

УДК 004.853

DOI: [10.26102/2310-6018/2022.39.4.010](https://doi.org/10.26102/2310-6018/2022.39.4.010)

## Выявление технологической взаимодополняемости предприятий на основе анализа патентного массива

Д.М. Коробкин✉, А.Ю. Безрученко, С.А. Фоменков, С.Г. Колесников

*Волгоградский государственный технический университет,  
Волгоград, Российская Федерация  
dkorobkin80@mail.ru✉*

**Резюме.** Технологическая взаимодополняемость (комплементарность) играет все более важную роль в принятии стратегических решений о том, с какими партнерами вести бизнес, а также для расширения разнообразия инновационной деятельности путем создания альянсов или слияний. Современные количественные измерения взаимодополняемости технологий в значительной степени основаны на исследованиях, основанных на патентных классификационных кодах. Однако этот подход игнорирует конкретные технологии в той или иной области и, следовательно, полученные результаты являются лишь общими указателями. Предлагается разработать подход к количественному измерению комплементарности технологий предприятий на основе статистического машинного анализа патентного массива. Целью исследования является разработка технологии кластеризации документов патентного ведомства США USPTO (United States Patent and Trademark Office) и выявления технологической взаимодополняемости предприятий на основе сравнения кластерной информации. Теоретическая ценность заключается в разработанных алгоритмах парсинга патентов; кластеризации патентного массива на основе метода hLDA (Hierarchical Latent Dirichlet Allocation, иерархическое латентное размещение Дирихле); формирования матрицы «Предприятия – Кластеры (Темы)»; визуализации связей предприятий в кластеризованном патентном пространстве. Практическая значимость работы заключается в разработанном программном приложении выявления технологической взаимодополняемости предприятий, эффективность работы которого была проверена на ряде тестовых примеров.

**Ключевые слова:** патент, USPTO, технологическая взаимодополняемость, кластеризация, hLDA.

**Благодарности:** исследование выполнено за счет гранта Российского научного фонда № 22-21-20125, <https://rscf.ru/project/22-21-20125/>, и Администрации Волгоградской области.

**Для цитирования:** Коробкин Д.М., Безрученко А.Ю., Фоменков С.А., Колесников С.Г. Выявление технологической взаимодополняемости предприятий на основе анализа патентного массива. *Моделирование, оптимизация и информационные технологии.* 2022;10(4). Доступно по: <https://moitvvt.ru/ru/journal/pdf?id=1264> DOI: 10.26102/2310-6018/2022.39.4.010

## Identification of technological complementarity of enterprises based on the analysis of the patent database

D.M. Korobkin✉, A.Y. Bezruchenko, S.A. Fomenkov, S.G. Kolesnikov

*Volgograd State Technical University,  
Volgograd, Russian Federation  
dkorobkin80@mail.ru✉*

**Abstract.** Technological complementarity plays an increasingly important role in making strategic decisions regarding the choice of partners to work with as well as for expanding the diversity of innovation activities through the creation of alliances or mergers. Modern quantitative measurements of technology complementarity are largely based on research founded on patent classification codes.

However, this approach ignores specific technologies in a particular field and, therefore, the results obtained are only general pointers. It is proposed to develop an approach to the quantitative measurement of the complementarity of enterprise technologies in reliance on statistical machine analysis of the patent array. The aim of the study is to develop a technology for clustering USPTO patent documents and identifying technological complementarity of enterprises based on the comparison of cluster information. The theoretical value lies in the following: the developed algorithms for patent parsing; clustering of the patent array based on the hLDA method; formation of "Enterprises – Clusters (Topics)" matrix; visualization of enterprise connections in a clustered patent space. The practical significance of the research lies in the developed software for identifying technological complementarity of enterprises, the effectiveness of which has been tested on a number of test examples.

**Keywords:** patent, USPTO, technological complementarity, clustering, hLDA.

**Acknowledgements:** the study was supported by the grant of the Russian Science Foundation No. 22-21-20125, <https://rscf.ru/en/project/22-21-20125/>, and the Administration of Volgograd Oblast.

**For citation:** Korobkin D.M., Bezruchenko A.Y., Fomenkov S.A., Kolesnikov S.G. Identification of technological complementarity of enterprises based on the analysis of the patent database. *Modeling, Optimization and Information Technology*. 2022;10(4). Available from: <https://moitvvt.ru/ru/journal/pdf?id=1264> DOI: 10.26102/2310-6018/2022.39.4.010 (In Russ.).

## Введение

Обладание инновациями является важным фактором успеха компании, и многие фирмы вкладывают большие деньги в научно-исследовательские и опытно-конструкторские работы (НИОКР). При этом часто в одной и той же области ведут исследования несколько компаний, и им может быть выгодно партнерство, поглощение [1] или слияние друг с другом. При этом известно, что наилучшие результаты получаются при высокой взаимодополняемости (комплементарности) технологий в этих компаниях.

В настоящее время исследование патентных массивов в основном сосредоточено на выявлении сходства патентов, и относительно мало исследований по определению взаимодополняемости патентов как основы для выявления комплементарности предприятий, использующих технологии, описанные в патентах. Исследователи [2] использовали модель LDA (Латентное размещение Дирихле) для создания промышленных деревьев при изучении патентов. Эксперты использовали дополнительные баллы для каждой технологической отрасли, чтобы рассчитать взаимодополняемость между патентами. Авторы [3] разработали комбинацию портфелей патентов на основе анализа ожидаемых отказов, включая анализ портфелей патентов, анализ отказов, инструменты ТРИЗ (Теория решения изобретательских задач) для решения проблем и оценку зонтичных портфелей патентов. В работе [4] используется метод корреляционного анализа, чтобы обнаружить дополнительные отношения между патентами.

Авторы работы [5] занимаются измерением точной технологической взаимодополняемости между различными предприятиями в области медицины на основе морфологического подхода. Вычисляется семантическое сходство между субъектами (S, Subject) и объектами действия (АО, Action-Object) на основе разработанного метатегауруса, затем формируются кластеры S и АО (ключевые технологические проблемы и методы) на основе матрицы семантического сходства. Строится технологическая морфологическая матрица нескольких измерений, и матрица заполняется структурами субъект-действие-объект (SAO) в соответствии с соответствующими ключевыми технологическими проблемами и методами для различных предприятий. Полученная матрица морфологии технологий используется для

измерения взаимодополняемости технологий между различными предприятиями на основе SAO. В работе [6] также используется морфологический подход для определения технологической взаимодополняемости предприятий. Проблема данного подхода заключается в недостаточной точности работы семантических анализаторов русскоязычных текстов для извлечения структур «субъект-действие-объект» (SAO).

В работе [7] предложен алгоритм вычисления взаимодополняемости патентов, который выявляет семантическую схожесть текстов патентов на основе модели глубокого обучения и количественно определяет взаимодополняемость патентов. Текст патента представляется векторами предложений с помощью модели LSTM, сгенерированные векторы предложений кластеризуются, выявляются темы. Принадлежность патента соответствующей теме определяется на основе модели классификации текста CNN.

Исследователи [8] предложили определение количественной оценки взаимодополняемости технологий между предприятиями на основе данных о патентах, полученных с помощью текстового анализа. Иерархическая латентная тематическая модель распределения Дирихле (hLDA) идентифицирует технологические темы, скрытые в патентных документах, вместе с иерархической структурой этих тем. Данный подход видится в качестве наиболее перспективного, однако в свободном доступе нет демонстрационной версии системы для проведения опытов.

Исходя из проведенного анализа существующих исследований, было решено реализовать систему, позволяющую вычислять и анализировать технологическую взаимодополняемость между предприятиями. В данном исследовании в качестве источника данных для вычисления комплементарности будем использовать базу патентных документов США USPTO (United States Patent and Trademark Office) [9] как наиболее полный и доступный источник мировой патентной информации. Внедрение программного модуля позволит облегчить и ускорить процесс поиска потенциальных партнеров по НИОКР по сравнению с ручной обработкой данных.

**Цель работы:** разработать программное приложение для кластеризации патентных документов USPTO на основе метода hLDA и выявления технологической взаимодополняемости предприятий на основе сравнения кластерной информации.

## Материалы и методы

### Методология определения технологической взаимодополняемости предприятий

Общий алгоритм процесса выявления технологической взаимодополняемости предприятий на основе кластеризации патентных документов USPTO представлен на Рисунке 1.

- 1) парсинг xml-файла с текстами патентов;
- 2) извлечение технических терминов из патентов;
  - 2.1) токенизация текста патентов;
  - 2.2) нормализация текста патентов (удаление стоп-слова);
  - 2.3) построение терм-документной матрицы на основе метрики tf-idf [10];
- 3) кластеризация обработанных текстов патентов (построение дерева тем и матрицы «Патенты – Темы»), кластеризация производится алгоритмом hLDA;
- 4) построение матрицы «Предприятия – Темы», матрица формируется подсчетом количества патентов каждой темы для каждого предприятия;
- 5) формирование матрицы взаимодополняемости.



Рисунок 1 – Общий алгоритм процесса выявления технологической взаимодополняемости предприятий на основе кластеризации патентных документов USPTO

Figure 1 – Algorithm of enterprise technological complementarity identification process based on clustering of USPTO patent documents

#### **Алгоритм кластеризации патентного массива на основе метода hLDA**

В работе модифицировалась классическая реализация алгоритма hLDA [11] (Рисунок 2). В алгоритме кластеризации используются параметры  $\alpha$ ,  $\gamma$ ,  $\eta$ .

Параметр  $\alpha$  сглаживает распределение слов по уровням, определяет, как много терминов приписывается более конкретным темам и как много более общим; чем выше, тем более равномерно, чем ниже, тем больше слов в общие темы ( $\alpha > 0$ , по умолчанию 10).

Параметр  $\gamma$  контролирует вероятность выбора нового пути (создания новой темы); чем выше  $\gamma$ , тем выше вероятность, но количество тем в основном зависит от  $\eta$  ( $\gamma > 0$ , по умолчанию 1).

Параметр  $\eta$  сглаживает распределение слов по темам, чем регулирует количество тем; чем меньше  $\eta$ , тем больше тем.  $0 < \eta < 1$ , по умолчанию 0.1.

Параметр  $\gamma$  увеличивает вероятность создания новых узлов, параметр  $\eta$  контролирует разреженность тем, параметр  $\alpha$  контролирует число тем, перенесенных с узла-корня на узлы-листья, т. е. контролирует, сколько слов назначено более конкретным темам и сколько более общим.



Рисунок 2 – Алгоритм кластеризации патентного массива на основе метода hLDA  
Figure 2 – Patent database clustering algorithm based on the hLDA method

Алгоритм инициализации дерева соответствует классической реализации метода hLDA, а в алгоритм итерации обучения дерева, показанный на Рисунке 3, были внесены изменения.

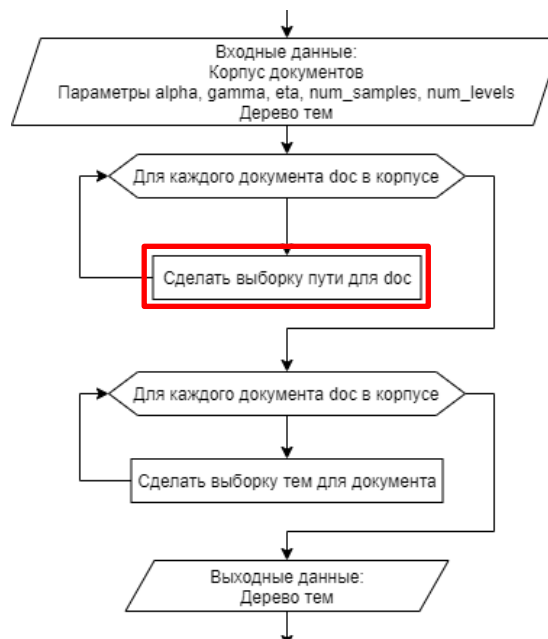


Рисунок 3 – Алгоритм итерации обучения дерева  
Figure 3 – Tree learning iteration algorithm

Алгоритм выборки пути для документа показан на Рисунке 4. Алгоритм выборки тем в данной работе соответствует классической реализации для метода hLDA.

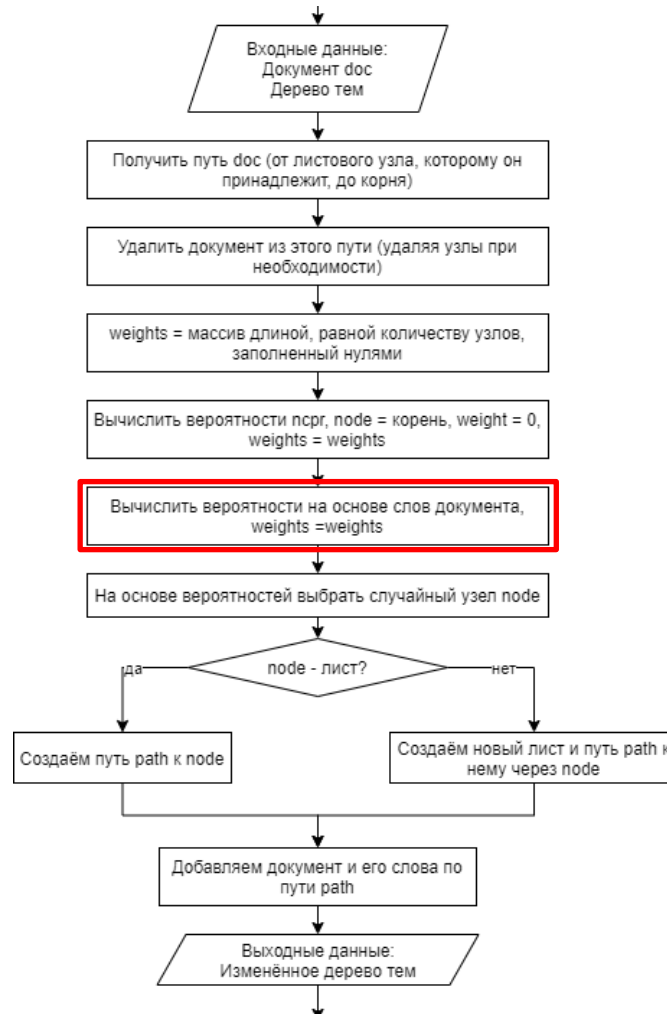


Рисунок 4 – Алгоритм выборки пути для документа  
Figure 4 – Path selection algorithm for a document

Модифицированный алгоритм вычисления вероятностей узлов на основе слов показан на Рисунке 5 ( $\lgamma$  – это натуральный логарифм гамма-функции).

Алгоритм вычисления вероятностей слов по всему дереву показан на Рисунке 6. В вышеуказанном алгоритме используется модифицированный алгоритм вычисления веса узла (Рисунок 7).

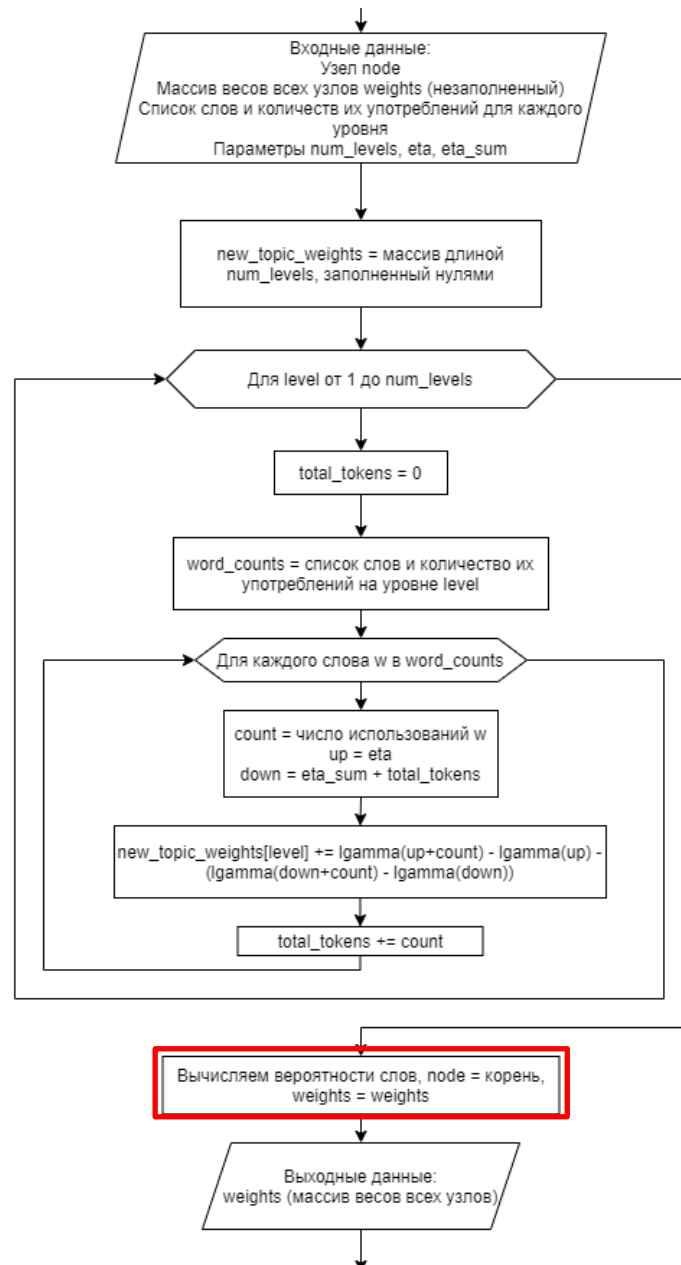


Рисунок 5 – Алгоритм вычисления вероятности на основе слов  
Figure 5 – Algorithm for calculating probability based on words

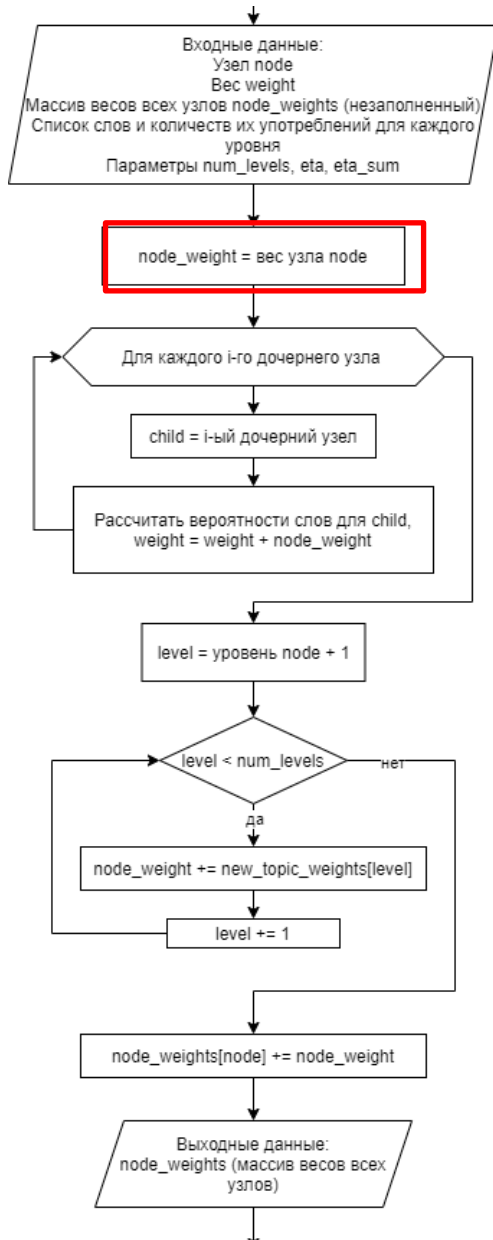


Рисунок 6 – Алгоритм вычисления вероятностей слов по дереву  
Figure 6 – Algorithm for calculating the probabilities of words on a tree

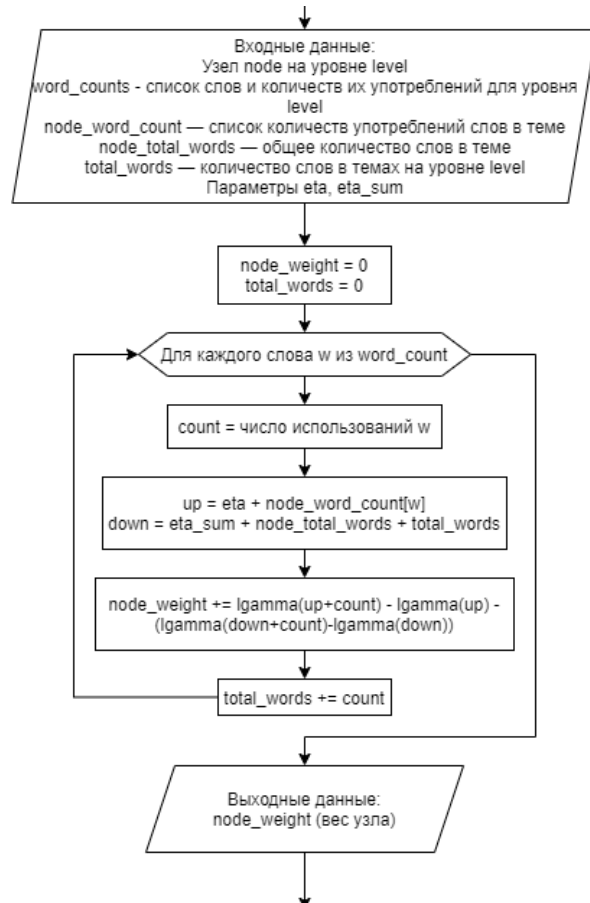


Рисунок 7 – Алгоритм вычисления веса узла  
Figure 7 – Algorithm for calculating the weight of a node

Изменения в алгоритме позволили ускорить соответствующую функцию примерно в 12 раз, а весь модуль кластеризации – примерно на 30 %. Для оптимизации классического версии алгоритма hLDA для кластеризации использовали замену одного из внутренних циклов формулой (1):

$$\sum_{w=0}^{wc} \sum_{i=0}^c \ln \left( \frac{\eta + wc_w + i}{\eta_{sum} + tw_n + tw_0 + i} \right) = \sum_{w=0}^{wc} \left( \frac{\ln(\Gamma(\eta + wc_w + c)) - \ln(\Gamma(\eta + wc_w))}{\ln(\Gamma(\eta_{sum} + tw_n + tw_0 + c)) + \ln(\Gamma(\eta_{sum} + tw_n + tw_0))} \right), \quad (1)$$



Это возможно, потому что внутренний цикл можно интерпретировать как разность логарифмов произведений арифметических прогрессий (которые можно вычислить с помощью гамма-функции) (Рисунки 8, 9).

```
for w in word_counts:
    count = word_counts[w]
    for i in range(count):
        node_weight += log( (self.eta + node_word_count[w] + i) /
                            (self.eta_sum + node_total_words + total_words) )
    total_words += 1
```

Рисунок 8 – Старый код  
Figure 8 – Old code

```
for w in word_counts:
    count = word_counts[w]
    up_part = self.eta + node_word_count[w]
    down_part = self.eta_sum + node_total_words + total_words
    node_weight += (math.lgamma(up_part+count) - math.lgamma(up_part)
                   - (math.lgamma(down_part+count) - math.lgamma(down_part)))
    total_words += count
```

Рисунок 9 – Новый код  
Figure 9 – New code

### Алгоритм построения матрицы взаимодополняемости

Расчет взаимодополняемости ведется по формуле:

$$Compl(A \leftarrow B) = \sum \frac{TN(B) - TN(A \& B)}{TN - TN(A)} \times \frac{PN(C_i)}{PN}, \quad (2)$$

где  $Compl(A \leftarrow B)$  – взаимодополняемость B по отношению к A,  $C_i$  – i-тая главная технологическая тема (т.е. более высокоуровневая тема в иерархии), TN – число подтем в  $C_i$ ,  $TN(A)$ ,  $TN(B)$ ,  $TN(A \& B)$  – число подтем в  $C_i$ , по которым у предприятия A, B, A и B есть патенты,  $PN(C_i)$  – число патентов в  $C_i$ ,  $PN$  – число патентов.

На основе формулы был сформирован алгоритм, показанный на Рисунке 10.

### Проектирование приложения определения технологической взаимодополняемости предприятий

Разрабатываемое программное приложение состоит из следующих частей:

- 1) модуль парсинга переводит изначальные файлы XML с патентами в удобный для последующей обработки вид;
- 2) модуль кластеризации патентов алгоритмом hLDA;
- 3) модуль анализа на основе данных из модулей парсинга и кластеризации строит матрицу «Предприятия – Темы» и матрицу взаимодополняемости предприятий;
- 4) модуль визуализации отображает матрицу взаимодополняемости предприятий.

Процесс анализа патентного массива показан на диаграмме потоков данных на Рисунке 11.

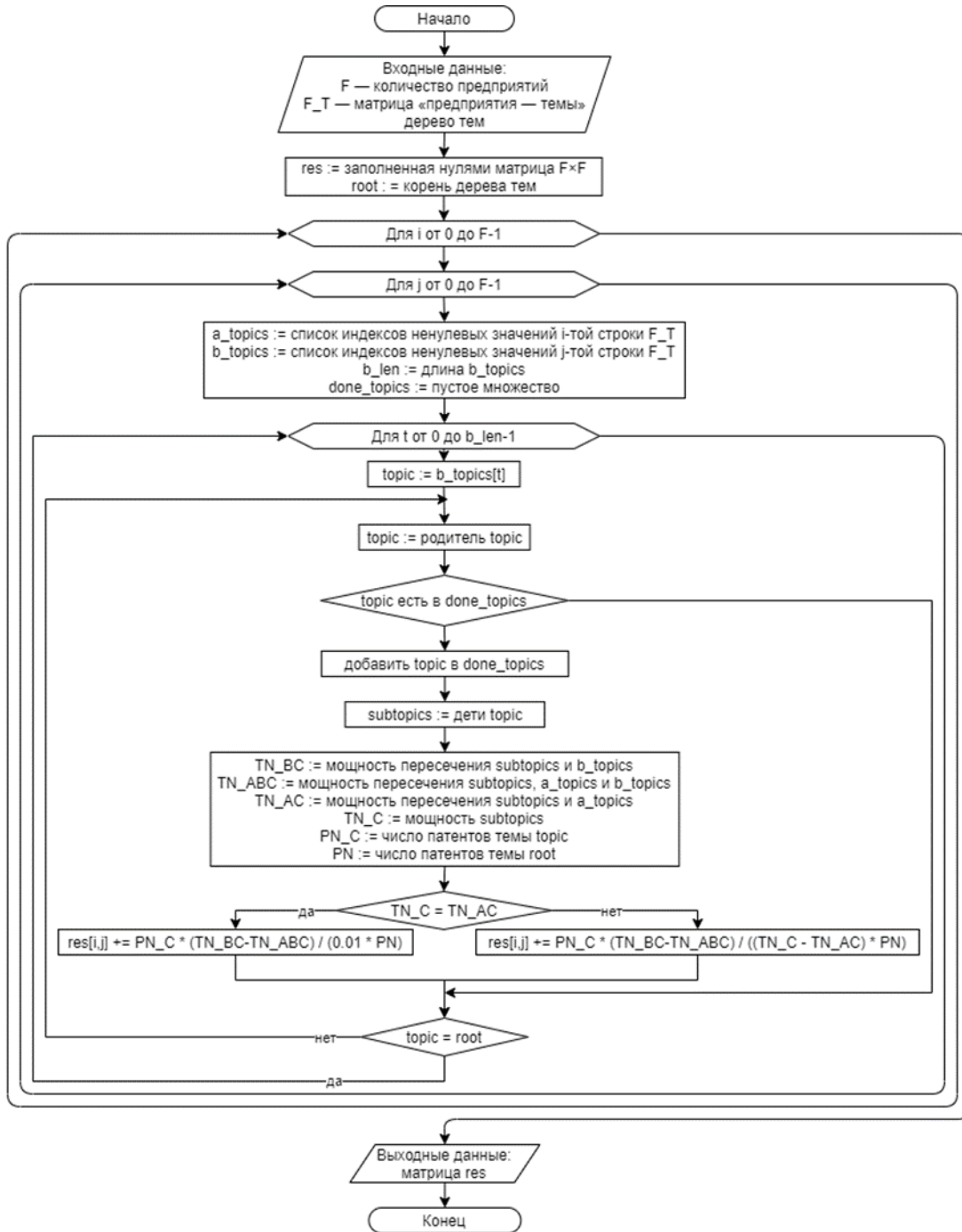


Рисунок 10 – Алгоритм построения матрицы взаимодополняемости  
 Figure 10 – Algorithm for constructing the complementarity matrix

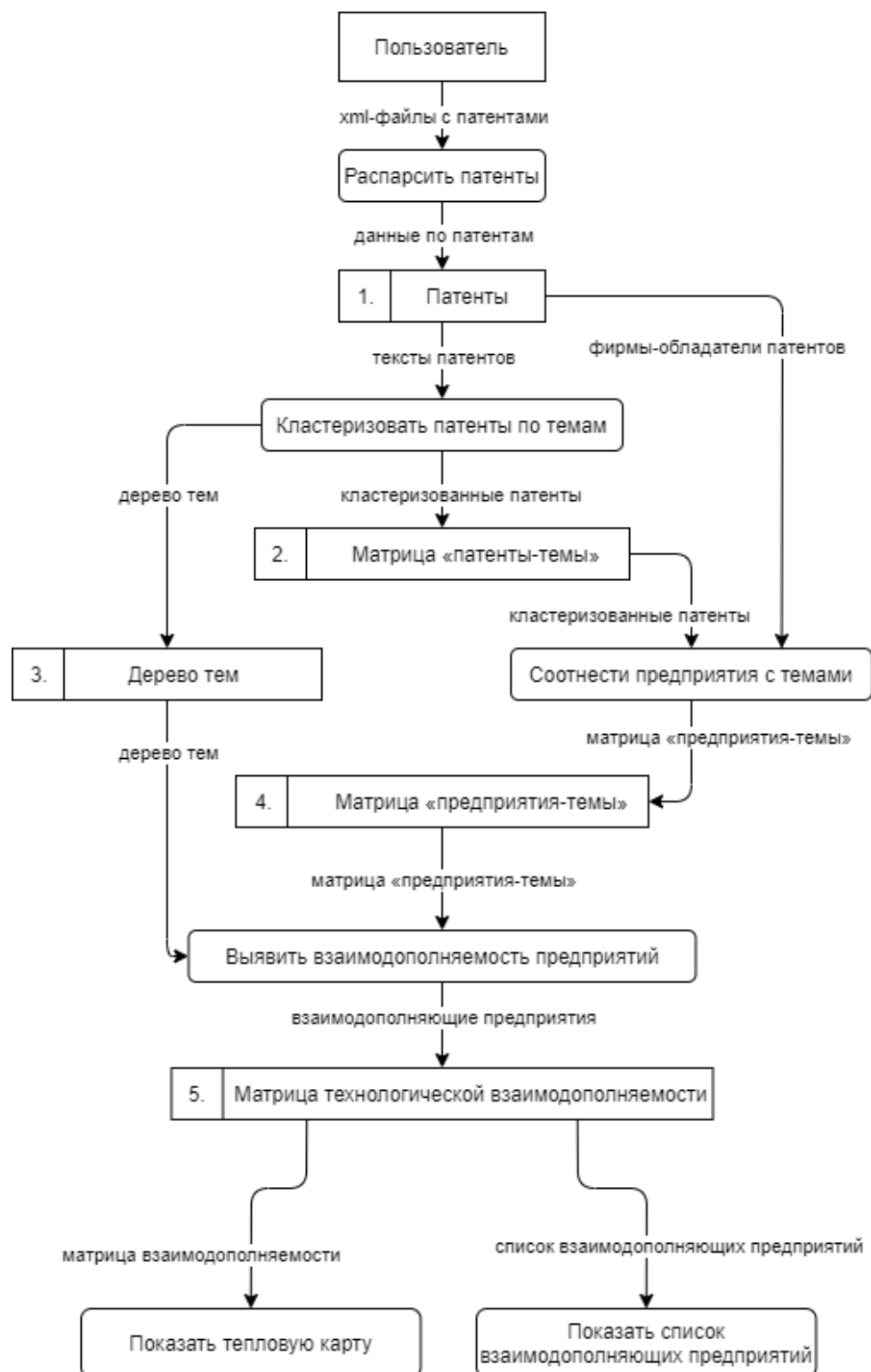


Рисунок 11 – Диаграмма потоков данных  
Figure 11 – Data flow diagram

### Архитектура системы

Серверная часть реализована на языке Python с помощью Django – фреймворка для веб-приложений на языке Python. Для извлечения данных из файлов с патентами было решено выбрать библиотеку BeautifulSoup. Для предобработки данных использовалась библиотека nltk [12]. Для кластеризации текстов патентов была выбрана

библиотека hlda [11], так как это единственная реализация hlda для python в свободном доступе. Для ускорения работы она была модифицирована. Для фронтенда был выбран JavaScript по причине того, что это самый распространённый язык для данного применения.

Парсер патентов должен принимать на вход один или несколько путей файлов патентов или папок с ними. Файлы патентов должны быть в формате USPTO.

На выходе парсер должен выводить в папку parser\_output:

- в подпапку data текст полей abstract, claims, description, все три поля в один файл, данные по каждому патенту в отдельный файл. Имя каждого файла — уникальный номер патента;

- в подпапку info файл или несколько файлов с информацией по патентам (Рисунок 12);

- в подпапку firms файл или несколько файлов с информацией по предприятиям, публикующим патенты (Рисунок 12).



Файл с информацией по патентам

Файл с информацией по фирмам-владельцам патентов

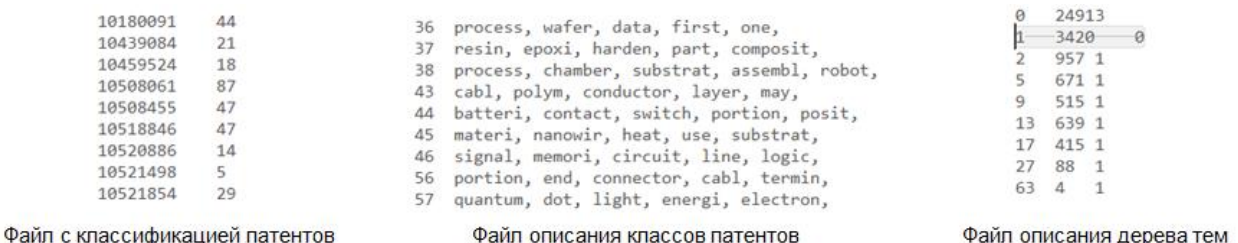
Рисунок 12 – Данные парсера  
Figure 12 – Parsing data

Модуль кластеризации патентов принимает на вход путь к папке с файлами текстов патентов и набор гиперпараметров алгоритма. На выходе он выдает три файла в папку cluster\_output:

- файл с классификацией патентов. Формат файла: каждый патент на отдельной строке, в которой указан id патента и номер соответствующей ему темы через знак табуляции (Рисунок 13);

- файл с описанием тем. Формат файла: каждая тема на отдельной строке, в которой указан номер темы и ключевые слова данной темы (Рисунок 13);

- файл с описанием дерева тем. Формат файла: пара номеров тем, разделенная символом табуляции, каждая пара на отдельной строке. Первое число из пары – номер темы-родителя темы со вторым номером (Рисунок 13).



Файл с классификацией патентов

Файл описания классов патентов

Файл описания дерева тем

Рисунок 13 – Данные кластеризации  
Figure 13 – Clustering data

Входные данные модуля анализа – данные, выданные парсером и кластеризатором. Выходные данные модуля – матрица взаимодополняемости (Рисунок 14).

Матрица «предприятия — темы» и матрица взаимодополняемости хранятся в нетекстовых файлах формата пру.

```
OMRON Corporation 0
Hewlett-Packard Development Company, L.P. 1
Leica Geosystems AG 2
Columbia Insurance Company 3
Apple Inc. 4
J. Choo Limited 5
Makita Corporation 6
Samsung Electronics Co., Ltd. 7
Sunbeam Products, Inc. 8
Chef'n Corporation 9
Rubbermaid Incorporated 10
Black & Decker Inc. 11
Andreas Stihl AG & Co. KG 12
```

Файл с номерами фирм

0.0000	0.0172	0.5648	0.0000	0.0000
0.0000	0.0000	0.5596	0.0000	0.0000
0.0000	0.0172	0.5648	0.0000	0.0000
0.0000	0.0000	0.5719	0.0000	0.0000
0.0000	0.0000	0.3179	0.0000	0.0000
0.0000	0.0000	0.5683	0.0000	0.0000
0.0000	0.0172	0.5648	0.0000	0.0000
0.0000	0.0172	0.5585	0.0000	0.0000
0.0000	0.0172	0.5588	0.0000	0.0000
0.0000	0.0172	0.5588	0.0000	0.0000

Фрагмент матрицы взаимодополняемости

```
{'3M Innovative Properties Company': {4: 1,
5: 1,
12: 5,
13: 4,
20: 6,
21: 5,
30: 1,
31: 3,
33: 9,
35: 4,
41: 2,
44: 1,
48: 4,
54: 1,
56: 1,
69: 1,
87: 1},
'ABB Technology AG': {9: 2,
13: 1,
17: 1,
20: 2,
29: 9,
42: 1,
44: 3,
56: 2,
76: 1},
```

Фрагмент матрицы «предприятие — темы»

Рисунок 14 – Данные анализа  
Figure 14 – Analysis data

### Результаты

Разработан программный модуль для кластеризации патентных документов USPTO на основе метода hLDA и выявления технологической взаимодополняемости предприятий на основе сравнения кластерной информации

Входные данные программы – xml файлы с патентными данными USPTO, зарегистрированными в промежутке времени между 2010 и 2020 годами, а также параметры алгоритма кластеризации (alpha, gamma, eta, количество итераций, высота дерева) и фильтрации (патентный класс, по которому отбираются патенты), порог количества патентов.

Выходные данные программы – представление технологической взаимодополняемости предприятий в виде тепловой карты (Рисунок 15) и в виде списка наиболее взаимодополняющих предприятий для выбранного (Рисунок 16).

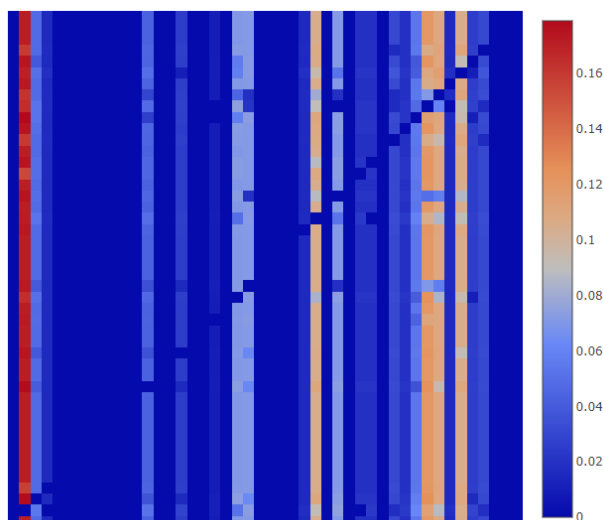


Рисунок 15 – Пример получаемой тепловой карты  
Figure 15 – Example of the resulting heat map

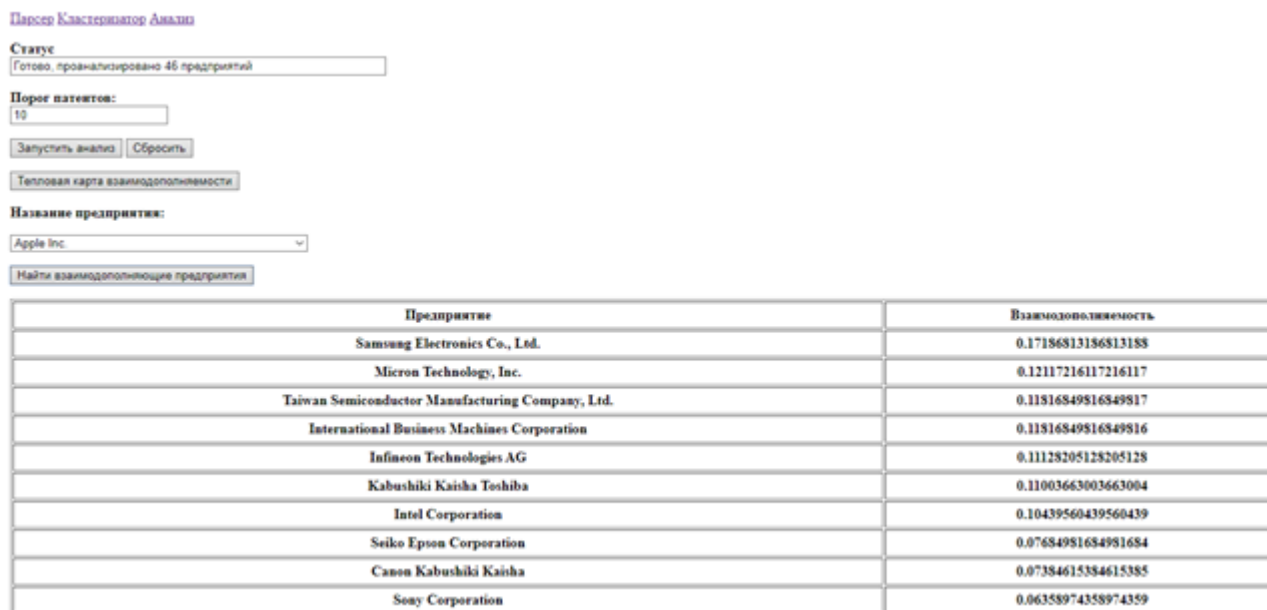


Рисунок 16 – Таблица наиболее взаимодополняющих предприятий  
Figure 16 – Table of the enterprises with the highest degree of complementarity

### Обсуждение

Теоретическая ценность заключается в разработанных алгоритмах парсинга патентов; кластеризации патентного массива на основе метода hLDA (изменения в классическом алгоритме позволили ускорить соответствующую функцию примерно в 12 раз, а весь модуль кластеризации – примерно на 30 %); формирования матрицы «Предприятия – Кластеры (Темы)»; визуализации связей предприятий в кластеризованном патентном пространстве.

Практическая значимость работы заключается в разработанном программном модуле выявления технологической взаимодополняемости предприятий. Данная программная разработка заняла второе место в Конкурсе научно-исследовательских работ студентов ВолГТУ 2021.

### Заключение

Целью данной работы являлась разработка программного модуля для кластеризации патентных документов USPTO и выявления технологической взаимодополняемости предприятий на основе сравнения кластерной информации. Определение количественной оценки взаимодополняемости технологий между предприятиями осуществляется на основе текстового анализа патентных документов. Иерархическая латентная тематическая модель распределения Дирихле (hLDA) позволяет идентифицировать технологические темы, скрытые в патентных документах, вместе с иерархической структурой этих тем. Расчет взаимодополняемости ведется на основе учета технологических тем, количества патентов в данных темах, зарегистрированных предприятиями.

### СПИСОК ИСТОЧНИКОВ

1. Guo Y. Research on the Risk Prevention of Overseas M & A of Chinese Enterprises. *Proc. of the 2022 7th International Conference on Financial Innovation and Economic Development*. 2022. DOI: 10.2991/aebmr.k.220307.085.

2. Purba N. S., Nooraeni R. Using LDA for Innovation Topic of Technology : Quantum Dots Patent Analysis. *Proc. of the 1st International Conference on Statistics and Analytics*. 2020;1. DOI: 10.4108/eai.2-8-2019.2290336.
3. Li H., Tan R., Xu B., Huo J. Patent design around method based on AFD for umbrella-type patent strategy. *Journal of Mechanical Engineering*. 2016;52(5):1–11. DOI: 10.3901/JME.2016.05.001.
4. Wang M.-Y. Exploring potential R&D collaborators with complementary technologies: the case of biosensors. *Technological Forecasting and Social Change*. 2012;79(5):862–874. DOI: 10.1016/j.techfore.2011.11.002.
5. Wang X., Li R., Liu Y., Lei M. A Morphology-Driven Method for Measuring Technology Complementarity: Empirical Study Involving Alzheimer's Disease. *Journal of Data and Information Science*. 2022;7(3):20–48. DOI: 10.2478/jdis-2022-001.
6. Borodin N., Korobkin D., Bezruchenko A., Fomenkov S. The search for R&D partners based on patent data. *Journal of Physics Conference Series*. 2021;2060(1):012022. DOI: 10.1088/1742-6596/2060/1/012022.
7. Sun L., Xu W. Enterprise Complementarity Based on Patent Information. *Wireless Communications and Mobile Computing*. 2022;2022(6):1–7. DOI: 10.1155/2022/5797285.
8. Wang X., Qiao Y., Hou Y., Zhang S., Han X. Measuring Technology Complementarity Between Enterprises With an hLDA Topic Model. *IEEE Transactions on Engineering Management*. 2021;68(5):1309–1320. DOI: 10.1109/TEM.2019.2958113.
9. Patents USPTO. United States Patent and Trademark Office. Доступно по: <https://www.uspto.gov/patents> (дата обращения: 07.11.2022).
10. Korobkin D., Fomenkov S., Fomenkova M., Vayngolts I., Kravets A. The Software for Computation the Criteria-Based Assessments of the Morphological Features of Technical Systems. *Cyber-Physical Systems*. 2021;350:161–172. DOI: 10.1007/978-3-030-67892-0\_14.
11. Gibbs sampler for the Hierarchical Latent Dirichlet Allocation topic model. GitHub. Доступно по: <https://github.com/joewandy/hlda> (дата обращения: 07.11.2022).
12. Natural Language Toolkit documentation. NLTK Project. Доступно по: <https://www.nltk.org> (дата обращения: 07.11.2022).

## REFERENCES

1. Guo Y. Research on the Risk Prevention of Overseas M & A of Chinese Enterprises. *Proc. of the 2022 7th International Conference on Financial Innovation and Economic Development*. 2022. DOI: 10.2991/aebmr.k.220307.085.
2. Purba N. S., Nooraeni R. Using LDA for Innovation Topic of Technology : Quantum Dots Patent Analysis. *Proc. of the 1st International Conference on Statistics and Analytics*. 2020;1. DOI: 10.4108/eai.2-8-2019.2290336.
3. Li H., Tan R., Xu B., Huo J. Patent design around method based on AFD for umbrella-type patent strategy. *Journal of Mechanical Engineering*. 2016;52(5):1–11. DOI: 10.3901/JME.2016.05.001.

4. Wang M.-Y. Exploring potential R&D collaborators with complementary technologies: the case of biosensors. *Technological Forecasting and Social Change*. 2012;79(5):862–874. DOI: 10.1016/j.techfore.2011.11.002.
5. Wang X., Li R., Liu Y., Lei M. A Morphology-Driven Method for Measuring Technology Complementarity: Empirical Study Involving Alzheimer's Disease. *Journal of Data and Information Science*. 2022;7(3):20–48. DOI: 10.2478/jdis-2022-001.
6. Borodin N., Korobkin D., Bezruchenko A., Fomenkov S. The search for R&D partners based on patent data. *Journal of Physics Conference Series*. 2021;2060(1):012022. DOI: 10.1088/1742-6596/2060/1/012022.
7. Sun L., Xu W. Enterprise Complementarity Based on Patent Information. *Wireless Communications and Mobile Computing*. 2022;2022(6):1–7. DOI: 10.1155/2022/5797285.
8. Wang X., Qiao Y., Hou Y., Zhang S., Han X. Measuring Technology Complementarity Between Enterprises With an hLDA Topic Model. *IEEE Transactions on Engineering Management*. 2021;68(5):1309–1320. DOI: 10.1109/TEM.2019.2958113.
9. Patents USPTO. United States Patent and Trademark Office. Available from: <https://www.uspto.gov/patents> (accessed on: 07.11.2022).
10. Korobkin D., Fomenkov S., Fomenkova M., Vayngolts I., Kravets A. The Software for Computation the Criteria-Based Assessments of the Morphological Features of Technical Systems. *Cyber-Physical Systems*. 2021;350:161–172. DOI: 10.1007/978-3-030-67892-0\_14.
11. Gibbs sampler for the Hierarchical Latent Dirichlet Allocation topic model. GitHub. Available from: [github.com/joewandy/hlda](https://github.com/joewandy/hlda) (accessed on: 07.11.2022).
12. Natural Language Toolkit documentation. NLTK Project. Available from: <https://www.nltk.org> (accessed on: 07.11.2022).

## ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

**Коробкин Дмитрий Михайлович**, кандидат технических наук, доцент кафедры САПРиПК, Волгоградский государственный технический университет, Волгоград, Российская Федерация.

*e-mail*: [dkorobkin80@mail.ru](mailto:dkorobkin80@mail.ru)

ORCID: [0000-0002-4684-1011](https://orcid.org/0000-0002-4684-1011)

**Dmitriy Mikhaylovich Korobkin**, Candidate of Technical Sciences, Assistant Professor at CAD Department, Volgograd State Technical University, Volgograd, Russian Federation.

**Безрученко Алексей Юрьевич**, магистрант, Волгоградский государственный технический университет, Волгоград, Российская Федерация.

*e-mail*: [alexei.bezruchenko@yandex.ru](mailto:alexei.bezruchenko@yandex.ru)

**Alexey Yurievich Bezruchenko**, Master's Student, Volgograd State Technical University, Volgograd, Russian Federation.

**Фоменков Сергей Алексеевич**, доктор технических наук, профессор кафедры САПРиПК, Волгоградский государственный технический университет, Волгоград, Российская Федерация.

*e-mail*: [saf550@yandex.ru](mailto:saf550@yandex.ru)

**Sergey Alekseevich Fomenkov**, Doctor of Technical Sciences, Professor at CAD Department, Volgograd State Technical University, Volgograd, Russian Federation.



**Колесников Сергей Григорьевич**, старший научный сотрудник кафедры САПриПК, Волгоградский государственный технический университет, Волгоград, Российская Федерация.

*e-mail:* [sk375@bk.ru](mailto:sk375@bk.ru)

**Kolesnikov Sergey Grigorievich**, Senior Researcher at CAD Department, Volgograd State Technical University, Volgograd, Russian Federation.

*Статья поступила в редакцию 08.11.2022; одобрена после рецензирования 30.11.2022; принята к публикации 13.12.2022.*

*The article was submitted 08.11.2022; approved after reviewing 30.11.2022; accepted for publication 13.12.2022.*