

## Выявление актуальных тем исследований по вопросам энергетической политики и безопасности в публикациях издательств Elsevier и MDPI за 2021–2023 гг.

### Часть 1. Кластеризация терминов и реферирование тем с помощью ChatGPT

**Б.Н. Чигарев**

Институт проблем нефти и газа РАН, г. Москва, Россия

E-mail: [bchigarev@ipng.ru](mailto:bchigarev@ipng.ru)

**Аннотация.** Данная статья посвящена выявлению актуальных тем исследований в публикациях за 2021–2023 гг. издательств Elsevier и MDPI по вопросам энергетической политики и безопасности. Проводимое исследование основывалось на выявлении ключевых слов/фраз из текстов заголовков и аннотаций, с использованием программы PatternRank, базирующейся на технологии KeyBERT и извлечении именных групп. Данные ключевые слова/фразы рассматривались как возможная альтернатива авторским ключевым словам. Библиометрический анализ таких терминов позволил выявить, используя программу VOSviewer, восемь кластеров, детально описывающих тематику энергетической политики и безопасности. Текстовое описание тематики отдельных кластеров дано с использованием технологии ChartGPT.

**Ключевые слова:** энергетическая политика, темы исследований, библиометрический анализ, KeyBERT, ChartGPT, VOSviewer.

**Для цитирования:** Чигарев Б.Н. Выявление актуальных тем исследований по вопросам энергетической политики и безопасности в публикациях издательств Elsevier и MDPI за 2021–2023 гг. Часть 1. Кластеризация терминов и реферирование тем с помощью ChatGPT// Актуальные проблемы нефти и газа. 2023. Вып. 2(41). С. 63–83. <https://doi.org/10.29222/ipng.2078-5712.2023-41.art5>

#### Введение

Цели данной публикации:

- Выявление актуальных тем исследований в публикациях за 2021–2023 гг. издательств Elsevier и MDPI по вопросам энергетической политики и безопасности.
- Демонстрация возможности извлечения ключевых слов и ключевых фраз, наиболее близко описывающих документ с использованием векторного представления слов KeyBERT<sup>1</sup> для использования их как альтернативы авторским ключевым словам.

#### Актуальность

##### проводимого исследования

Изучение зарубежных научных публикаций способствует более полному пониманию конкретной темы за счет использования различных точек зрения, подходов и интересов разных стран и крупных экономик. Это может способствовать разработке инновационных решений сложных задач энергетической политики и безопасности [1, 2].

<sup>1</sup> <https://maartengr.github.io/KeyBERT/> – домашняя страница проекта KeyBERT.

Вопросы энергетической политики и безопасности имеют решающее значение для стабильности и устойчивости энергетической системы и, в конечном счете, экономики. Они становятся вопросом национальной безопасности и внешней политики, когда происходят значительные изменения в мировой экономике и политике [3, 4].

Выделение ключевых слов в тексте позволяет быстро определить ключевые вопросы и темы в тексте. Использование ключевых слов в качестве поисковых терминов особенно важно в научных исследованиях, где необходимо обрабатывать большие объемы информации. К сожалению, авторские ключевые слова часто бывают недоступны при экспорте библиометрических данных из реферативных баз данных с открытым доступом.

Например, при экспорте данных из базы The Lens по теме «Энергетическая политика» (Year Published = (2020–) Field of Study = (Energy policy), Scholarly Works 4479, актуально на 25.03.2023) из 4479 записей только 220 содержат ключевые слова в поле Keywords.

Актуальность извлечения ключевых фраз и слов из текстов, в том числе из научных публикаций, хорошо представлена в работах [5–7]. Более того, ряд авторов использует лингвистический анализ текстов и векторное представление слов и предложений для выделения коротких предложений, используемых для аннотирования научных статей, примером может служить работа [8].

KeyBERT, метод машинного обучения, был применен в статье [9]

для извлечения ключевых слов из достоверных и ложных новостей. На следующем этапе эти ключевые слова были использованы для получения актуальных поисковых запросов в Google Trends API. Примечание: файлы RSS<sup>2</sup> являются хорошим примером текстов, содержащих заголовки, тело новости и ссылку на источник, но не содержащих список ключевых слов.

Файлы RSS удобны для сбора информации по политическим и экономическим вопросам с сайтов аналитических агентств.

Однако научных публикаций по использованию KeyBERT для извлечения ключевых фраз/слов и последующего применения VOSviewer, программы для построения и визуализации библиометрических сетей, в частности кластеризации ключевых слов на основе их совместной встречаемости, найти не удалось. Само определение актуальных тем исследований на основе совместной встречаемости ключевых слов широко используется, хорошим примером может служить работа [10].

Вышесказанное определило научную новизну данной статьи, заключающуюся в выявлении актуальных исследовательских тем по вопросам энергетической политики и безопасности по публикациям за 2021–2023 гг. издательств Elsevier и MDPI с использованием технологии KeyBERT для извлечения ключевых фраз/слов из заголовков и аннотаций статей и последующей кластеризации этих терминов для описания актуальных тем исследований.

---

<sup>2</sup> <https://ru.wikipedia.org/wiki/RSS> – RSS (англ. Rich Site Summary – обогащенная сводка сайта).

### Материалы и методы

*Библиометрические данные* по вопросам энергетической политики и безопасности за 2021–2023 гг. были собраны из двух источников: ScienceDirect издательства Elsevier и сайта mdpi.com издательства MDPI. Данные актуальны на 14.03.2023. Данные скачивались в формате RIS, далее переводились в формат CSV и объединялись в один файл. Использовались два запроса: “energy policy” и “energy security”. Запросы не объединялись, что позволило сравнить результаты каждого запроса. Использовались данные только для двух типов материалов: Review articles и Research articles. Для данных ScienceDirect их соотношения составили 111 и 1194 статьи для запроса “energy policy” и 92 и 612 для запроса “energy security”, соответственно. Для данных из издательства MDPI получены 523 результата для “energy policy” и 433 для “energy security”. В обоих случаях поиск проводился по полям: Title, Abstract, Keywords. Указанные цифры не вычислялись по экспортированным библиометрическим данным, а брались непосредственно с используемых платформ. Более детальные характеристики библиометрических данных, собранных по указанным запросам представлены в Прил. 1.

Для работы по извлечению ключевых слов из текстов заголовков и аннотаций все данные объединялись, а дублирующие записи удалялись. Итоговое число записей, которое использовалось в работе, составило 2612 строки.

Существенным преимуществом библиометрических данных ScienceDirect и

MDPI является высокая наполненность поля Author keywords, что позволило провести сравнение сгенерированных ключевых слов и авторских ключевых слов.

*Методика выявления ключевых слов/фраз из текстов заголовков и аннотаций* базировалась на использовании BERT (Bidirectional Encoder Representations from Transformers) – модели обработки естественного языка, предложенной исследователями из Google Research в 2018 г. [11]. Наиболее часто для этого используется KeyBERT, простая в использовании техника извлечения ключевых слов, основанная на BERT для создания списка ключевых слов и ключевых фраз, наиболее схожих с документом [12].

Однако, как показано в работе [13], KeyBERT позволяет указать диапазон n-грамм<sup>3</sup> для определения длины извлекаемых ключевых фраз, но при этом возникают две проблемы: пользователи обычно не знают оптимального диапазона n-грамм, а возвращаемые ключевые фразы иногда грамматически неправильны. Для устранения указанных проблем авторы предлагают предварительно извлекать именные словосочетания и в дальнейшем использовать их в качестве кандидатов для выбора ключевых фраз, наиболее хорошо согласующихся с текстом. Авторы разработали технику PatternRank для извлечения ключевых фраз с помощью KeyphraseVectorizers и KeyBERT.

---

<sup>3</sup> N-грамма – последовательность из n смежных элементов. С семантической точки зрения это может быть последовательность слогов, слов или букв.

Именно PatternRank использовался в данном исследовании для извлечения ключевых фраз из текстов заголовков и аннотаций публикаций, относящихся к тематике энергетической политики и безопасности.

Прежде чем была проведена процедура извлечения ключевых фраз из текстов, последние подвергались предобработке. Удалялись сокращения и пояснения, приведенные в скобках, например, Internet of Things (IoT) или methane (CH<sub>4</sub>). Удалялись или замещались кавычки, непечатаемые символы и т. д.

В csv-файлах, содержащих собранные библиометрические данные, названия полей заменялись на названия, используемые в файлах, экспортируемых из Scopus. А извлеченные ключевые слова и фразы размещались в поле Index Keywords.

Подготовленные таким образом табличные данные использовались в программе VOSviewer [14] для построения сети ключевых слов и их кластеризации.

Полученные кластеры ключевых слов рассматривались как описание тем публикаций.

По составленным таблицам ключевых слов для каждого кластера с помощью сервиса ChartGPT генерировался текст, относящийся к данному набору ключевых слов. Полученный текст использовался как поясняющий тематику данного кластера.

Примечание: в данной статье термины «ключевые слова» и «ключевые фразы» используются как синонимы, так как и в авторских, и в сгенерированных ключевых терминах они могут встречаться как состоящие из одного слова, так и из нескольких слов.

*Пояснение:* в разделах «Сгенерированное и переведенное...»

*приводится фактически подстрочный перевод с английского сгенерированных текстов с попыткой сохранить баланс читаемости и минимального искажения сгенерированного текста на английском языке. Поэтому текст не подвергался литературному и смысловому редактированию.*

## Результаты и обсуждения

### Результаты извлечения ключевых слов с использованием PatternRank

PatternRank для извлечения ключевых слов был выбран не сразу. Предварительно проверялась возможность использования TextBlob [15], Yake! [16], Sifaka [17] и RaKUn 2.0 [18]. Сравнение носило сугубо оценочный характер и не может претендовать на детальное исследование. По субъективному мнению автора, Yake! и Sifaka показали хорошие результаты, но в данном исследовании ставилась, в том числе, задача использования концепции векторного представления слов, а в указанных двух программах данный подход не используют. Техническая сторона данного исследования проводилась в стиле проверки концепции (proof of concept) и демонстрации практической осуществимости того, что технология использования векторного представления слов целесообразна для применения в библиометрических исследованиях.

Удобство использования PatternRank заключалось также в том, что сразу генерировался список ключевых слов для каждой записи, состоящей из заголовка и аннотации отдельной публикации. В качестве примера приведем списки ключевых слов/фраз для ряда публикаций:

- electricity program effectiveness, electricity program, PV capacity, renewable energy, PV installation;

- renewable electricity, energy development, wind energy development, electricity bills, wind energy;

- energy planning, emission buildings, carbon cities, evaluation methodologies, energy exploitation;

- Ukrainian energy sector, energy efficiency gap, energy efficiency gaps, energy investments, green energy investments;

- municipal waste recycling, municipal waste, CO<sub>2</sub> emissions, sustainable energy policies, sustainable energy;

- urban heat islands effect, urban green resources, green infrastructure, climate, urban areas;

- green energy technologies, green energy deployment, renewable energy development, renewable energy, green technological development;

- local energy system, local energy, energy system, climate change mitigation, decomposition model;

- social distancing policies, strict social distancing policies, energy policies, electricity consumption behaviors, electricity demands;

- German photovoltaic auctions; auction theory modelling, auction theory, optimal bidding strategy, bidding strategies.

Явным преимуществом сгенерированных ключевых слов является то, что выборочная проверка показала их встречаемость в текстах заголовков и аннотаций, в то время как авторские ключевые слова не обязательно встречаются в этих текстах. Например, в статье [19] в списке авторских ключевых слов есть термины “neural networks” и “machine learning” – их нет ни в заголовке, ни в аннотации, но они встречаются в полном

тексте статьи. Как отмечалось во введении, авторские ключевые слова часто отсутствуют в записях открытых реферативных баз и статьи приходится искать по терминам, встречающимся в заголовках и аннотациях.

Сгенерированные ключевые слова, представленные как Index keywords в таблице библиометрических данных в формате Scopus, при использовании их в программе VOSviewer дают визуализацию сети совместной встречаемости ключевых терминов, представленную на рис. 1.

Для сравнения сеть и кластеризация авторских ключевых слов представлена в Прил. 2.

Сравнение некоторых параметров сгенерированных и авторских ключевых слов:

- общее число уникальных терминов: 7832 и 7238;

- число полученных кластеров при одинаковых параметрах VOSviewer: 8 и 9;

- число одинаковых терминов в списках из 5000 с наибольшим числом общих связей: 755;

- число одинаковых терминов в списках из 250 с наибольшим числом общих связей: 95

- число сгенерированных терминов из 250, содержащихся в списке из 5000 авторских ключевых слов: 188.

- показатель в степенном законе распределения терминов, соответственно – 2,495 и 2,493.

Из приведенных данных следует, что, несмотря на некоторое сходство в статистических показателях обоих типов терминов, списки самих терминов существенно различаются.

Список 20 сгенерированных слов, имеющих максимальную общую силу связи, представлен в табл. 1.

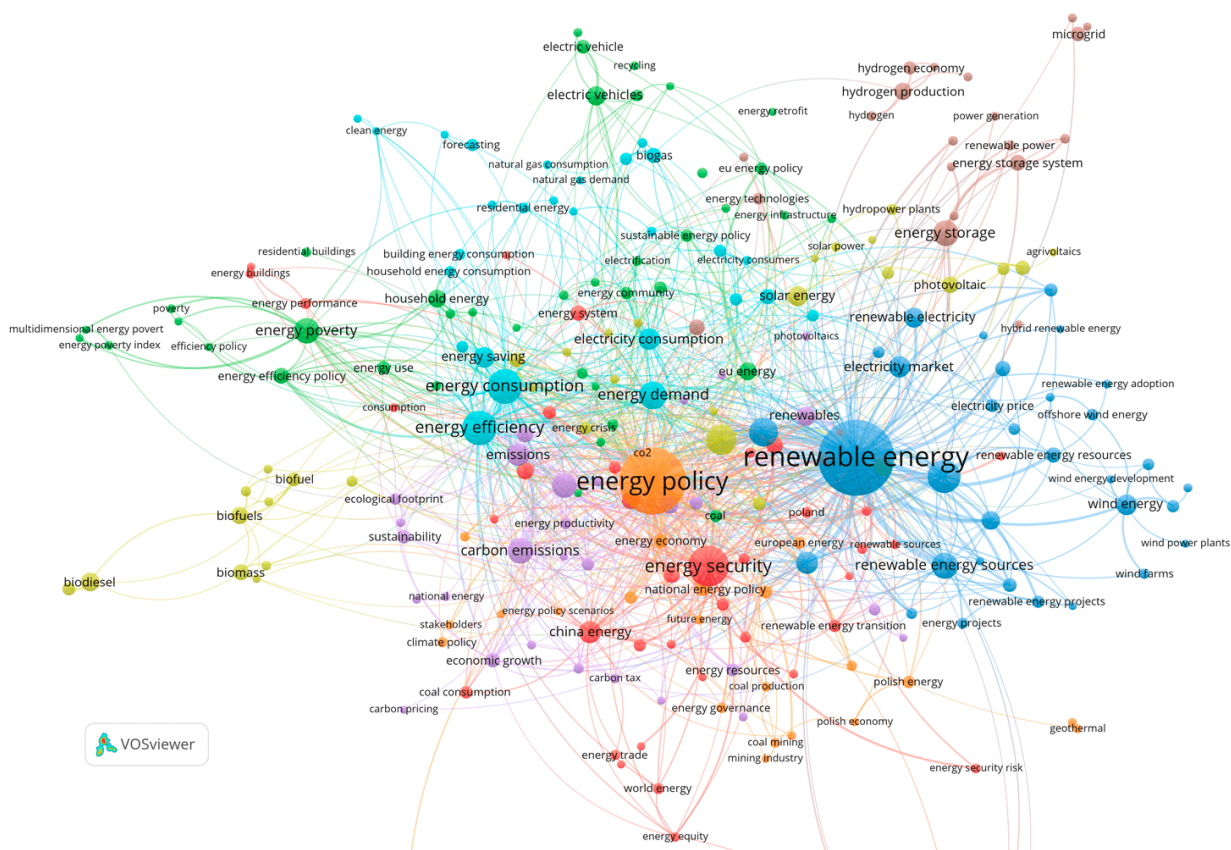


Рис. 1. Сеть и кластеризация сгенерированных ключевых слов

Параметры: общее число уникальных терминов – 7832, из них 250 встречаются 5 или более раз; минимальное число терминов в кластере – 20, число кластеров – 8

Таблица 1

Сгенерированные ключевые слова, имеющие максимальное значение total link strength

KWs	Occur	TLS	KWs	Occur	TLS
renewable energy	412	732	renewable energy consumption	35	81
energy policy	323	470	emissions	39	75
energy consumption	87	167	co2 emissions	46	73
energy security	118	160	energy poverty	48	73
sustainable energy	67	140	carbon emissions	48	65
energy sector	61	129	energy storage	47	65
renewable energy policy	79	128	energy development	25	57
energy efficiency	85	114	renewable energy development	26	57
renewable energy sources	49	112	china energy	35	53
energy demand	55	94	energy production	23	53

Сокращения: KWs – название ключевого термина;  
 Occur (occurrences) – частота встречаемости ключевого термина;  
 TLS (total link strength) – общая сила связи.

Слова с максимальным значением общей силы связей должны чаще совместно встречаться. Запрос: «“renewable energy” AND “energy policy” AND “energy consumption” AND “energy security” AND “sustainable energy”» к платформе ScienceDirect дает 2941 результатов. Для сравнения – 5 авторских ключевых слов с максимальным значением общей силы связей: «“energy policy” AND “renewable energy” AND “energy security” AND “energy transition” AND “energy efficiency”» при запросе к платформе ScienceDirect дает 1432 результата. Ключевые слова: “energy policy”, “renewable energy” и “energy security” являются общими в двух запросах, значит “energy consumption” и “sustainable energy” из списка сгенерированных ключевых слов лучше сочетаются с общим контекстом, чем авторские ключевые слова: “energy transition” и “energy efficiency”.

Платформа ScienceDirect ограничивает число ключевых фраз в запросе до 8.

По запросу: «“renewable energy” AND “energy policy” AND “energy consumption” AND “energy security” AND “sustainable energy” AND “energy sector” AND “renewable energy policy” AND “energy efficiency”» получено 384 результатов, а по запросу из списка авторских ключевых слов: «“energy policy” AND “renewable energy” AND “energy security” AND “energy transition” AND “energy efficiency” AND “energy resources” AND “climate change” AND “sustainability”» – 689 результатов.

Таким образом, однозначного преимущества использования в запросах сгенерированных или авторских ключевых слов не наблюдается. Но тот факт, что 5 и 8 ключевых слов в запросе дают достаточное количество результатов, говорит об их соответствии одной тематике.

### **Тематика отдельных кластеров из сгенерированных ключевых слов**

Примечание: учитывая, что целью данной публикации служит выявление актуальных тем исследований в публикациях за 2021–2023 гг. издательств Elsevier и MDPI по вопросам энергетической политики и безопасности, в данном разделе представлены переводы ключевых терминов, а термины на английском языке, полученные кластеризацией с использованием программы VOSviewer, даны в Прил. 3. После приведения списка ключевых терминов и их встречаемости, дано краткое описание тематики кластера, сгенерированное с использованием ChartGPT и переведенное на русский язык. Данный подход позволяет показать, как тематика кластера отражается в широком социально политическом контексте, на котором обучен ChartGPT. Учитывая быстрое развитие технологии ChartGPT, отметим, что приведенные результаты актуальны на 21.02.2023.

#### **Первый кластер**

*Ключевые термины и их встречаемость:* энергетическая безопасность (118), энергетика Китая (35), рынок энергетики (23), выбросы парниковых газов (22), ядерная энергетика (21), планирование в энергетике (19), энергетическая система (17), энергетика (13), управление в энергетике (12), многокритериальные методы принятия решений (12), потребление угля (11), глобальная энергетика (11), международная энергетика (11), переход к возобновляемым источникам энергии (11), энергетическая промышленность (9), энергетические показатели (9), ядерная энергетика (9), Польша (9), устойчивый энергетический переход (9), мировая энергетика (9),

источники энергии (8), торговля энергией (8), энергетический переход (8), производство электроэнергии (7), энергетика зданий (7), энергетическая справедливость (7), риски энергетической безопасности (7), потребление (6), показатели энергетической безопасности (6), выбросы газа (6), мощности возобновляемой энергетики (6), возобновляемые источники (6), устойчивая энергетическая безопасность (6), критерии принятия решения (5), критериальный анализ принятия решения (5), энергетическая независимость (5), индекс энергетической безопасности (5), энергоснабжение (5), глобальная энергетическая безопасность (5), глобальная торговля энергией (5), атомные электростанции (5), здания с нулевым энергопотреблением (5).

#### **Сгенерированное и переведенное описание тематики первого кластера**

Энергетическая безопасность относится к числу важнейших приоритетов для стран всего мира. Китай является крупным мировым игроком на энергетическом рынке и вкладывает значительные средства в строительство атомных электростанций. Однако обеспокоенность по поводу прозрачности и подотчетности в энергетическом секторе Китая порождает вопросы о показателях энергетической безопасности страны. Эффективное планирование и управление энергетикой имеет важное значение для обеспечения устойчивого энергетического перехода и сокращения выбросов парниковых газов. Многокритериальный анализ решений может быть использован для оценки оптимального энергетического баланса для данного региона с учетом таких факторов, как мощность возобновляемых

источников энергии, потребление угля и энергоэффективность зданий. Польша является ярким примером страны, стремящейся к энергетической независимости и работающей над переходом на возобновляемые источники энергии. Здания с нулевым энергопотреблением являются инновационным решением для снижения энергопотребления и обеспечения устойчивой энергетической безопасности.

#### **Второй кластер**

*Ключевые термины и их встречаемость:* энергетическая бедность (48), электромобили (27), бытовая энергетика (24), энергетика Евросоюза (23), политика энергоэффективности (18), электромобиль (14), использование энергии (13), уголь (12), политика устойчивой энергетики (12), энергетическое сообщество (11), зеленая энергетика (11), энергетическая политика ЕС (10), выбросы парниковых газов (10), энергетическая справедливость (9), энергетическая политика (9), энергетические инновации (8), электрификация (7), энергетические инициативы (7), бытовое энергопотребление (7), жилые дома (7), сельская энергетика (7), общинная энергетика (6), модели энергопотребления (6), энергетическая демократия (6), энергетическая инфраструктура (6), индекс энергетической бедности (6), климат ЕС (6), зарядка электромобилей (6), бедность (6), переработка (6), возобновляемая генерация (6), бытовая энергетика (5), эффективная политика (5), сокращение выбросов (5), модернизация энергетики (5), топливная бедность (5), многоплановая энергетическая бедность (5), многоплановый индекс энергетической бедности (5), электрификация сельских районов (5), умная энергетика (5).



**Сгенерированное и переведенное описание тематики второго кластера**

Энергетическая бедность является серьезной проблемой, затрагивающей миллионы людей во всем мире, особенно в развивающихся странах. Для решения этой проблемы и обеспечения энергетической справедливости необходимы политические меры и инициативы в области устойчивой энергетики. Электрификация транспорта с помощью электромобилей (EV) является важной частью сокращения выбросов парниковых газов и продвижения политики устойчивой энергетики. Политика энергоэффективности может помочь снизить потребление энергии и способствовать сокращению выбросов, особенно в жилых домах. Рециклинг и производство возобновляемой энергии являются важными компонентами устойчивой энергетической инфраструктуры. Многомерный индекс энергетической бедности может помочь измерить и решить проблему энергетической нищеты. Электрификация сельских районов и инициативы в области коммунальной энергетики позволяют развивать низкоуглеродную энергетику в районах с недостаточным уровнем сервиса. Энергетическая демократия способна содействовать расширению участия общественности в разработке энергетической политики и инноваций. Интеллектуальные энергетические решения помогают снизить энергопотребление домохозяйств и способствуют формированию устойчивых энергетических моделей.

**Третий кластер**

*Ключевые термины и их встречаемость:* возобновляемая энергетика (412), политика в области возобновляемой энергетики (79), энергетический сектор (61), возобновляемые источники энергии (49), потребление возобновляемой энергии (35), рынок электроэнергии (34), энергия ветра (31), возобновляемая электроэнергия (30), развитие возобновляемой энергетики (26), возобновляемые энергетические ресурсы (17), технологии возобновляемой энергетики (17), сектор возобновляемой энергетики (14), цена электроэнергии (13), производство возобновляемой энергии (13), проекты возобновляемой энергетики (12), энергия ветра (12), морская энергия ветра (10), региональная энергетика (10), системы возобновляемой энергетики (10), ветряная турбина (10), энергетические проекты (9), производство возобновляемой энергии (9), ветряные электростанции (8), морские ветряные электростанции (6), морская ветроэнергетика (6), внедрение возобновляемой энергетики (6), гибридная возобновляемая энергетика (5), рынок возобновляемой энергетики (5), использование возобновляемой энергии (5), устойчивое электроснабжение (5), развитие ветроэнергетики (5), ветряные электростанции (5).

**Сгенерированное и переведенное описание тематики третьего кластера**

Возобновляемая энергетика быстро развивается во всем мире, что обусловлено необходимостью сокращения выбросов парниковых газов и перехода к более устойчивой энергетической системе.

Многие страны проводят стратегию развития возобновляемой энергетики, направленную на использование возобновляемых источников энергии и стимулирование потребления возобновляемой энергии.

Рынок электроэнергии изменяется, стремясь приспособиться к возобновляемым источникам энергии, включая энергию ветра. Энергия ветра является одной из наиболее распространенных технологий использования возобновляемых источников энергии и применяется как в наземных, так и в морских ветропарках. Региональная энергетическая политика предусматривает развитие возобновляемой энергетики в конкретных районах. Системы возобновляемой энергетики становятся все более сложными, включая гибридные системы, сочетающие различные источники возобновляемой энергии. Рынок возобновляемых источников энергии растет в различных секторах экономики. Проекты по использованию возобновляемых источников энергии разрабатываются по всему миру, включая оффшорные ветряные электростанции.

#### **Четвертый кластер**

*Ключевые термины и их встречаемость:* устойчивая энергетика (67), солнечная энергия (29), биодизель (25), развитие энергетики (25), биотопливо (22), фотоэлектричество (20), биомасса (19), цена на энергию (16), солнечная фотоэлектрическая энергия (16), биотопливо (14), производство биодизеля (13), устойчивое развитие энергетики (13), энергетический кризис (11), гидроэнергетика (11), солнечные фотоэлектростанции (11), энергия биомассы (9), гидроэлектростанции (9), производство биотоплива (8), ресурсы биомассы (7), солнечная энергия (7), альтернативные виды топлива (6),

производство биоэнергии (6), продовольственная безопасность (6), цели устойчивого развития (6), сточные воды (6), агровольтаика (5), газификация биомассы (5), глобальный энергетический кризис (5), фотоэлектрическая энергия (5), устойчивое энергопотребление (5), энергия воды (5), безопасность воды (5).

#### **Сгенерированное и переведенное описание тематики четвертого кластера**

Использование чистой энергии необходимо для устойчивого будущего, а солнечная энергия и биодизель являются важными возобновляемыми источниками энергии. Биотопливо, включая биодизель, позволяет снизить выбросы парниковых газов от транспорта. Фотоэлектрические технологии широко используются для производства солнечной энергии и становятся все более доступными. Биомасса является другим важным возобновляемым источником энергии и может быть использована для производства энергии с помощью таких процессов, как газификация биомассы. Гидроэнергетика относится к зрелым технологиям производства возобновляемой энергии с помощью гидроэлектростанций. Устойчивое развитие энергетики необходимо для решения глобального энергетического кризиса и достижения целей устойчивого развития, включая продовольственную и водную безопасность. Агровольтаика — это инновационное решение, сочетающее производство солнечной энергии с сельскохозяйственной деятельностью для обеспечения устойчивого землепользования и повышения устойчивого потребления энергии. Сточные воды могут служить источником возобновляемой энергии благодаря таким процессам, как анаэробное брожение.

### Пятый кластер

*Ключевые термины и их встречаемость:* выбросы углерода (48), выбросы  $\text{CO}_2$  (46), выбросы (39), возобновляемые источники энергии (27), производство энергии (23), устойчивое развитие (20), экономический рост (15), устойчивость (14), углеродная нейтральность (13), энергетические ресурсы (13), выбросы углекислого газа (12), экологический след (10), инвестиции в энергетику (10), производительность энергии (9), авторегрессия (8),  $\text{CO}_2$  (8), улавливание углерода (7), налог на углерод (7), изменение климата (6), коинтеграция (6), распределенные энергоресурсы (6), экологическая устойчивость (6), национальная энергетика (6), фотовольтаика (6), переход к устойчивому развитию (6), цены на углерод (5), энергетические субсидии (5), инвестиции в возобновляемые источники энергии (5), стохастический анализ границ (5), потребление воды (5), водные ресурсы (5).

### Сгенерированное и переведенное описание тематики пятого кластера

Сокращение выбросов углерода, в том числе выбросов  $\text{CO}_2$ , имеет решающее значение для достижения устойчивого развития и борьбы с изменением климата. Использование возобновляемых источников энергии может помочь снизить выбросы углерода при производстве энергии, способствуя росту экономики и устойчивому развитию. Углеродная нейтральность, при которой выбросы углерода уравниваются поглощением углерода, также является важной целью устойчивого развития. Эффективное и рациональное использование энергии снижает воздействие на окружающую среду и

способствует экологической безопасности. Инвестиции в энергетику, включая инвестиции в возобновляемые источники энергии, способны ускорить переход к стабильному производству энергии.

Улавливание углерода и ценообразование, например, налоги на углерод, позволяют сократить выбросы углерода как в энергетическом секторе, так и в других секторах экономики.

Распределенные энергетические ресурсы, такие как солнечные фотоэлектрические батареи, помогают сократить выбросы углерода и способствуют переходу к устойчивому развитию. Использование воды и водных ресурсов должно учитываться при производстве энергии и реализации мер по достижению устойчивого развития.

### Шестой кластер

*Ключевые термины и их встречаемость:* потребление энергии (87), энергоэффективность (85), спрос на энергию (55), потребление электроэнергии (33), энергосбережение (32), биогаз (17), спрос на электроэнергию (17), производство биогаза (12), производство энергии (12), прогнозирование (12), производство электроэнергии (11), электроэнергетика (10), меры по энергоэффективности (10), потребление энергии в зданиях (9), бытовая энергетика (9), потребление энергии домохозяйствами (8), биогазовые установки (7), бытовая электроэнергия (7), потребители электроэнергии (6), использование энергии (6), потребление природного газа (6), потребление электроэнергии населением (6), чистая энергия (5), прогноз (5), домохозяйства (5), спрос на природный газ (5), потребление электроэнергии населением (5).

### **Сгенерированное и переведенное описание тематики шестого кластера**

Сокращение потребления энергии имеет решающее значение для устойчивого развития, а меры по повышению энергоэффективности снижают спрос на энергию. Потребление электроэнергии, вносящее основной вклад в энергопотребление, может быть снижено за счет мер по энергосбережению и внедрения чистых источников энергии. Производство биогаза позволяет снизить зависимость от потребления природного газа и получать чистую энергию. Прогнозы спроса и производства энергии используются для обоснования энергетической политики и инвестиционных решений. Сектор электроэнергетики, включающий потребителей электроэнергии и домашние хозяйства, необходимо учитывать при работе над снижением энергопотребления и стимулированием производства чистой энергии. Потребление энергии домашними хозяйствами, включая потребление энергии в зданиях, является важной областью для принятия мер по повышению энергоэффективности. Спрос на природный газ и его потребление следует учитывать в рамках усилий по продвижению чистой энергии и сокращению выбросов углерода.

### **Седьмой кластер**

*Ключевые термины и их встречаемость:* энергетическая политика (323), национальная энергетическая политика (16), экономика энергетики (14), энергетическая устойчивость (12), европейская энергетика (12), энергетическая политика будущего (12),

польская энергетика (12), климатическая политика (8), управление энергетикой (8), формирование энергетической политики (8), добыча угля (7), потребители энергии (7), энергия будущего (7), централизованное теплоснабжение (6), геотермальная (6), геотермальная энергия (6), новая энергетическая политика (6), заинтересованные стороны (6), меры энергетической политики (5), сценарии энергетической политики (5), горнодобывающая промышленность (5), польская экономика (5).

### **Сгенерированное и переведенное описание тематики седьмого кластера**

Энергетическая политика – это набор правил, норм и рекомендаций, определяющих как страна производит, распределяет и использует энергию. Она направлена на обеспечение энергетической безопасности, устойчивости и доступности энергии для всех граждан. Политика, способствующая повышению энергоэффективности и энергосбережению, помогает потребителям экономить деньги на счетах за электроэнергию и одновременно снижает углеродный след. Для эффективной и действенной реализации энергетической политики необходимо надлежащее управление энергетическим сектором. Заинтересованные стороны в энергетическом секторе включают энергетические компании, потребителей, разработчиков политики и экологические группы. Сценарии энергетической политики позволяют разработчикам политики определить потенциальное воздействие различных мер энергетической политики на горнодобывающую промышленность.

### Восьмой кластер

*Ключевые термины и их встречаемость:* хранение энергии (47), производство водорода (22), система хранения энергии (19), политика энергетического перехода (18), микросеть (16), водородная экономика (13), энергетические технологии (10), возобновляемая энергия (8), аккумуляторное хранение энергии (7), водород (7), возобновляемые источники энергии переменной мощности (7), технологии хранения энергии (6), производство электроэнергии (6), внедрение возобновляемых источников энергии (6), возобновляемое использование водорода (6), устойчивые энергетические системы (6), хранение тепловой энергии (6), производство зеленого водорода (5), возобновляемые ресурсы (5), интеллектуальные сети (5).

### Сгенерированное и переведенное описание тематики восьмого кластера

Хранение энергии приобретает все большее значение по мере продвижения мира к декарбонизированному и устойчивому будущему. Существует множество технологий хранения энергии, включая аккумуляторы, тепловое хранение энергии и производство водорода. Производство водорода является ключевой технологией для хранения энергии, поскольку водород можно производить из возобновляемых ресурсов, таких как энергия ветра и солнца, и использовать для выработки электроэнергии или в качестве топлива для транспорта. При производстве зеленого водорода используется возобновляемая энергия в процессе электролиза. Интеллектуальные сети и микросети повышают эффективность и надежность системы электроснабжения за счет интеграции возобновляемых источников энергии и технологий хранения энергии.

Политика энергетического перехода увеличивает использование возобновляемых источников энергии, стимулируя технологии возобновляемых источников энергии и системы хранения энергии.

### Заключение

Среди актуальных тем исследований в публикациях за 2021–2023 гг. издательств Elsevier и MDPI по вопросам энергетической политики и энергетической безопасности доминируют проблемы, связанные с внедрением возобновляемых источников энергии, изменения климата, устойчивого развития, энергетической бедности, регионального развития энергетики, методов хранения энергии, вырабатываемой возобновляемыми источниками, вопросов интеграции энергетических систем, водородной энергетики, задачи энергосбережения и участия потребителей в выработке энергетической политики. Однако тематика конкретных экономических аспектов энергетической политики и энергетической безопасности отражена слабо как в списках ключевых слов, так и в текстах, сгенерированных с использованием технологии ChartGPT.

Полученные в результате анализа библиометрических данных восемь тематических кластеров хорошо описывают различные аспекты доминирующих проблем энергетической политики и энергетической безопасности, присутствующих в научных исследованиях, опубликованных издательствами Elsevier и MDPI в 2021–2023 гг.

Продемонстрирована возможность применения методов извлечения ключевых слов и ключевых фраз, наиболее близко описывающих документ с использованием векторного представления слов KeyBERT для использования их как альтернативы авторским ключевым словам.

Показано, что такие ключевые слова обладают некоторыми статистическими характеристиками, схожими с аналогичными характеристиками авторских ключевых слов, но список сгенерированных ключевых слов лишь частично пересекается с авторскими ключевыми словами. Использование сгенерированных ключевых слов может быть полезно при проведении библиометрических исследований, основанных на данных, не включающих в себя авторские ключевые слова, но содержащих тексты заголовков и аннотаций.

Показано, что использование технологий типа ChatGPT дает достаточно хорошие результаты, основанные на доминирующих текстовых шаблонах, по которым производилось обучение моделей данных.

Следует отметить и негативную сторону использования данных технологий – это преобладание шаблонных фраз и

предложений в сгенерированных текстах. Векторизация слов и текстов на основе их контекста приводит к доминированию слов и фраз самого контекста. В данных технологиях не просматривается алгоритмов снижения весовой значимости часто употребляемых текстовых шаблонов, что приводит к их доминированию. Для критического анализа текстов и сбора актуальной информации необходимо наличие операторов, исключающих или снижающих значимость доминирующих текстовых шаблонов. В классических запросах присутствует оператор “NOT” – реализовать нечто подобное без дообучения существующих моделей является проблемой, требующей отдельного исследования.

Целесообразно провести аналогичное исследование с использованием методов извлечения ключевых слов из текстов, не использующих векторное представление слов и текстов, например, Yake!.

*Статья написана в рамках выполнения государственного задания ИППГ РАН (тема «Фундаментальный базис энергоэффективных, ресурсосберегающих и экологически безопасных, инновационных и цифровых технологий поиска, разведки и разработки нефтяных и газовых месторождений, исследование, добыча и освоение традиционных и нетрадиционных запасов и ресурсов нефти и газа; разработка рекомендаций по реализации продукции нефтегазового комплекса в условиях энергоперехода и политики ЕС по декарбонизации энергетики (фундаментальные, поисковые, прикладные, экономические и междисциплинарные исследования)», № 122022800270-0).*

#### Литература

1. *Sokolowski M.M., Heffron R.J.* Defining and conceptualising energy policy failure: The when, where, why, and how // *Energy Policy*. 2022. Vol. 161. P. 112745. <https://doi.org/10.1016/j.enpol.2021.112745>
2. *Gatto A.* The energy futures we want: A research and policy agenda for energy transitions // *Energy Research & Social Science*. 2022. Vol. 89. P. 102639. <https://doi.org/10.1016/j.erss.2022.102639>
3. *Meyer K.E., Fang T., Panibratov A.Yu.* et al. International business under sanctions // *Journal of World Business*. 2023. Vol. 58, No. 2. P. 101426. <https://doi.org/10.1016/j.jwb.2023.101426>
4. *Новак А.* Российский и мировой ТЭК: вызовы и перспективы // *Энергетическая политика*. 2022. № 4. С. 6–15. [https://doi.org/10.46920/2409-5516\\_2022\\_4170\\_7](https://doi.org/10.46920/2409-5516_2022_4170_7)

5. Pascual Espada J., Solís Martínez J., Cid Rico I., Emilio Velasco Sánchez L. Extracting keywords of educational texts using a novel mechanism based on linguistic approaches and evolutive graphs // *Expert Systems with Applications*. 2023. Vol. 213. P. 118842. <https://doi.org/10.1016/j.eswa.2022.118842>
6. Khan O.A., Wasi S., Siddiqui M.S., Karim A. Keyword extraction for medium-sized documents using corpus-based contextual semantic smoothing // *Complexity*. 2022. Vol. 2022. P. 7015764. <https://doi.org/10.1155/2022/7015764>
7. Nomoto T. Keyword extraction: A modern perspective // *SN Computer Science*. 2022. Vol. 4, No. 1. P. 92. <https://doi.org/10.1007/s42979-022-01481-7>
8. La Quatra M., Cagliero L. Transformer-based highlights extraction from scientific papers // *Knowledge-Based Systems*. 2022. Vol. 252. P. 109382. <https://doi.org/10.1016/j.knosys.2022.109382>
9. Kelebercová L., Munk M. Search queries related to COVID-19 based on keyword extraction // *Procedia Computer Science*. 2022. Vol. 207. P. 2618–2627. <https://doi.org/10.1016/j.procs.2022.09.320>
10. Yang K., Lu J., Duan L. et al. Research hotspots and theme trends in post-traumatic growth: A co-word analysis based on keywords // *International Journal of Nursing Sciences*. 2023. Vol. 10, No. 2. P. 268–275. <https://doi.org/10.1016/j.ijnss.2023.03.001>
11. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for language understanding // *Proceedings of the 2019 NAACL Conference: Human Language Technologies*. Stroudsburg, PA: Association for Computational Linguistics, 2019. Vol. 1. P. 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
12. Grootendorst M. KeyBERT: BibTeX // *Zenodo*. 2021. 25 January. <https://doi.org/10.5281/zenodo.4461265>
13. Schopf T., Klimek S., Matthes F. PatternRank: Leveraging pretrained language models and part of speech for unsupervised keyphrase extraction: // *Proceedings of the 14<sup>th</sup> International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. Vailletta: SciTePress, 2022. Vol. 1. P. 243–248. <https://doi.org/10.5220/0011546600003335>
14. Van Eck N.J., Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping // *Scientometrics*. 2010. Vol. 84, No. 2. P. 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
15. Aljedaani W., Rustam F., Mkaouer M.W. et al. Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry // *Knowledge-Based Systems*. 2022. Vol. 255. P. 109780. <https://doi.org/10.1016/j.knosys.2022.109780>
16. Campos R., Mangaravite V., Pasquali A. et al. YAKE! Keyword extraction from single documents using multiple local features // *Information Sciences*. 2020. Vol. 509. P. 257–289. <https://doi.org/10.1016/j.ins.2019.09.013>
17. VandenBerg C., Callan J. Sifaka: Text mining above a search API // *arXiv*. 2018. 5 October. <https://doi.org/10.48550/ARXIV.1810.02907>
18. Škrlj B., Koloski B., Pollak S. Retrieval-efficiency trade-off of unsupervised keyword extraction // *Lecture Notes in Computer Science*. 2022. Vol. 13601. P. 379–393. [https://doi.org/10.1007/978-3-031-18840-4\\_27](https://doi.org/10.1007/978-3-031-18840-4_27)
19. Kim S.Y., Ganesan K., Dickens P., Panda S. Public sentiment toward solar energy – Opinion mining of Twitter using a Transformer-Based Language Model // *Sustainability*. 2021. Vol. 13, No. 5. P. 2673. <https://doi.org/10.3390/su13052673>

## ПРИЛОЖЕНИЕ

## Приложение 1

**Источник: ScienceDirect**

Запрос: «Title, abstract, keywords: “energy policy”, 2021–2023»

Наиболее часто встречаемые предметные области (**Subject areas**) и число публикаций отнесенных к ним: Energy (712); Environmental Science (218); Engineering (197); Social Sciences (170); Economics, Econometrics and Finance (94); Business, Management and Accounting (24); Decision Sciences (23); Agricultural and Biological Sciences (18); Earth and Planetary Sciences (16); Materials Science (16).

Названия изданий (**Publication title**) с наибольшим числом публикаций: Energy Policy (210); Energy (98); iScience (98); Renewable and Sustainable Energy Reviews (88); Energy Research & Social Science (88); Journal of Cleaner Production (68); Renewable Energy (59); Energy Economics (50); Applied Energy (45); Energy Strategy Reviews (31); Energy Reports (31); Energy and Buildings (28); Sustainable Energy Technologies and Assessments (22).

Запрос: «Title, abstract, keywords: “energy security”, 2021–2023»

Наиболее часто встречаемые предметные области (**Subject areas**) и число публикаций отнесенных к ним: Energy (409); Environmental Science (152); Engineering (129); Social Sciences (84); Chemical Engineering (35); Materials Science (34); Agricultural and Biological Sciences (28); Economics, Econometrics and Finance (26); Chemistry (19); Earth and Planetary Sciences (15).

Названия изданий (**Publication title**) с наибольшим числом публикаций: Energy (57); Journal of Cleaner Production (46); Renewable Energy (43); Renewable and Sustainable Energy Reviews (42); Energy Policy (40); Energy Research & Social Science (30); Energy Reports (27); Energy Strategy Reviews (21); Fuel (20); Energy Economics (17); Resources Policy (16).

**Источник: MDPI**

Запрос: «Title, abstract, keywords: “energy policy”, 2021–2023»

Наиболее часто встречаемые предметные области (**Subjects**) и число публикаций отнесенных к ним: Environmental & Earth Sciences (559); Engineering (370); Chemistry & Materials Science (357); Physical Sciences (350); Business & Economics (188); Social Sciences, Arts and Humanities (181); Computer Science & Mathematics (39); Public Health & Healthcare (25); Biology & Life Sciences (21).

Названия изданий (**Journals**) с наибольшим числом публикаций: Energies (322); Sustainability (173); IJERPH (13); Applied Sciences (8); Water (6).

Запрос: «Title, abstract, keywords: “energy security”, 2021–2023»

Наиболее часто встречаемые предметные области (**Subjects**) и число публикаций отнесенных к ним: Environmental & Earth Sciences (380); Engineering (295); Chemistry & Materials Science (283); Physical Sciences (240); Business & Economics (107); Social Sciences, Arts and Humanities (97); Computer Science & Mathematics (44); Biology & Life Sciences (43); Public Health & Healthcare (23).

Названия изданий (**Journals**) с наибольшим числом публикаций: Energies (206); Sustainability (95); Applied Sciences (15); Water (13); Sensors (11); Molecules (8).





**Приложение 3****Кластер 1**

energy security (118.0), china energy (35.0), energy market (23.0), greenhouse gas emissions (22.0), nuclear energy (21.0), energy planning (19.0), energy system (17.0), energy (13.0), energy management (12.0), multi criteria decision (12.0), coal consumption (11.0), global energy (11.0), international energy (11.0), renewable energy transition (11.0), energy industry (9.0), energy performance (9.0), nuclear power (9.0), poland (9.0), sustainable energy transition (9.0), world energy (9.0), energy sources (8.0), energy trade (8.0), energy transition (8.0), electricity production (7.0), energy buildings (7.0), energy equity (7.0), energy security risk (7.0), consumption (6.0), energy security performance (6.0), gas emissions (6.0), renewable energy capacity (6.0), renewable sources (6.0), sustainable energy security (6.0), criteria decision (5.0), criteria decision analysis (5.0), energy independence (5.0), energy security index (5.0), energy supply (5.0), global energy security (5.0), global energy trade (5.0), nuclear power plants (5.0), zero energy buildings (5.0)

**Кластер 2**

energy poverty (48.0), electric vehicles (27.0), household energy (24.0), eu energy (23.0), energy efficiency policy (18.0), electric vehicle (14.0), energy use (13.0), coal (12.0), sustainable energy policy (12.0), energy community (11.0), green energy (11.0), eu energy policy (10.0), ghg emissions (10.0), energy justice (9.0), energy policymaking (9.0), energy innovation (8.0), electrification (7.0), energy initiatives (7.0), household energy use (7.0), residential buildings (7.0), rural energy (7.0), community energy (6.0), energy consumption patterns (6.0), energy democracy (6.0), energy infrastructure (6.0), energy poverty index (6.0), eu climate (6.0), ev charging (6.0), poverty (6.0), recycling (6.0), renewable generation (6.0), domestic energy (5.0), efficiency policy (5.0), emissions reduction (5.0), energy retrofit (5.0), fuel poverty (5.0), multidimensional energy poverty (5.0), multidimensional energy poverty index (5.0), rural electrification (5.0), smart energy (5.0)

**Кластер 3**

renewable energy (412.0), renewable energy policy (79.0), energy sector (61.0), renewable energy sources (49.0), renewable energy consumption (35.0), electricity market (34.0), wind energy (31.0), renewable electricity (30.0), renewable energy development (26.0), renewable energy resources (17.0), renewable energy technologies (17.0), renewable energy sector (14.0), electricity price (13.0), renewable energy generation (13.0), renewable energy projects (12.0), wind power (12.0), offshore wind energy (10.0), regional energy (10.0), renewable energy systems (10.0), wind turbine (10.0), energy projects (9.0), renewable energy production (9.0), wind farms (8.0), offshore wind farms (6.0), offshore wind power (6.0), renewable energy adoption (6.0), hybrid renewable energy (5.0), renewable energy market (5.0), renewable energy use (5.0), sustainable electricity (5.0), wind energy development (5.0), wind power plants (5.0)

**Кластер 4**

sustainable energy (67.0), solar energy (29.0), biodiesel (25.0), energy development (25.0), biofuels (22.0), photovoltaic (20.0), biomass (19.0), energy price (16.0), solar photovoltaic (16.0), biofuel (14.0), biodiesel production (13.0), sustainable energy development (13.0), energy crisis (11.0),

hydropower (11.0), solar pv (11.0), biomass energy (9.0), hydropower plants (9.0), biofuel production (8.0), biomass resources (7.0), solar power (7.0), alternative fuels (6.0), bioenergy production (6.0), food security (6.0), sustainable development goals (6.0), wastewater (6.0), agrivoltaics (5.0), biomass gasification (5.0), global energy crisis (5.0), photovoltaic power (5.0), sustainable energy consumption (5.0), water energy (5.0), water security (5.0)

#### **Кластер 5**

carbon emissions (48.0), co2 emissions (46.0), emissions (39.0), renewables (27.0), energy production (23.0), sustainable development (20.0), economic growth (15.0), sustainability (14.0), carbon neutrality (13.0), energy resources (13.0), carbon dioxide emissions (12.0), ecological footprint (10.0), energy investments (10.0), energy productivity (9.0), autoregressive (8.0), co2 (8.0), carbon capture (7.0), carbon tax (7.0), climate change (6.0), cointegration (6.0), distributed energy resources (6.0), environmental sustainability (6.0), national energy (6.0), photovoltaics (6.0), sustainability transitions (6.0), carbon pricing (5.0), energy subsidies (5.0), renewable energy investments (5.0), stochastic frontier analysis (5.0), water consumption (5.0), water resources (5.0)

#### **Кластер 6**

energy consumption (87.0), energy efficiency (85.0), energy demand (55.0), electricity consumption (33.0), energy saving (32.0), biogas (17.0), electricity demand (17.0), biogas production (12.0), energy generation (12.0), forecasting (12.0), electricity generation (11.0), electricity sector (10.0), energy efficiency measures (10.0), building energy consumption (9.0), residential energy (9.0), household energy consumption (8.0), biogas plants (7.0), residential electricity (7.0), electricity consumers (6.0), energy usage (6.0), natural gas consumption (6.0), residential energy consumption (6.0), clean energy (5.0), forecast (5.0), households (5.0), natural gas demand (5.0), residential electricity consumption (5.0)

#### **Кластер 7**

energy policy (323.0), national energy policy (16.0), energy economy (14.0), energy sustainability (12.0), european energy (12.0), future energy policy (12.0), polish energy (12.0), climate policy (8.0), energy governance (8.0), energy policy making (8.0), coal mining (7.0), coal production (7.0), energy consumers (7.0), future energy (7.0), district heating (6.0), geothermal (6.0), geothermal energy (6.0), new energy policy (6.0), stakeholders (6.0), energy policy measures (5.0), energy policy scenarios (5.0), mining industry (5.0), polish economy (5.0)

#### **Кластер 8**

energy storage (47.0), hydrogen production (22.0), energy storage system (19.0), energy transition policy (18.0), microgrid (16.0), hydrogen economy (13.0), energy technologies (10.0), renewable power (8.0), battery energy storage (7.0), hydrogen (7.0), variable renewable energy (7.0), energy storage technologies (6.0), power generation (6.0), renewable energy penetration (6.0), renewable hydrogen (6.0), sustainable energy systems (6.0), thermal energy storage (6.0), green hydrogen production (5.0), renewable resources (5.0), smart grid (5.0)

# Identification of actual energy policy and security research topics in Elsevier and MDPI publications for 2021–2023.

## Part 1. Term clustering and topic referencing with ChatGPT

**B.N. Chigarev**

Oil and Gas Research Institute, Russian Academy of Sciences, Moscow, Russia  
E-mail: [bchigarev@ipng.ru](mailto:bchigarev@ipng.ru)

**Abstract.** This article aims to identify relevant research topics in Elsevier and MDPI publications for the years 2021–2023 on energy policy and security issues. The research was based on the identification of keywords/phrases from title texts and abstracts using PatternRank software based on KeyBERT technology and name group extraction. These keywords/phrases were considered as a possible alternative to the author keywords. The bibliometric analysis of these terms using VOSviewer identified eight clusters describing energy policy and security issues in detail. Textual descriptions of the topics in each cluster are provided using ChartGPT technology.

**Keywords:** energy policy, research topics, bibliometric analysis, KeyBERT, ChartGPT, VOSviewer.

**Citation:** *Chigarev B.N.* Identification of actual energy policy and security research topics in Elsevier and MDPI publications for 2021–2023. Part 1. Term clustering and topic referencing with ChatGPT // Actual Problems of Oil and Gas. 2023. Iss. 2(41). P. 63–83. <https://doi.org/10.29222/ipng.2078-5712.2023-41.art5> (In Russ.).

### References

1. *Sokolowski M.M., Heffron R.J.* Defining and conceptualising energy policy failure: The when, where, why, and how // Energy Policy. 2022. Vol. 161. P. 112745. <https://doi.org/10.1016/j.enpol.2021.112745>
2. *Gatto A.* The energy futures we want: A research and policy agenda for energy transitions // Energy Research & Social Science. 2022. Vol. 89. P. 102639. <https://doi.org/10.1016/j.erss.2022.102639>
3. *Meyer K.E., Fang T., Panibratov A.Yu.* et al. International business under sanctions // Journal of World Business. 2023. Vol. 58, No. 2. P. 101426. <https://doi.org/10.1016/j.jwb.2023.101426>
4. *Novak A.* Russian and global fuel and energy complex: challenges and prospects // Energeticheskaya Politika. 2022. No. 4. P. 6–15. [https://doi.org/10.46920/2409-5516\\_2022\\_4170\\_7](https://doi.org/10.46920/2409-5516_2022_4170_7) (In Russ.).
5. *Pascual Espada J., Solís Martínez J., Cid Rico I., Emilio Velasco Sánchez L.* Extracting keywords of educational texts using a novel mechanism based on linguistic approaches and evolutive graphs // Expert Systems with Applications. 2023. Vol. 213. P. 118842. <https://doi.org/10.1016/j.eswa.2022.118842>
6. *Khan O.A., Wasi S., Siddiqui M.S., Karim A.* Keyword extraction for medium-sized documents using corpus-based contextual semantic smoothing // Complexity. 2022. Vol. 2022. P. 7015764. <https://doi.org/10.1155/2022/7015764>
7. *Nomoto T.* Keyword extraction: A modern perspective // SN Computer Science. 2022. Vol. 4, No. 1. P. 92. <https://doi.org/10.1007/s42979-022-01481-7>

8. *La Quatra M., Cagliero L.* Transformer-based highlights extraction from scientific papers // Knowledge-Based Systems. 2022. Vol. 252. P. 109382. <https://doi.org/10.1016/j.knosys.2022.109382>
9. *Kelebercová L., Munk M.* Search queries related to COVID-19 based on keyword extraction // Procedia Computer Science. 2022. Vol. 207. P. 2618–2627. <https://doi.org/10.1016/j.procs.2022.09.320>
10. *Yang K., Lu J., Duan L.* et al. Research hotspots and theme trends in post-traumatic growth: A co-word analysis based on keywords // International Journal of Nursing Sciences. 2023. Vol. 10, No. 2. P. 268–275. <https://doi.org/10.1016/j.ijnss.2023.03.001>
11. *Devlin J., Chang M.-W., Lee K., Toutanova K.* BERT: Pre-training of Deep Bidirectional Transformers for language understanding // Proceedings of the 2019 NAACL Conference: Human Language Technologies. Stroudsburg, PA: Association for Computational Linguistics, 2019. Vol. 1. P. 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
12. *Grootendorst M.* KeyBERT: BibTeX // Zenodo. 2021. 25 January. <https://doi.org/10.5281/zenodo.4461265>
13. *Schopf T., Klimek S., Matthes F.* PatternRank: Leveraging pretrained language models and part of speech for unsupervised keyphrase extraction: // Proceedings of the 14<sup>th</sup> International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. Valletta: SciTePress, 2022. Vol. 1. P. 243–248. <https://doi.org/10.5220/0011546600003335>
14. *Van Eck N.J., Waltman L.* Software survey: VOSviewer, a computer program for bibliometric mapping // Scientometrics. 2010. Vol. 84, No. 2. P. 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
15. *Aljedaani W., Rustam F., Mkaouer M.W.* et al. Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry // Knowledge-Based Systems. 2022. Vol. 255. P. 109780. <https://doi.org/10.1016/j.knosys.2022.109780>
16. *Campos R., Mangaravite V., Pasquali A.* et al. YAKE! Keyword extraction from single documents using multiple local features // Information Sciences. 2020. Vol. 509. P. 257–289. <https://doi.org/10.1016/j.ins.2019.09.013>
17. *VandenBerg C., Callan J.* Sifaka: Text mining above a search API // arXiv. 2018. 5 October. <https://doi.org/10.48550/ARXIV.1810.02907>
18. *Škrlj B., Koloski B., Pollak S.* Retrieval-efficiency trade-off of unsupervised keyword extraction // Lecture Notes in Computer Science. 2022. Vol. 13601. P. 379–393. [https://doi.org/10.1007/978-3-031-18840-4\\_27](https://doi.org/10.1007/978-3-031-18840-4_27)
19. *Kim S.Y., Ganesan K., Dickens P., Panda S.* Public sentiment toward solar energy – Opinion mining of Twitter using a Transformer-Based Language Model // Sustainability. 2021. Vol. 13, No. 5. P. 2673. <https://doi.org/10.3390/su13052673>