучению, либо студенты должны быть знакомы с этими концепциями заранее. Ключевые слова из классов 2–4 стоит внести в основной блок дисциплин, а сло-

ва из класса 5 можно рассматривать как темы для углубленного изучения.



6. Обсуждение и заключение

Для анализа перспективных технологий, которые формируют тематическое наполнение образовательной программы, в статье разработана методика численных оценок перспективных технологий на основе динамики изменения числа патентов. Был предложен масштабируемый инструмент численного анализа для выбора перспективных технологий, включающий два параметра – скорость роста/снижения и количество патентов, обе характеристики рассчитываются на последнем стабильном этапе жизненного цикла развития технологий. Для примера было осуществлено формирование тематического наполнение образовательной программы «цифровой маркетинг». В качестве источника для получения перечня ключевых слов была использована искусственная нейронная сеть ChatGPT (модель GPT-3.5). Для 296 ключевых слов были получены временные ряды с ежегодным количеством патентных документов, содержащих упоминание конкретного навыка или технологии. Методика была применена к собранным данным. В результате была получена плоскость параметров, на основе которой были

идентифицированы 5 классов ключевых слов. На основе этого были выбраны актуальные направления для формирования образовательной программы в области «цифрового маркетинга». Предложенную методику идентификации актуальных технологий и навыков можно применять для выявления наиболее значимых направлений в рамках выбранной предметной области. Применение патентных данных для количественного анализа может давать преимущество при составлении учебных программ, так как перспективные навыки и направления могут быть выявлены до их активного внедрения на рынок труда. Методика в текущем виде не подразумевает автоматизацию поиска пороговых значений и периодов для сбора количественных данных. Их выбор остается за экспертом, однако в будущем можно рассмотреть методы многокритериальной оптимизации для поиска эффективных пороговых значений для предложенных условий. Кроме того, дальнейшие исследования могут включать методы машинного обучения, дополнять и уточнять приведенные условия для определения актуальных направлений, а также рассматривать тексты патентных документов более детально для повышения точности оценки.

References

- [1] Ilin D.Y., Nikulchev E.V., Bubnov G.G., Mateshuk E.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;2(38):71-88. (In Russ., abstract in Eng.) EDN: ZBKKLH
- [2] Uglev V.A. Evaluate Curricula Balance for Software Engineering Education with using UGVA Method. *Modern Information Technologies and IT-Education*. 2021;17(3):684-696. (In Russ., abstract in Eng.) <https://doi.org/10.25559/SITITO.17.202103.684-696>
- [3] Nikulchev E.V., Ilin D.Y., Bubnov G.G. The actualization of educational programs based on patent analysis as an indicator of the development of innovative technologies. *Cloud of science*. 2017;4(4):513-524. (In Russ., abstract in Eng.) EDN: YSCXIB
- [4] Yarullin D.V. Intelligent Control System for IT Specialists Training Based on Denotative Analytics. *Applied Mathematics and Control Sciences*. 2022;(3):141-164. (In Russ., abstract in Eng.) <https://doi.org/10.15593/2499-9873/2022.03.08>
- [5] Belov M.A., Grishko S.I., Cheremisina E.N., Tokareva N.A. Training of IT Specialists in the Conditions of Global Digital Transformation. The Concept of Automated Management of Competency Profiles in the Educational Programs of the Future. *Modern Information Technologies and IT-Education*. 2021;17(3):658-669. (In Russ., abstract in Eng.) <https://doi.org/10.25559/SITITO.17.202103.658-669>
- [6] Daim T.U., Rueda G., Martin H., Gerdsri P. Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technological Forecasting and Social Change*. 2006;73(8):981-1012. <https://doi.org/10.1016/j.techfore.2006.04.004>
- [7] Zhou Y., Dong F., Kong D., Liu Y. Unfolding the convergence process of scientific knowledge for the early identification of emerging technologies. *Technological Forecasting and Social Change*. 2019;144:205-220. <https://doi.org/10.1016/j.techfore.2019.03.014>
- [8] Breitzman A., Thomas P. Inventor team size as a predictor of the future citation impact of patents. *Scientometrics*. 2015;103(2):631-647. <https://doi.org/10.1007/s11192-015-1550-5>
- [9] Kyebambe M.N., Cheng G., Huang Y., He C., Zhang Z. Forecasting emerging technologies: A supervised learning approach through patent analysis. *Technological Forecasting and Social Change*. 2017;125:236-244. <https://doi.org/10.1016/j.techfore.2017.08.002>
- [10] Aristodemou L., Tietze F. The state-of-the-art on Intellectual Property Analytics (IPA): A literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data. *World Patent Information*. 2018;55:37-51. <https://doi.org/10.1016/j.wpi.2018.07.002>
- [11] Zhou Y., Dong F., Liu Y., Li Z., Du J., Zhang L. Forecasting emerging technologies using data augmentation and deep learning. *Scientometrics*. 2020;123:1-29. <https://doi.org/10.1007/s11192-020-03351-6>
- [12] Vicente-Gomila J.M., Artacho-Ramirez M.A., Ting M., Porter A.L. Combining tech mining and semantic TRIZ for technology assessment: Dye-sensitized solar cell as a case. *Technological Forecasting and Social Change*. 2021;169:120826. <https://doi.org/10.1016/j.techfore.2021.120826>
- [13] Joung J., Kim K. Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data. *Technological Forecasting and Social Change*. 2017;114:281-292. <https://doi.org/10.1016/j.techfore.2016.08.020>
- [14] Puccetti G., Giordano V., Spada I., Chiarello F., Fantoni G. Technology identification from patent texts: A novel named entity recognition method. *Technological Forecasting and Social Change*. 2023;186:122160. <https://doi.org/10.1016/j.techfore.2022.122160>
- [15] Gozuacik N., Sakar C.O., Ozcan S. Technological forecasting based on estimation of word embedding matrix using LSTM networks. *Technological Forecasting and Social Change*. 2023;191:122520. <https://doi.org/10.1016/j.techfore.2023.122520>



- [16] Chi Y.C., Wang H.C. Establish a patent risk prediction model for emerging technologies using deep learning and data augmentation. *Advanced Engineering Informatics*. 2022;52:101509. <https://doi.org/10.1016/j.aei.2021.101509>
- [17] Hain D.S., Jurowetzki R., Buchmann T., Wolf P. A text-embedding-based approach to measuring patent-to-patent technological similarity. *Technological Forecasting and Social Change*. 2022;177:121559. <https://doi.org/10.1016/j.techfore.2022.121559>
- [18] Kraus S., Kumar S., Lim W.M., Kaur J., Sharma A., Schiavone F. From moon landing to metaverse: Tracing the evolution of Technological Forecasting and Social Change. *Technological Forecasting and Social Change*. 2023;189:122381. <https://doi.org/10.1016/j.techfore.2023.122381>
- [19] Ulrich C., Frieske B., Schmid S.A., Friedrich H.E. Monitoring and Forecasting of Key Functions and Technologies for Automated Driving. *Forecasting*. 2022;4(2):477-500. <https://doi.org/10.3390/forecast4020027>
- [20] Calleja-Sanz G., Olivella-Nadal J., Solé-Parellada F. Technology Forecasting: Recent Trends and New Methods. In: Machado C., Davim J. (eds.) Research Methodology in Management and Industrial Engineering. *Management and Industrial Engineering*. Cham: Springer; 2020. p. 45-69. https://doi.org/10.1007/978-3-030-40896-1_3
- [21] Liu Y., Wang G., Zhou Y., Liu Y. Advanced Technology Evolution Pathways of Nanogenerators: A Novel Framework Based on Multi-Source Data and Knowledge Graph. *Nanomaterials*. 2022;12(5):838. <https://doi.org/10.3390/nano12050838>
- [22] Feng L., Liu K., Wang J., Lin K.Y., Zhang K., Zhang L. Identifying Promising Technologies of Electric Vehicles from the Perspective of Market and Technical Attributes. *Energies*. 2022;15(20):7617. <https://doi.org/10.3390/en15207617>
- [23] Nikulchev E., Ilin D., Bubnov G., Mateshuk E. Scalable service for predictive learning based on the professional social networking sites. *International Journal of Advanced Computer Science and Applications*. 2017;8(5):9-15. <http://dx.doi.org/10.14569/IJACSA.2017.080502>
- [24] Percia D.D., Blonay W., Gillard S., Maillart T., Mermoud A., Maréchal L., Tsesmelis M. Identification of Future Cyberdefense Technology by Text Mining. *Cyberdefense. International Series in Operations Research & Management Science*. 2023;342:69-86. https://doi.org/10.1007/978-3-031-30191-9_5
- [25] Adamuthe A.C., Thampi G.T. Technology forecasting: A case study of computational technologies. *Technological Forecasting and Social Change*. 2019;143:181-189. <https://doi.org/10.1016/j.techfore.2019.03.002>
- [26] Haleem A., Mannan B., Luthra S., Kumar S., Khurana S. Technology forecasting (TF) and technology assessment (TA) methodologies: a conceptual review. *Benchmarking: An International Journal*. 2019;26(1):48-72. <https://doi.org/10.1108/BIJ-04-2018-0090>
- [27] Zamani M., Yalcin H., Naeini A.B., Zeba G., Daim T.U. Developing metrics for emerging technologies: identification and assessment. *Technological Forecasting and Social Change*. 2022;176:121456. <https://doi.org/10.1016/j.techfore.2021.121456>

Поступила 25.06.2024; одобрена после рецензирования 04.08.2024; принята к публикации 17.09.2024.
Submitted 25.06.2024; approved after reviewing 04.08.2024; accepted for publication 17.09.2024.

Об авторах:

Никульчев Евгений Витальевич, профессор кафедры цифровых технологий обработки данных Института кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119452, Российская федерация, г. Москва, пр. Вернадского, д. 78), доктор технических наук, профессор, профессор РАО, ORCID: <https://orcid.org/0000-0003-1254-9132>, nikulchev@mail.ru

Грюкан Ярослав Сергеевич, магистрант кафедры цифровых технологий обработки данных Института кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119452, Российская федерация, г. Москва, пр. Вернадского, д. 78), ORCID: <https://orcid.org/0009-0004-2179-8995>, gryukan@mirea.ru

Ильин Дмитрий Юрьевич, доцент кафедры цифровых технологий обработки данных Института кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119452, Российская федерация, г. Москва, пр. Вернадского, д. 78), кандидат технических наук, ORCID: <https://orcid.org/0000-0002-0241-2733>, i@dmitryilin.com

Все авторы прочитали и одобрили окончательный вариант рукописи.

About the authors:

Evgeny V. Nikulchev, Professor of the Chair of Digital Data Processing Technologies, Institute for Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78 Vernadsky Ave., Moscow 119454, Russian Federation), Dr. Sci. (Tech.), Professor, Processor of Russian Academy of Education, ORCID: <https://orcid.org/0000-0003-1254-9132>, nikulchev@mail.ru

Yaroslav S. Gryukan, Master Degree Student of the Chair of Digital Data Processing Technologies, Institute for Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78 Vernadsky Ave., Moscow 119454, Russian Federation), ORCID: <https://orcid.org/0009-0004-2179-8995>, gryukan@mirea.ru

Dmitry Yu. Ilin, Associate Professor of the Chair of Digital Data Processing Technologies, Institute for Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78 Vernadsky Ave., Moscow 119454, Russian Federation), Cand. Sci. (Eng.), ORCID: <https://orcid.org/0000-0002-0241-2733>, i@dmitryilin.com

All authors have read and approved the final manuscript.

