

**ЭНЕРГЕТИЧЕСКИЕ РЕСУРСЫ ЗЕМНОЙ КОРЫ: ВЫЗОВЫ СОВРЕМЕННОЙ ЭКОНОМИКИ**

Оригинальная статья

УДК [303.6+303.7]:001.8

<https://doi.org/10.29222/ipng.2078-5712.2024-15-2.art5>**Выявление актуальных задач энергетического перехода в публикациях агрегатора контента научных публикаций Scilit. Часть 2. Кластеризация публикаций****Б.Н. Чигарев** ✉

Институт проблем нефти и газа РАН, Москва, Россия

**Аннотация.** *Актуальность.* Доступ российских исследователей к Scopus и Web of Science стал ограничен, поэтому актуальным становится использование открытых реферативных баз данных. *Цель работы.* Выявление актуальных задач энергетического перехода в публикациях, представленных в агрегаторе контента для научных публикаций с бесплатным доступом Scilit. *Материалы и методы.* В исследовании использована 10121 библиометрическая запись статей за 2019–2023 гг. Публикации систематизировались с использованием алгоритма Gibbs sampling for Dirichlet mixture model. Темы публикаций внутри полученных кластеров анализировались с помощью демоверсии программы Carrot2. Ранжирование публикаций осуществлялось с помощью утилиты sumy с алгоритмом lex-rank. *Результаты.* Выявленные актуальные темы посвящены системным проблемам энергетических комплексов, включая интеграцию различных источников генерации энергии, накопление энергии в «аккумуляторах» или «зеленом водороде» и оптимизацию их работы. Большое внимание уделено социальным аспектам энергетического перехода, особенно актуальным для сельских территорий и регионов с низким уровнем экономического развития. *Выводы.* Без финансовой поддержки и наличия соответствующей инфраструктуры для местных энергетических сообществ энергетический переход может быть ими отвергнут. Домохозяйства следует поощрять к использованию более чистых источников энергии, менее вредных для здоровья и окружающей среды.

**Ключевые слова:** энергетический переход, актуальные темы исследований, агрегатор контента Scilit, библиометрический анализ

**Финансирование:** работа выполнена в рамках государственного задания ИПНГ РАН (тема № 122022800270-0).

**Для цитирования:** Чигарев Б.Н. Выявление актуальных задач энергетического перехода в публикациях агрегатора контента научных публикаций Scilit. Часть 2. Кластеризация публикаций // Актуальные проблемы нефти и газа. 2024. Т. 15, № 2. С. 174–199. <https://doi.org/10.29222/ipng.2078-5712.2024-15-2.art5>

---

✉ Чигарев Борис Николаевич, e-mail: [bchigarev@ipng.ru](mailto:bchigarev@ipng.ru)

© Чигарев Б.Н., 2024



Контент доступен под лицензией Creative Commons Attribution 4.0 License.

## Введение

Прекращение доступа российских исследователей к реферативным базам данных Scopus и Web of Science затрудняет определение актуальных направлений исследований, но не исключает возможности их проведения, способствуя переходу на реферативные базы данных открытого доступа.

Как правило, Scopus и Web of Science являются более полными в плане заполнения полей библиометрических записей. Открытые базы данных способны охватывать большой перечень публикаций, при этом для конкретного исследования могут потребоваться не все поля библиометрических данных.

Автору не удалось найти на платформе Scilit исследований, посвященных энергетическому переходу, в которых бы использовались методы кластеризации текстов и алгоритмы обобщения. В то же время выбранные методы, как показано ниже в обзоре литературы, адекватны поставленной задаче.

Цель исследования – изучить потенциал Scilit для анализа актуальных исследовательских задач энергетического перехода с использованием методов кластеризации коротких текстов GSDMM (Gibbs sampling for Dirichlet mixture model), Lingo3G (из демоверсии Carrot2) и суммаризации текстов алгоритмом lex-rank программы sumy.

## Краткий обзор литературы

В этом разделе представлен краткий обзор публикаций, доступных на платформе Scilit, которые имеют отношение к методам кластеризации и обобщения, используемым в данной работе. Обзор публикаций по теме энергетического перехода в целом приведен в первой части

исследования «Часть 1. Кластеризация ключевых слов» [1].

Запрос к Scilit – “gsdmm” выдает 29 результатов без ограничений по времени и типу публикаций. Большинство из них посвящено кластеризации текстов, что указывает на адекватность выбора автора.

Рассмотрим статьи, наиболее полно раскрывающие особенности использования данного метода при кластеризации текстов.

В работе [2] приведено сравнение применения двух алгоритмов латентного размещения Дирихле (LDA) и модели полиномиальной смеси Дирихле (GSDMM) на коротких текстах. Важность данного сравнения: 1) заголовки публикаций и аннотации являются короткими текстами; 2) алгоритм латентного размещения Дирихле – наиболее часто применяемый метод для тематического анализа текстов. Показано, что эффективность GSDMM на коротких текстах превосходит LDA в плане согласованности и устойчивости тем, при условии, что документ принадлежит к одной теме.

Авторы работы [3] рассмотрели проблему поиска оптимальных значений гиперпараметров алгоритма GSDMM. Оценки гиперпараметров, полученные их методом, не совпадают с обычно используемыми в литературе значениями и приводят к иной кластеризации текстов. Авторы также отмечают, что оптимизированный GSDMM дает более однородный набор слов в каждой теме по сравнению с K-Means. Проблема выбора гиперпараметров является классической, но кластеризация коротких текстов зависит не только от них, но и от предобработки текстов, выбора словаря и даже от вида текстов: аннотации статей или тексты твитов (не опубликованные автором эксперименты использования GSDMM применительно к аннотациям).

В статье [4] оценивалась эффективность кластеризации веб-сервисов с использованием методов тематического моделирования и различных алгоритмов кластеризации. В качестве метода уменьшения размерности и представления признаков предлагается алгоритм выборки Гиббса для модели полиномиальной смеси Дирихле (GSDMM). Результаты показывают, что GSDMM с K-Means или агломеративной кластеризацией превосходит другие методы. Сокращение размерности достигает 90,88%, 88,84% и 93,13% на трех наборах данных, полученных в режиме реального времени.

GSDMM не только позволяет тестировать использование различных гиперпараметров, но и формировать словарь, по которому и производится сравнение текстов, что дает большие возможности для анализа полученных результатов.

Приведенный выше очень краткий литературный обзор, наряду с собственными экспериментами по применению методов кластеризации коротких текстов, послужил обоснованием выбора GSDMM.

В данной работе детализация тем публикаций каждого кластера, полученная с помощью GSDMM, была проведена с использованием демонстрационной версии программы Carrot2 и алгоритма Lingo3G.

В системе Scilit использование демонстрационной версии программы Carrot2 и алгоритма Lingo3G представлено в препринте [5], посвященном изучению социальных аспектов в области энергетических исследований с помощью методов библиометрии. Исследование основано на данных публикаций журнала

Energy Research & Social Science, экспортированных из базы данных ScienceDirect за 2019–2023 гг.

Публикация авторов, разработчиков программы Carrot2, объясняющая работу алгоритма Lingo [6], вышла в 2004 г., но научных публикаций об его использовании немного. При этом на самом сайте программы (<https://carrotsearch.com/lingo3g/>) можно найти детальную информацию о применении алгоритма Lingo3G.

В работах индонезийских авторов [7, 8] для кластеризации документов использовался алгоритм Lingo программы Carrot2 Workbench, которая доступна в открытом доступе на GitHub. Эксперименты автора данной статьи показали, что субъективно алгоритм Lingo3G дает лучшую кластеризацию, чем Lingo, поэтому в исследованиях использовались демоверсия Carrot2 и алгоритм Lingo3G.

Выбор утилиты командной строки sumy для извлечения резюме из текстов обусловлен тем, что в нем реализован целый ряд алгоритмов: Luhn, Edmundson, Latent Semantic Analysis, LSA, LexRank and TextRank, SumBasic, KL-Sum, а также тем, что утилита имеет хорошую поддержку на GitHub (<https://github.com/misobelica/sumy>).

Для ознакомления с использованием методов экстрактивного суммирования текста и, в частности, утилиты sumy, целесообразно обратиться к работе [9]. В ней после соответствующей предварительной обработки контент был обобщен с помощью программ на языке Python: NLTK, Spacy, Genism и sumy методом экстрактивного суммирования текста, который представляет собой компактное резюме текстовых данных, подготовленное на основе ключевых слов, уже присутствующих в документе.

### Материалы и методы

В данной работе использовалась 10121 библиометрическая запись, экспортированная из Scilit, агрегатора контента для научных публикаций, по запросу: “energy transition” in Common Fields [Title, Abstract, Keyword]. Дополнительная фильтрация включала: 2019–2023 гг., тип публикаций – только статьи на английском языке (актуально на 24.01.2024).

Из Scilit экспортировались файлы “citations” в формате RIS, так как они содержат поле AB (“Abstracts”).

Тексты полей “Titles” («Заголовки») и “Abstracts” («Аннотации») были подвергнуты нормализации с помощью утилиты anyascii<sup>1</sup> (подобие транслитерации).

Осуществлялась дополнительная предобработка текста, в частности, удалялись пояснения в различных скобках, остатки тегов разметки и т. д.

Для использования SQL и других скриптов и программ файлы в формате RIS переводились в TSV.

Перед кластеризацией библиометрических записей проверялась заполненность полей “Abstract” и “DOI”.

Кластеризация записей производилась при следующих параметрах:

```
gsdmm (файл с полями Tiles Abstracts DOI).txt Dict.txt GSDMM_ -a 0.1 -b 0.1 -m 1000 -k 10
```

Содержание полученных 10 кластеров анализировалось с использованием демоверсии приложения Carrot2 и алгоритма кластеризации Lingo3G. Подробнее с программой и условиями ее использования можно ознакомиться на сайте: <https://search.carrotsearch.com/#!/about>.

<sup>1</sup> Утилита преобразования символов Unicode в их наилучшее ASCII-представление. URL: <https://github.com/anyascii/anyascii> (дата обращения: 24.01.2024).

Этим условиям отвечали 9743 записи. Библиометрические записи были сгруппированы с использованием алгоритма GSDMM, разработанного Yin and Wang [10] для кластеризации коротких текстовых документов.

Для GSDMM важно, по каким терминам будут сравниваться записи и по какому словарю.

Для этого из текстов записей были удалены стоп-слова, взятые из программ GATE [11] и SpaCy<sup>2</sup>, а термины были лемматизированы с помощью утилиты sed и списка лемм для английского языка, собранного на GitHub.

После дополнительной предобработки текстов заголовков и аннотаций, составлялся список встречаемости терминов в тексте.

В этом списке были сохранены термины, встречающиеся 100 и более раз, и удалены первые 5 терминов с наибольшей частотой встречаемости. После первой итерации кластеризации записей были определены и удалены из словаря термины, часто встречающиеся во всех кластерах.

Таким образом, был составлен словарь из 2012 терминов, который использовался в алгоритме GSDMM.

Использовались следующие параметры, отличные от параметров по умолчанию:

– “minClusterSize”: 0.1, определяет минимально допустимый размер кластера по отношению к размеру родительского кластера;

<sup>2</sup> Программа обработки естественного языка (NLP) на Python. URL: <https://github.com/explosion/spaCy> (дата обращения: 24.01.2024).

– “maxClusterSize”: 0.3, определяет максимально допустимый размер кластера по отношению к размеру родительского кластера;

– “clusterCountBase”: 3, число кластеров, обнаруженных за каждый проход кластеризации. Чем больше значение этого параметра, тем больше общее число кластеров.

Полученные для каждого кластера (методом GSDMM) результаты отображались в виде графического представления распределения записей по темам и кратких аннотаций двух публикаций, наиболее релевантных выбранному подкластеру (см. итоги выбора под круговыми диаграммами).

Релевантные публикации выбирались с помощью утилиты sumy с алгоритмом lex-rank<sup>1</sup>.

При кратком изложении содержания публикации по возможности сокращали объем реферата в два-три раза по сравнению с аннотацией.

Использование sumy позволяло определить релевантность статьи в контексте всего списка анализируемых работ, относящихся к данному кластеру.

Краткие содержания статей составлялись таким образом, чтобы их текст был примерно в два-три раза меньше, чем аннотация.

## Результаты и обсуждения

### **Кластеры, полученные в результате применения алгоритма GSDMM к текстам заголовков и аннотаций 9743 библиометрических записей**

В табл. 1 и 2 представлены наиболее часто встречаемые термины для каждого из 10 кластеров, полученные путем применения алгоритма GSDMM к текстам заголовков и аннотаций 9743 библиометрических записей. TF – частота встречаемости термина в кластере.

**Табл. 1.** Распределение 30 наиболее часто встречаемых терминов по кластерам 0–4

**Table 1.** Distribution of the 30 most frequently occurring terms by the Clusters 0–4

Кластер 0		Кластер 1		Кластер 2		Кластер 3		Кластер 4	
Термин	TF	Термин	TF	Термин	TF	Термин	TF	Термин	TF
1	2	3	4	5	6	7	8	9	10
policy	586	policy	593	household	164	electricity	594	climate	612
economic	519	research	515	clean	146	power	551	global	590
effect	508	social	494	fuel	128	scenario	533	policy	581
country	458	change	466	rural	121	technology	492	country	556
environmental	430	understand	423	health	102	gas	479	change	543
data	375	article	402	policy	98	sector	477	economic	470
growth	356	local	396	cooking	96	demand	475	gas	448
sustainable	353	approach	393	air	88	fuel	460	article	417
consumption	347	climate	379	effect	85	cost	458	fuel	400
change	333	framework	374	data	84	climate	452	challenge	396
global	317	community	366	china	82	production	437	fossil	375
research	317	actor	344	improve	77	policy	435	sector	373

<sup>1</sup> Модуль автоматического резюмирования текстовых документов. URL: <https://github.com/miso-belica/sumy> (дата обращения: 24.01.2024).

Продолжение табл. 1

Table 1 continued

1	2	3	4	5	6	7	8	9	10
significant	303	political	337	income	75	global	429	world	358
economy	298	explore	336	pollution	70	generation	412	sustainable	348
green	291	challenge	326	gas	66	economic	388	economy	347
technology	288	power	319	survey	61	reduction	382	state	332
promote	285	sustainable	313	area	60	change	374	market	324
china	268	identify	308	solid	58	supply	363	green	320
factor	267	literature	303	consumption	56	fossil	359	industry	316
climate	266	address	297	significant	56	term	329	oil	308
level	266	perspective	281	biomass	55	environmental	322	power	306
provide	261	technology	270	coal	52	sustainable	321	international	303
evidence	259	support	267	traditional	52	compare	308	technology	302
panel	250	argue	266	heat	51	country	293	environmental	291
empirical	242	global	263	higher	51	resource	287	term	282
resource	241	context	261	factor	50	assessment	278	research	273
examine	237	examine	258	level	50	greenhouse	278	resource	271
relationship	237	aim	257	significantly	49	capacity	276	government	267
investigate	232	economic	257	promote	48	co2	276	production	264
clean	230	draw	256	quality	48	consumption	275	investment	262

Табл. 2. Распределение 30 наиболее часто встречаемых терминов по кластерам 5–9

Table 2. Distribution of the 30 most frequently occurring terms by the Clusters 5–9

Кластер 5		Кластер 6		Кластер 7		Кластер 8		Кластер 9	
Термин	TF	Термин	TF	Термин	TF	Термин	TF	Термин	TF
1	2	3	4	5	6	7	8	9	10
research	558	state	195	power	599	policy	311	advance	99
sustainable	502	structure	173	propose	419	survey	207	surface	99
technology	434	investigate	166	electricity	376	household	205	density	98
challenge	412	effect	159	generation	352	factor	198	environmental	98
approach	407	spectrum	141	grid	345	support	197	main	98
policy	395	density	139	cost	326	social	190	addition	96
aim	371	optical	138	demand	300	research	186	industrial	96
provide	371	method	134	storage	289	data	184	propose	96
identify	354	experimental	125	approach	280	electricity	169	operate	95
review	352	absorption	121	operation	266	change	164	significant	95
climate	350	reveal	116	method	262	level	164	demand	93
environmental	335	temperature	115	design	259	solar	163	electrochemical	91
change	323	level	114	data	249	influence	155	enable	91
support	310	obtain	114	provide	244	technology	155	highlight	90
sector	303	excitation	113	technology	242	identify	153	1.5	9
literature	300	calculation	112	challenge	238	aim	151	academic	9
framework	296	work	112	capacity	235	effect	151	action	9

Продолжение табл. 2

Table 2 continued

1	2	3	4	5	6	7	8	9	10
economic	291	electronic	106	compare	232	local	149	center	9
social	282	theory	104	solar	232	sustainable	144	central	9
propose	261	time	104	load	228	community	141	collection	9
design	259	electron	103	performance	226	environmental	140	company	9
solution	258	band	102	wind	223	understand	140	consequently	9
strategy	253	compare	101	time	220	provide	139	curve	9
global	252	material	101	economic	218	investigate	136	damage	9
address	251	demonstrate	100	integrate	214	country	132	decarbonisation	9
article	246	ion	100	optimization	206	cost	129	double	9
build	238	parameter	100	simulation	206	explore	129	effectiveness	9
plan	238	transfer	100	plan	205	climate	128	frame	9
understand	235	interaction	99	solution	204	economic	123	installation	9
project	234	indicate	98	scenario	202	government	123	look	9

Особенностью алгоритма GSDMM является то, что записи кластеризуются в соответствии с распределением в них терминов, а сами термины могут встречаться в записях, принадлежащих к разным кластерам.

Более подробно темы полученных кластеров анализируются в следующем разделе.

#### **Тематика кластеров, полученная с использованием демоверсии программы Carrot2 и алгоритма Lingo3G**

Ниже представлены диаграммы распределения публикаций по темам и подтемам для каждого из 10 кластеров, полученных в предыдущем разделе. Под каждым рисунком указана выбранная субтема, по которой были выявлены публикации с высоким рейтингом. Далее следуют примеры публикаций с наивысшим рангом (определены по алгоритму lex-rank утилиты sumu) и краткое описание их содержания.

Диаграмма распределения публикаций для нулевого кластера приведена на рис. 1.

Выбрана подтема “Carbon Emission” («Выбросы углерода») темы “Energy Consumption” («Потребление энергии»), 74 публикации.

#### **Примеры статей, соответствующих тематике кластера**

*Публикация* “Nexus between energy consumption and carbon dioxide emission: evidence from 10 highest fossil fuel and 10 highest renewable energy-using economies” [12] («Взаимосвязь между потреблением энергии и выбросами углекислого газа: данные по 10 странам с наибольшим потреблением ископаемого топлива и 10 странам с наибольшим потреблением возобновляемых источников энергии» – перевод автора).

*Краткое содержание.* В исследовании рассматривается взаимосвязь между экономическим ростом, открытостью торговли, потреблением энергии, валовым накоплением капитала, ростом населения и продолжительностью жизни в 10 ведущих странах, использующих возобновляемые источники энергии (TRU) и ископаемое топливо (TFU), в период с 1991 по 2020 гг.

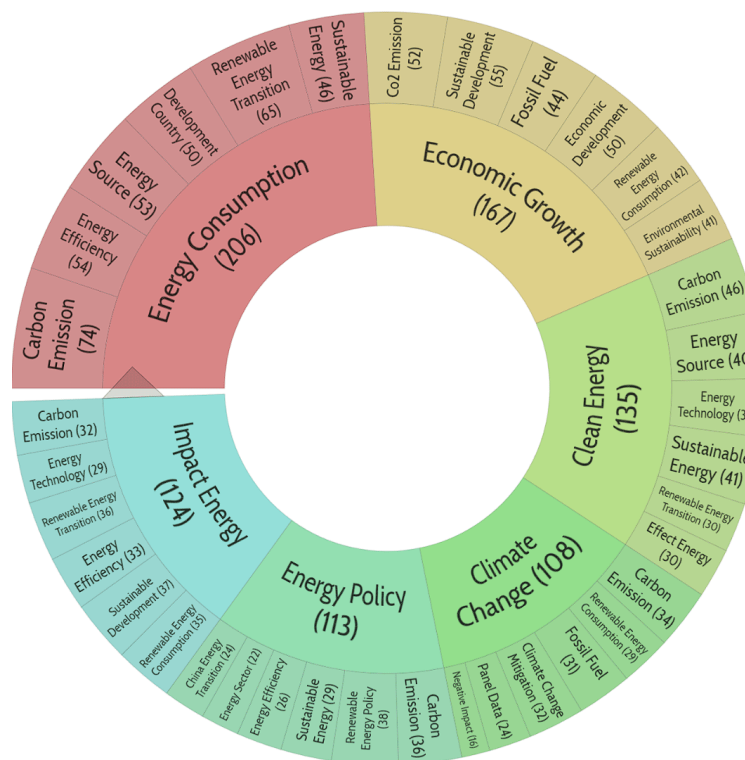


Рис. 1. Распределение публикаций по темам и подтемам для кластера 0

Fig. 1. Distribution of publications by topics and subtopics for the Cluster 0

Результаты показывают, что экономический рост, потребление ископаемого топлива, открытость торговли и валовое накопление капитала увеличивают уровень выбросов углекислого газа (CO<sub>2</sub>) в краткосрочном периоде для стран TRU и TFU, за исключением валового накопления капитала для стран TFU. Однако экономический рост увеличивает выбросы CO<sub>2</sub> только для стран TRU, в то время как потребление ископаемого топлива увеличивает выбросы CO<sub>2</sub> для обеих групп экономик в долгосрочной перспективе. Напротив, возобновляемые источники энергии снижают выбросы CO<sub>2</sub> в краткосрочной и долгосрочной перспективе, а человеческий капитал – только в краткосрочной.

Публикация “Can clean energy adoption and international trade contribute to the achievement of India’s 2070 carbon neutrality agenda? Evidence using quantile ARDL

measures” [13] («Могут ли внедрение чистых источников энергии и международная торговля способствовать реализации программы углеродной нейтральности Индии на 2070 год? Доказательства с использованием квантильных показателей ARDL» – перевод автора).

Краткое содержание. Индия – одна из крупнейших развивающихся экономик мира, которая в основном характеризуется высокой энергоемкостью и зависимостью от ископаемого топлива. В данном исследовании используются как традиционные модели авторегрессии с распределенным лагом (ARDL), так и недавно разработанные квантильные модели ARDL (QARDL). Результаты ARDL показывают, что увеличение потребления возобновляемой энергии на 1% способствует сокращению выбросов CO<sub>2</sub> на 0,8%, в то время как экономический рост увеличивает выбросы CO<sub>2</sub> в долгосрочной перспективе.



Результаты QARDL показывают, что экономический рост положительно, а потребление возобновляемой энергии отрицательно влияют на выбросы CO<sub>2</sub>

во всех квантилях выбросов CO<sub>2</sub>.

На рис. 2 приведена диаграмма распределения публикаций для первого кластера.

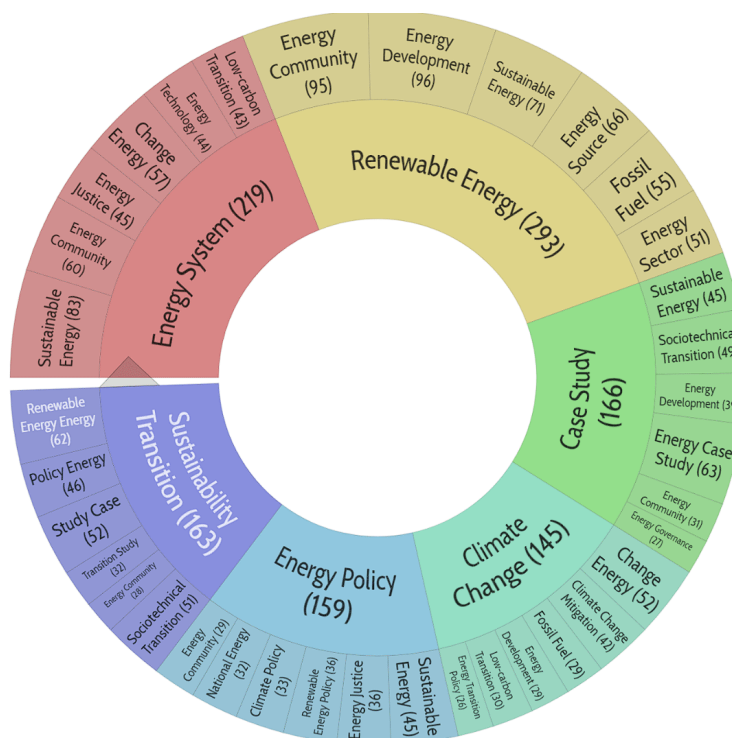


Рис. 2. Распределение публикаций по темам и подтемам для кластера 1

Fig. 2. Distribution of publications by topics and subtopics for the Cluster 1

Выбрана подтема “Sustainable Energy” («Устойчивая энергетика») темы “Energy System” («Энергетическая система»), 83 публикации.

**Примеры статей, соответствующих тематике кластера**

Публикация “Social innovation supports inclusive and accelerated energy transitions with appropriate governance” [14] («Социальные инновации способствуют инклюзивному и ускоренному энергетическому переходу при надлежащем управлении» – перевод автора).

*Краткое содержание.* В данном исследовании изучается роль социальных инноваций в ускорении перехода к устойчивой и справедливой энергетике на основе использования смешанных методов

анализа данных, полученных в результате экспертных интервью, изучения документов, экспериментов, репрезентативных опросов и опросов экспертов. Исследование позволило сделать четыре ключевых вывода: (1) понимание социальных инноваций в энергетике требует признания основных социальных практик и изменений в социальных отношениях; (2) управление, политические структуры и национальный контекст в значительной степени определяют динамику социальных инноваций; (3) на процессы социальных инноваций влияют различные политические факторы; (4) социальные инновации в энергетическом секторе, как правило, имеют сильное общественное признание среди граждан и местных сообществ.

Публикация “Toward a comprehensive framework of social innovation for climate neutrality: A systematic literature review from business/production, public policy, environmental sciences, energy, sustainability and related fields” [15] («В поисках всеобъемлющей системы социальных инноваций для обеспечения климатической нейтральности: Систематический обзор деловой литературы, ...» – перевод автора).

Краткое содержание. Социальные инновации рассматриваются как ключевой фактор, способствующий реализации мер по борьбе с изменением климата и декарбонизации. Помимо технологических инноваций, новые социальные практики

могут привести к трансформации социотехнических систем в сторону более демократичного энергетического перехода и широкого участия граждан в действиях по защите климата. В ходе анализа существующей литературы были определены следующие ключевые аспекты социальных инноваций: контекст, исходные условия, деятельность в области социальных инноваций и ее результаты.

Диаграмма распределения публикаций для второго кластера представлена на рис. 3.

Выбрана подтема “Clean Fuel” («Чистое топливо») темы “Household Energy Transition” («Энергетический переход домохозяйств»), 19 публикаций.

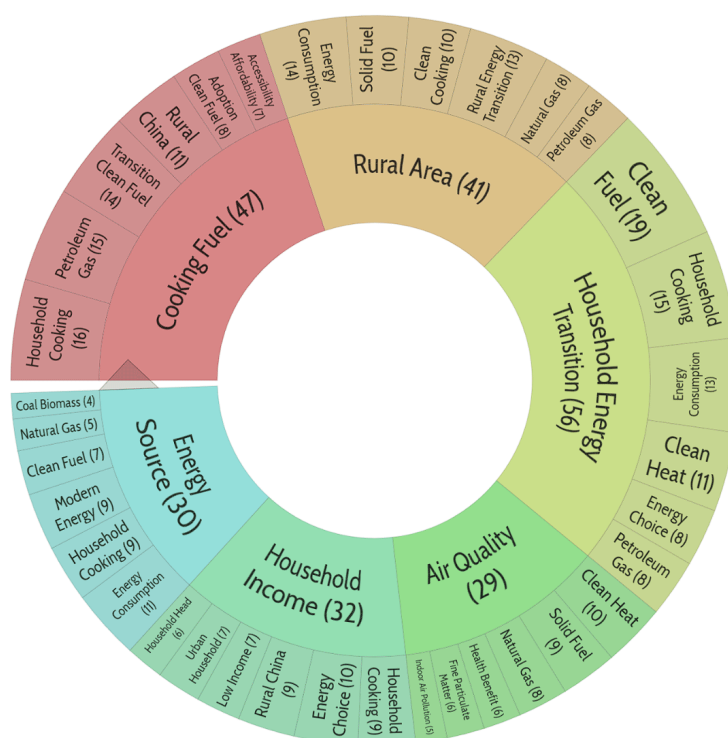


Рис. 3. Распределение публикаций по темам и подтемам для кластера 2

Fig. 3. Distribution of publications by topics and subtopics for the Cluster 2

**Примеры статей, соответствующих тематике кластера**

Публикация “Cooking fuel choices and subjective well-being in rural China: Implications for a complete

energy transition” [16] («Выбор топлива для приготовления пищи и субъективное благополучие в сельских районах Китая: Перспективы полного энергетического перехода» – перевод автора).

*Краткое содержание.* В данной работе рассматривается влияние выбора топлива для приготовления пищи на субъективное благополучие людей, измеряемое категориями счастья и удовлетворенности жизнью, с использованием данных открытого доступа из исследования динамики трудовых ресурсов Китая за 2016 год. Эмпирические результаты показывают, что полный переход на энергоснабжение значительно повышает уровень счастья и удовлетворенности жизнью. Полный энергетический переход повышает субъективное благополучие людей в восточной и центральной частях Китая, но не оказывает никакого влияния на людей в западной части страны. Переход на экологически чистое топливо также значительно улучшает субъективное самочувствие людей.

*Публикация* “A systematic review of household energy transition in low and

middle income countries” [17] («Систематический обзор энергетического перехода домашних хозяйств в странах с низким и средним уровнем дохода» – перевод автора).

*Краткое содержание.* Авторы данного исследования провели систематический обзор, в котором рассмотрели процесс принятия решений в домохозяйствах и факторы, влияющие на использование твердого топлива или отказ от него, а также переход на более чистые виды топлива и усовершенствованные плиты с использованием биомассы. На основе результатов обзора построена концептуальная схема принятия решений о потреблении энергии в быту и факторов, влияющих на них.

На рис. 4 приведена диаграмма распределения публикаций для третьего кластера.

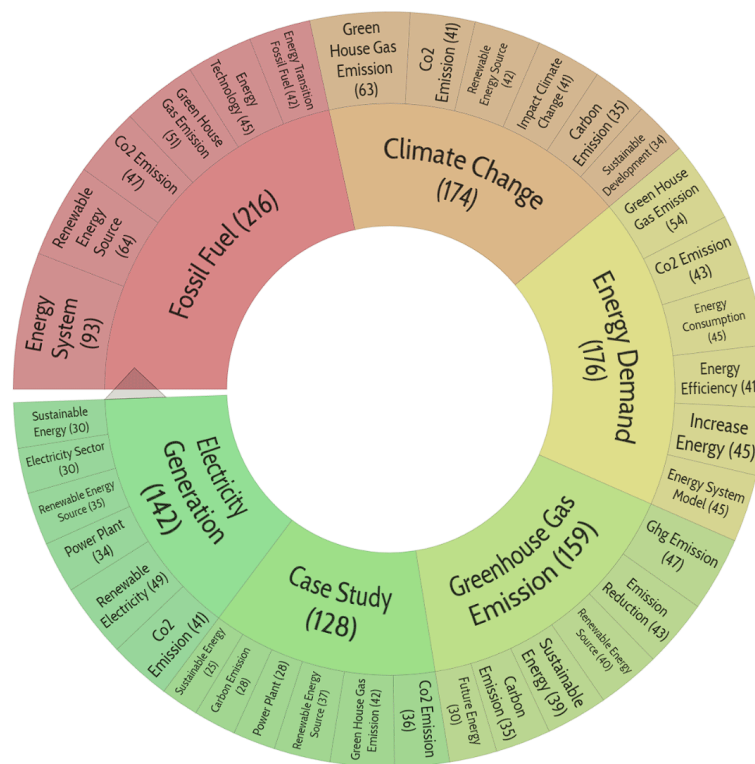


Рис. 4. Распределение публикаций по темам и подтемам для кластера 3

Fig. 4. Distribution of publications by topics and subtopics for the Cluster 3

Выбрана подтема “Energy System” («Энергетическая система») темы “Fossil Fuel” («Ископаемое топливо»), 93 публикации.

**Примеры статей,  
соответствующих тематике  
кластера**

*Публикация* “Sustainable energy transition for renewable and low carbon grid electricity generation and supply” [18] («Переход к устойчивой энергетике для производства и поставки электроэнергии из возобновляемых источников и низкоуглеродных сетей» – перевод автора).

*Краткое содержание.* В 2020 г. на ископаемое топливо приходилось 61,3% мирового производства электроэнергии. Переход к устойчивой энергетической системе требует политических инициатив, направленных на повышение эффективности использования невозобновляемых источников. Устойчивость энергетике лучше всего анализировать с помощью пятимерного подхода, включающего экологическую, экономическую, социальную, техническую и институциональную/политическую устойчивость. Для создания устойчивой электроэнергетической системы необходимы технологии, политика, стратегии, инфраструктура, такая как интеллектуальные сети, и модели с сочетанием возобновляемых и низкоуглеродных источников энергии.

*Публикация* “Pathway towards achieving 100% renewable electricity by 2050 for South Africa” [19] («Путь к достижению 100-процентной доли возобновляемой электроэнергии к 2050 г. для Южной Африки» – перевод автора).

*Краткое содержание.* Южная Африка переходит на экономически эффективную, не содержащую ископаемых углеродов энергетическую систему, стремясь смягчить

воздействие на климат, связанное с взаимосвязью водных и энергетических ресурсов. Ключевым вариантом для страны является отказ от энергосистем, зависящих от угля. Моделирование пяти сценариев показывает, что солнечная фотоэлектрическая и ветровая энергии могут преодолеть угольную зависимость в энергетическом секторе. Выровненная стоимость электроэнергии незначительно увеличивается с 49,2 евро/МВт-ч в 2015 г. до 50,8 евро/МВт-ч в сценарии наилучшей политики и значительно – до 104,9 евро/МВт-ч в сценарии текущей политики к 2050 г. Стоимость электроэнергии без учета затрат на выбросы парниковых газов также несколько возрастает: с 44,1 евро/МВт-ч в 2015 г. до 47,1 евро/МВт-ч в сценарии наилучшей политики и до 62,8 евро/МВт-ч в сценарии текущей политики. В сценарии наилучшей политики без учета затрат на выбросы парниковых газов 96% приходится на возобновляемые источники энергии, а оставшиеся 4% – на уголь и газовые турбины.

Диаграмма распределения публикаций для четвертого кластера представлена на рис. 5.

Выбрана подтема “Oil Gas” («Нефть Газ») темы “Fossil Fuel” («Ископаемое топливо»), 71 публикация.

**Примеры статей,  
соответствующих тематике  
кластера**

*Публикация* “The clean energy claims of BP, Chevron, ExxonMobil and Shell: A mismatch between discourse, actions and investments” [20] («Заявления BP, Chevron, ExxonMobil и Shell о чистой энергии: Несоответствие между дискурсом, действиями и инвестициями» – перевод автора).

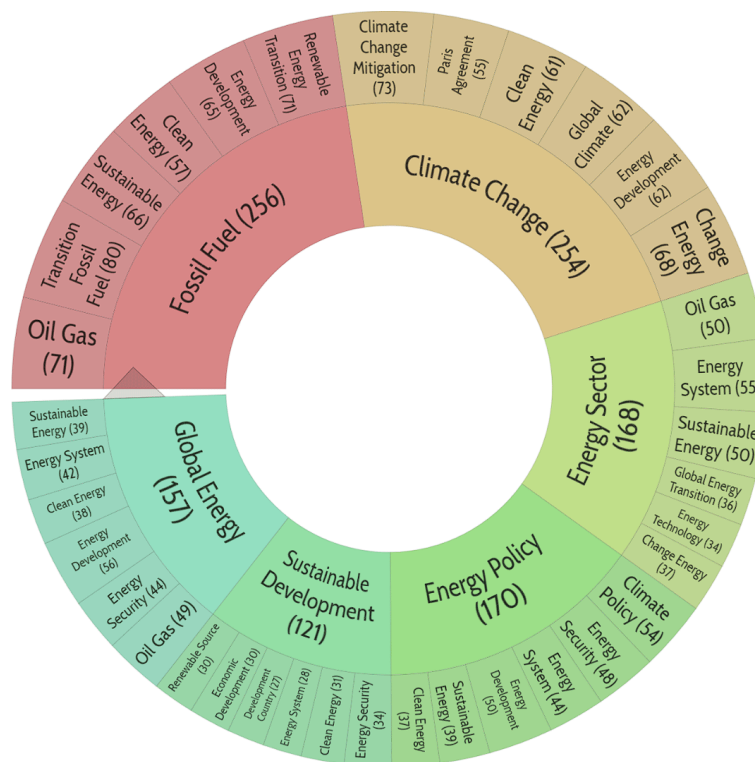


Рис. 5. Распределение публикаций по темам и подтемам для кластера 4

Fig. 5. Distribution of publications by topics and subtopics for the Cluster 4

*Краткое содержание.* Глобальная экономика должна быть декарбонизирована к середине столетия, чтобы избежать изменения климата, а это требует значительной трансформации бизнес-моделей, основанных на использовании ископаемого топлива. Крупнейшие компании обсуждают вопросы чистой энергетики и изменения климата, заявляют о стратегиях декарбонизации и инвестируют в альтернативные источники энергии. Данные за период 2009–2020 гг. свидетельствуют о растущей тенденции к декарбонизации и стратегиям использования чистой энергии. Однако финансовый анализ показывает сохраняющуюся зависимость от ископаемого топлива и незначительные расходы на чистую энергию. Перехода к бизнес-моделям на основе чистой энергии не происходит, поскольку масштабы инвестиций и действий не соответствуют рассуждениям.

*Публикация* “The Russian coal industry in an uncertain world: Finally pivoting to Asia?” [21] («Российская угольная промышленность в нестабильном мире: наконец-то поворот в Азию» – перевод автора).

*Краткое содержание.* В статье рассматривается российская угольная промышленность, крупнейший в мире экспортер ископаемого топлива и третий по величине экспортер угля, с использованием концепции тройной встроенности (Triple Embeddedness Framework – TEF). Оценивается способность отрасли адаптироваться к меняющимся социально-политическим и технико-экономическим условиям. Анализ показывает, что российская угольная промышленность не готовится к долгосрочным изменениям на международных рынках угля, что создает риски для местных сообществ, региональной экономики и глобальной декарбонизации.

Российские экспортеры угля пытаются ускорить свою переориентацию на Восток после введения Европейским союзом санкций в связи с СВО на Украине, а российская угольная инфраструктура

расширяется для обслуживания рынка Азиатско-Тихоокеанского региона.

На рис. 6 приведена диаграмма распределения публикаций для пятого кластера

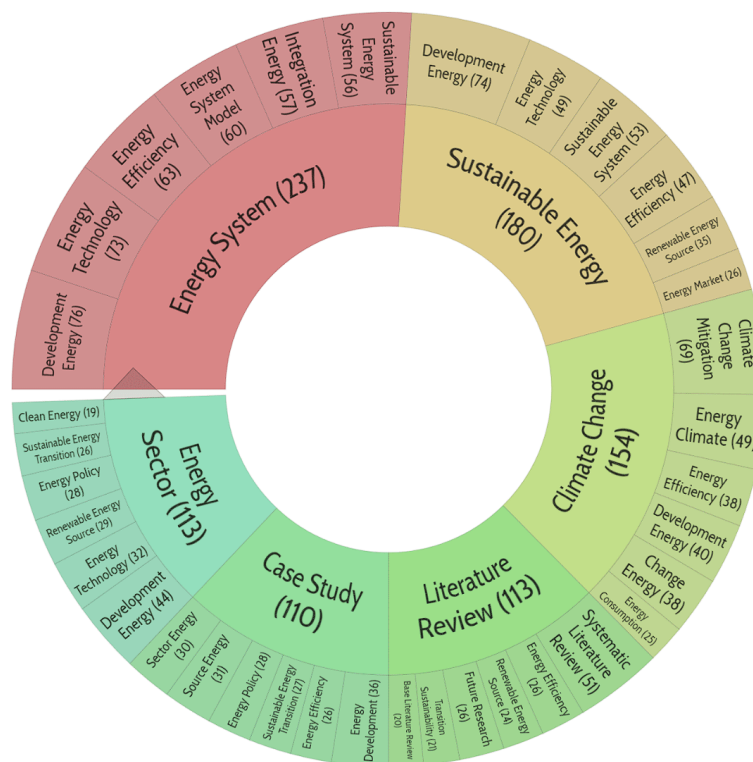


Рис. 6. Распределение публикаций по темам и подтемам для кластера 5

Fig. 6. Distribution of publications by topics and subtopics for the Cluster 5

Выбрана подтема “Development Energy” («Энергетика развития») темы “Energy System” («Энергетическая система»), 76 публикаций.

**Примеры статей, соответствующих тематике кластера**

Публикация “The future of the Polish energy mix in the context of social expectations” [22] («Будущее энергетического баланса Польши в контексте социальных ожиданий» – перевод автора).

Краткое содержание. Авторы провели исследование, чтобы определить, готовы ли местные власти к формированию содержательной энергетической политики

в соответствии с польской энергетической политикой и региональными условиями, учитывая мнение жителей региона. Авторы обнаружили, что не существует комплексных исследований мнений представителей местных органов власти об энергетической политике. Исследование показало, что хотя большинству респондентов известны предположения о польской энергетической политике до 2030 года, почти четверть не осознает, что политика изменится в ближайшем будущем. Большинство респондентов считают, что польское правительство должно подготовить соответствующую информационную кампанию, связанную с энергетической политикой.

Авторы предлагают повысить осведомленность населения, чтобы сформировать энергетическую экономику регионов. Исследование также выявило предпочтения представителей местных органов власти относительно оптимального энергетического баланса в Польше и их субъективную оценку общественного мнения относительно ожидаемой направленности энергетической политики.

*Публикация* “Open data and energy analytics – An analysis of essential information for energy system planning, design and operation” [23] («Открытые данные и энергетическая аналитика – анализ необходимой информации для планирования, проектирования и эксплуатации энергосистем» – перевод автора).

*Краткое содержание.* Энергетический переход трансформирует глобальные и национальные энергетические системы, требуя эффективных стратегий принятия решений. В данной статье анализируются

открытые энергетические модели и данные, выделяется важная информация для междисциплинарных исследований. В ней освещаются ключевые концепции и перспективы моделирования, которые влияют на планирование, проектирование и эксплуатацию энергетических систем. Экосистемы взаимодействующих открытых данных и моделей имеют решающее значение для разработки энергетических услуг и технологий следующего поколения. Инновационные бизнес-модели требуют прозрачного мониторинга, проверки и отслеживания эффективности на разных уровнях.

Диаграммы распределения публикаций для шестого и седьмого кластеров представлены на рис. 7 и 8. Публикации кластера 6 не имеют прямого отношения к процессу энергетического перехода. Они касаются вопросов физических низкоэнергетических переходов и в данной статье не анализируются.

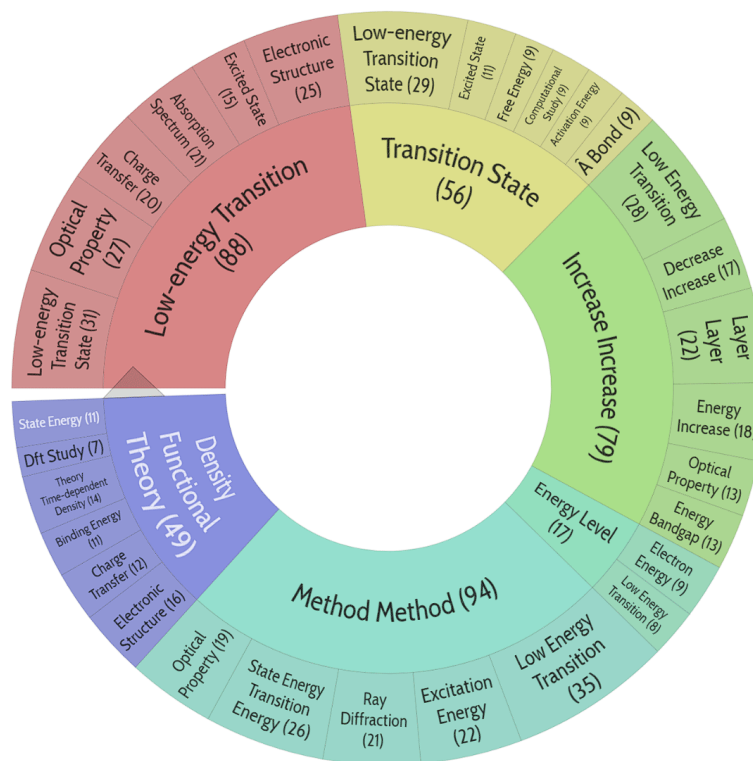


Рис. 7. Распределение публикаций по темам и подтемам для кластера 6

Fig. 7. Distribution of publications by topics and subtopics for the Cluster 6

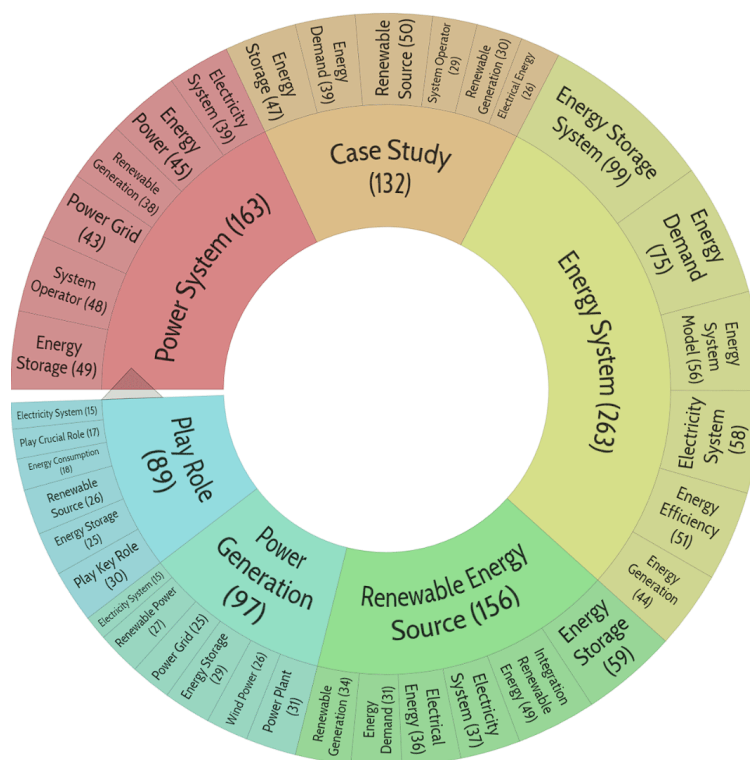


Рис. 8. Распределение публикаций по темам и подтемам для кластера 7

Fig. 8. Distribution of publications by topics and subtopics for the Cluster 7

Выбрана подтема “Energy Storage System” («Система накопления энергии») темы “Energy System” («Энергетическая система»), 99 публикаций.

**Примеры статей, соответствующих тематике кластера**

*Публикация.* “Battery storage configuration of AC/DC hybrid distribution networks” [24] («Конфигурация аккумуляторных батарей в гибридных распределительных сетях переменного/постоянного тока» – перевод автора).

*Краткое содержание.* Нестабильность возобновляемых источников энергии создает проблемы для устойчивого функционирования энергосистемы. Аккумуляторные системы хранения энергии (BESS) в распределительных сетях переменного/постоянного тока могут эффективно решить эти проблемы.

В данной работе предлагается двух-уровневая оптимизационная модель распределения энергии BESS для гибридных сетей AC/DC с целью решения задач технико-экономической оптимизации. Результаты моделирования показывают, что согласованная работа BESS и гибридных сетей AC/DC может эффективно подавить колебания напряжения и улучшить показатель эффективности BESS на протяжении всего жизненного цикла.

*Публикация* “Hybrid energy storage design and dispatch strategy evaluation with sensitivity analysis: Techno-economic-environmental assessment” [25] («Разработка гибридных накопителей энергии и оценка стратегии диспетчеризации с помощью анализа чувствительности: технико-экономическая и экологическая оценка» – перевод автора).



*Краткое содержание.* Целью данного исследования являлась разработка оптимизированной гибридной системы накопления энергии с использованием аккумуляторов и суперконденсаторов для объединения с крупномасштабной солнечной фотоэлектрической системой. В исследовании оцениваются две стратегии диспетчеризации – следование нагрузке и циклическая зарядка – для трех различных конфигураций энергосистемы. В исследо-

вании используются локализованные данные о потреблении энергии в условиях Малайзии. Предложенная система показывает 30-процентное улучшение годовой пропускной способности при диспетчеризации с учетом нагрузки и позволяет увеличить долю возобновляемых источников энергии на 5%.

На рис. 9 приведена диаграмма распределения публикаций для восьмого кластера

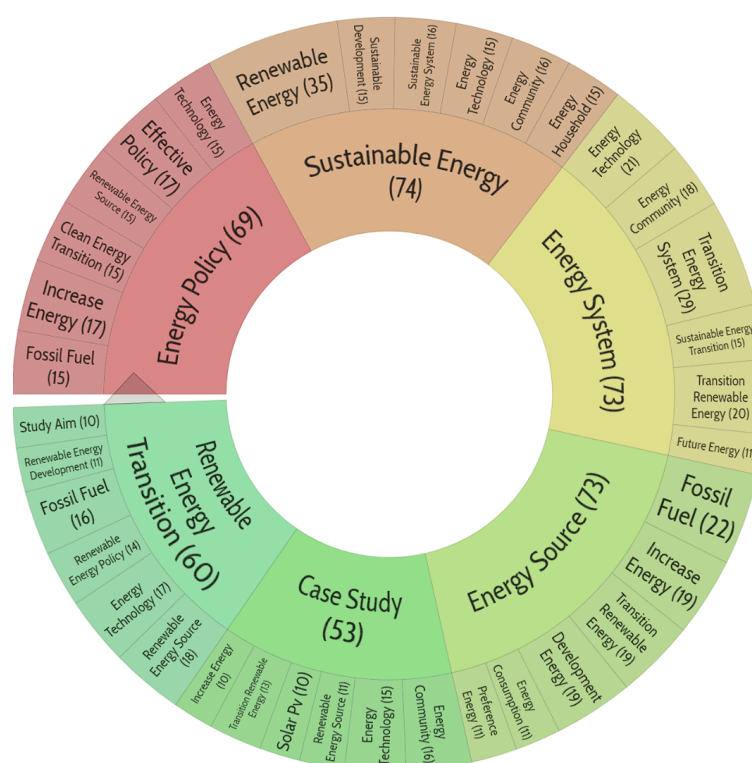


Рис. 9. Распределение публикаций по темам и подтемам для кластера 8

Fig. 9. Distribution of publications by topics and subtopics for the Cluster 8

Выбрана подтема “Renewable Energy” («Возобновляемая энергия») темы “Sustainable Energy” («Устойчивая энергетика»), 35 публикаций.

**Примеры статей, соответствующих тематике кластера**

Публикация “Sustainable energy transition and circular economy: The heterogeneity of potential investors in rural

community renewable energy projects” [26] («Устойчивый энергетический переход и экономика замкнутого цикла: неоднородность потенциальных инвесторов в проекты по возобновляемой энергетике в сельских общинах» – перевод автора).

*Краткое содержание.* В данной статье анализируется готовность инвестирования в возобновляемые источники энергии в небольшой галисийской деревне.

Исследование сосредоточено на различных финансовых, социально-демографических и социально-психологических особенностях населения. Результаты показывают неоднородность индивидуальных установок и опасений, которые влияют на готовность инвестирования в проекты по возобновляемым источникам энергии в сельской местности. Исследование показывает, что организаторы проектов и политики должны учитывать эти факторы при разработке стимулов и стратегий, направленных на развитие возобновляемой энергетики в сельской местности.

*Публикация* “On the future(s) of energy communities in the German energy transition: A derivation of transformation pathways” [27] («О будущем энергетических сообществ в условиях энергетического перехода в Германии: определение путей трансформации» – перевод автора).

*Краткое содержание.* Европейский Союз стремится к активному участию

граждан в устойчивом энергетическом переходе, особенно в энергетических сообществах, для успешного преобразования энергетической системы. В настоящее время энергетические сообщества в Европе невелики. В данной работе была представлена модель для изучения участия граждан в энергетическом переходе в Германии. Данные для модели были получены на основе тематических исследований, интервью и опросов. Текущие тенденции свидетельствуют о том, что переход на основе энергетических сообществ, основанный на участии граждан, может оказаться неуспешным. В статье делается вывод, что для активного участия различных социальных групп в производстве энергии необходимо изменить ряд условий, таких как отмена мешающих нормативных актов и расширение схем финансовой поддержки для низших социально-экономических групп.

Диаграмма распределения публикаций для девятого кластера представлена на рис. 10.

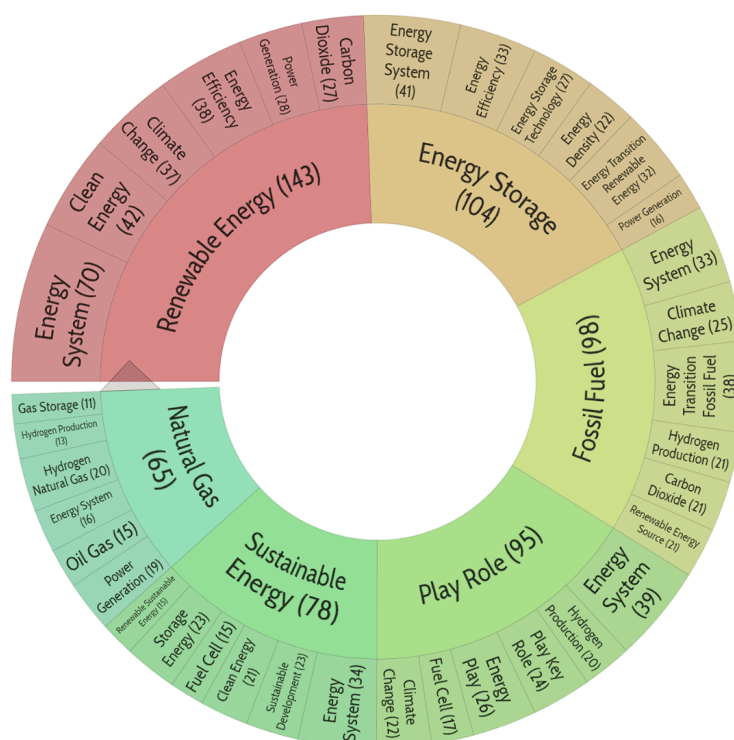


Рис. 10. Распределение публикаций по темам и подтемам для кластера 9

Fig. 10. Distribution of publications by topics and subtopics for the Cluster 9

Выбрана подтема “Energy System” («Энергетическая система?») темы “Renewable Energy” («Возобновляемая энергия»), 35 публикаций.

**Примеры статей,  
соответствующих тематике  
кластера**

*Публикация* “Study on performance of a green hydrogen production system integrated with the thermally activated cooling” [28] («Исследование эффективности системы производства «зеленого водорода», интегрированной с термически контролируемым охлаждением» – перевод автора).

*Краткое содержание.* Возобновляемые источники энергии играют все более важную роль в мировом электроснабжении, однако для крупномасштабного производства требуются экологически безопасные системы хранения энергии. «Зеленый водород» является перспективным вариантом благодаря высокой плотности энергии, незначительной массе, распространенности в окружающей среде и отсутствию парниковых газов. В данной статье представлена численная модель системы производства «зеленого водорода» с использованием теплового накопителя, рекуперирующего тепло, генерируемое всей системой.

*Публикация* “Taxonomy of the fuels in a whole-energy system” [29] («Таксономия видов топлива в комплексной энергетической системе» – перевод автора).

*Краткое содержание.* В статье приводится обзор научной литературы, терминологии стандарта ISO и предлагаются новые определения синтетического топлива, электротоплива и альтернативных видов топлива. Синтетическое топливо – это жидкое топливо, получаемое путем

газификации угля, природного газа или биомассы. Электротопливо аккумулирует электроэнергию в химических соединениях, а альтернативное топливо используется в качестве заменителя ископаемых источников нефти при производстве энергии.

*Примечание:* данная статья, опубликованная в журнале *Frontiers in Energy Research*, представляет собой всеобъемлющий обзор темы, указанной в ее названии, изложенной в терминах, характерных для статей в подкластере “Renewable Energy” AND “Energy System”, которые алгоритм *sumo* классифицировал как «термины, отражающие тему подкластера».

Аналогичным образом подходящие тема и подтема могут быть выбраны в соответствии с интересами конкретного эксперта в данной области.

### **Выводы**

Актуальные темы публикаций связаны с системными проблемами энергетических комплексов, включая интеграцию различных видов источников энергии, систем хранения энергии, таких как «зеленый водород» и «аккумуляторы», и оптимизацию их работы. Это согласуется с трактовкой энергетического перехода, который определяется как переход от ископаемых источников энергии к возобновляемым.

В публикациях на рассматриваемую тему большое внимание уделяется социальным аспектам энергетического перехода, что особенно актуально для сельских районов и регионов с низким уровнем экономического развития. Без финансирования и создания инфраструктуры, необходимой для местных энергетических сообществ, они могут отвергнуть энергетический переход.

Переход на более чистые и безопасные для здоровья источники энергии для отопления и приготовления пищи – одна из главных задач, стоящих перед домохозяйствами. Вопрос в том, как и за счет каких ресурсов эти благие намерения будут реализованы.

Актуальную тему будущего исследования можно сформулировать следующим образом: создание и финансирование разномасштабных энергетических систем, учитывающих

социальные и экономические возможности местных сообществ.

Перспектива исследования:

– проанализировать возможность использования других методов построения словаря для алгоритма GSDMM и предварительной обработки текста;

– проанализировать вопрос системной интеграции в исследованиях в сфере энергетики, возникающий в период энергетического перехода.

### Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

### Список источников

1. *Чигарев Б.Н.* Выявление актуальных задач энергетического перехода в публикациях агрегатора контента научных публикаций Scilit. Часть 1. Кластеризация ключевых слов // Актуальные проблемы нефти и газа. 2024. Т. 15, № 1. С. 72–94. <https://doi.org/10.29222/ipng.2078-5712.2024-15-1.art6>
2. *Mazarura J., de Waal A.* A comparison of the performance of latent Dirichlet allocation and the Dirichlet multinomial mixture model on short text // 2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), Stellenbosch, South Africa, 30 November – 2 December 2016. <https://doi.org/10.1109/RoboMech.2016.7813155>
3. *Abdelmotaleb H., Wojtys M., McNeile C.* A comparison of a novel optimized GSDMM Model with K-means clustering for topic modelling of free text // Journal of Machine Intelligence and Data Science. 2023. Vol. 4. P. 52–62. <https://doi.org/10.11159/jmids.2023.07>
4. *Agarwal N., Sikka G., Awasthi L.K.* Evaluation of web service clustering using Dirichlet Multinomial Mixture model based approach for Dimensionality Reduction in service representation // Information Processing & Management. 2020. Vol. 57, No. 4. P. 102238. <https://doi.org/10.1016/j.ipm.2020.102238>
5. *Chigarev B.* Social aspects in Energy Research & Social Science journal publications for 2019–2023. Bibliometric analysis // Preprints. 2023. <https://doi.org/10.20944/preprints202309.0744.v1>
6. *Osiński S., Stefanowski J., Weiss D.* Lingo: Search results clustering algorithm based on singular value decomposition // Advances in Soft Computing. 2004. Vol. 25. P. 359–368. [https://doi.org/10.1007/978-3-540-39985-8\\_37](https://doi.org/10.1007/978-3-540-39985-8_37)
7. *Mustikasari D.* Analisis tema skripsi mahasiswa menggunakan document clustering dengan algoritma Lingo [Analysis of student thesis themes using document clustering with Lingo algorithm] // KINETIK. 2017. Vol. 2, No. 2. P. 131–140. <https://doi.org/10.22219/kinetik.v2i2.180>
8. *Mustika D., Adji T.B., Kadir A.* Analisis potensi daerah melalui metode document clustering Laporan pelaksanaan kegiatan Kuliah kerja nyata-pembelajaran pemberdayaan masyarakat [Analyzing regional potential through document clustering method: Report on the implementation of learning

activities. Lecture on the real work – community empowerment] // *Jurnal Edukasi dan Penelitian Informatika*. 2015. Vol. 1, No. 1. <https://doi.org/10.26418/jp.v1i1.9978>

9. *Rama Rao G., Avanikeeth, Prudvini Ch.* et al. News text summarization based on multi-feature and fuzzy logic // *International Journal of Advanced Research in Science, Communication and Technology*. 2020. Vol. 3, No. 1. P. 47–51. <https://doi.org/10.48175/IJARSCT-6213>

10. *Yin J., Wang J.* A Dirichlet multinomial mixture model-based approach for short text clustering // *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computer Machinery, 2014. P. 233–242. <https://doi.org/10.1145/2623330.2623715>

11. *Cunningham H., Tablan V., Roberts A., Bontcheva K.* Getting more out of biomedical documents with GATE's full lifecycle open source text analytics // *PLoS Computational Biology*. 2013. Vol. 9, No. 2. P. e1002854. <https://doi.org/10.1371/journal.pcbi.1002854>

12. *Ali M., Seraj M.* Nexus between energy consumption and carbon dioxide emission: evidence from 10 highest fossil fuel and 10 highest renewable energy-using economies // *Environmental Science and Pollution Research*. 2022. Vol. 29, No. 58. P. 87901–87922. <https://doi.org/10.1007/s11356-022-21900-9>

13. *Das N., Murshed M., Rej S.* et al. Can clean energy adoption and international trade contribute to the achievement of India's 2070 carbon neutrality agenda? Evidence using quantile ARDL measures // *International Journal of Sustainable Development & World Ecology*. 2023. Vol. 30, No. 3. P. 262–277. <https://doi.org/10.1080/13504509.2022.2139780>

14. *Sovacool B.K., Brugger H., Brunzema I.* et al. Social innovation supports inclusive and accelerated energy transitions with appropriate governance // *Communications Earth & Environment*. 2023. Vol. 4, No. 1. P. 289. <https://doi.org/10.1038/s43247-023-00952-w>

15. *Bresciani S., Rizzo F., Deserti A.* Toward a comprehensive framework of social innovation for climate neutrality: A systematic literature review from business/production, public policy, environmental sciences, energy, sustainability and related fields // *Sustainability*. 2022. Vol. 14, No. 21. P. 13793. <https://doi.org/10.3390/su142113793>

16. *Ma W., Vatsa P., Zheng H.* Cooking fuel choices and subjective well-being in rural China: Implications for a complete energy transition // *Energy Policy*. 2022. Vol. 165. P. 112992. <https://doi.org/10.1016/j.enpol.2022.112992>

17. *Guta D., Baumgartner J., Jack D.* et al. A systematic review of household energy transition in low and middle income countries // *Energy Research & Social Science*. 2022. Vol. 86. P. 102463. <https://doi.org/10.1016/j.erss.2021.102463>

18. *Kabeyi M.J.B., Olanrewaju O.A.* Sustainable energy transition for renewable and low carbon grid electricity generation and supply // *Frontiers in Energy Research*. 2022. Vol. 9. P. 743114. <https://doi.org/10.3389/fenrg.2021.743114>

19. *Oyewo A.S., Aghahosseini A., Ram M.* et al. Pathway towards achieving 100% renewable electricity by 2050 for South Africa // *Solar Energy*. 2019. Vol. 191. P. 549–565. <https://doi.org/10.1016/j.solener.2019.09.039>

20. *Li M., Trencher G., Asuka J.* The clean energy claims of BP, Chevron, ExxonMobil and Shell: A mismatch between discourse, actions and investments // *PLoS ONE*. 2022. Vol. 17, No. 2. P. e0263596. <https://doi.org/10.1371/journal.pone.0263596>

21. *Overland I., Loginova J.* The Russian coal industry in an uncertain world: Finally pivoting to Asia? // *Energy Research & Social Science*. 2023. Vol. 102. P. 103150. <https://doi.org/10.1016/j.erss.2023.103150>
22. *Drożdż W., Mróz-Malik O., Kopiczko M.* The future of the Polish energy mix in the context of social expectations // *Energies*. 2021. Vol. 14, No. 17. P. 5341. <https://doi.org/10.3390/en14175341>
23. *Manfren M., Nastasi B., Groppi D., Astiaso Garcia D.* Open data and energy analytics – An analysis of essential information for energy system planning, design and operation // *Energy*. 2020. Vol. 213. P. 118803. <https://doi.org/10.1016/j.energy.2020.118803>
24. *Meng H., Jia H., Xu T.* et al. Battery storage configuration of AC/DC hybrid distribution networks // *CSEE Journal of Power and Energy Systems*. 2023. Vol. 9, No. 3. P. 859–872. <https://doi.org/10.17775/CSEEJPES.2021.07630>
25. *Citalingam K., Go Y.I.* Hybrid energy storage design and dispatch strategy evaluation with sensitivity analysis: Techno-economic-environmental assessment // *Energy Storage*. 2022. Vol. 4, No. 5. P. e353. <https://doi.org/10.1002/est2.353>
26. *Romero-Castro N., Ángeles López-Cabarcos M., Miramontes-Viña V., Ribeiro-Soriano D.* Sustainable energy transition and circular economy: The heterogeneity of potential investors in rural community renewable energy projects // *Environment, Development and Sustainability*. 2023. <https://doi.org/10.1007/s10668-022-02898-z>
27. *Broska L.H., Vögele S., Shamon H., Wittenberg I.* On the future(s) of energy communities in the German energy transition: A derivation of transformation pathways // *Sustainability*. 2022. Vol. 14, No. 6. P. 3169. <https://doi.org/10.3390/su14063169>
28. *Moulebe L.P., Touati A., Akpoviro E.O., Rabbah N.* Study on performance of a green hydrogen production system integrated with the thermally activated cooling // *Acta Innovations*. 2022. No. 47. P. 5–19. <https://doi.org/10.32933/ActaInnovations.47.1>
29. *Rixhon X., Limpens G., Jeanmart H., Contino F.* Taxonomy of the fuels in a whole-energy system // *Frontiers in Energy Research*. 2021. Vol. 9. P. 660073. <https://doi.org/10.3389/fenrg.2021.660073>

#### **Информация об авторе**

*Борис Николаевич Чигарев* – к.ф.-м.н., ведущий инженер по научно-технической информации, Институт проблем нефти и газа РАН, Москва, Россия; SPIN-код: 7610-8398, <https://orcid.org/0000-0001-9903-2800>; e-mail: [bchigarev@ipng.ru](mailto:bchigarev@ipng.ru)

**Поступила в редакцию 26.02.2024**

**ENERGY RESOURCES OF THE EARTH'S CRUST: CHALLENGES OF THE MODERN ECONOMY**

Original article

<https://doi.org/10.29222/ipng.2078-5712.2024-15-2.art5>**Identification of topical issues of the energy transition in publications of the scientific content aggregator Scilit.  
Part 2. Clustering of publications****B.N. Chigarev** ✉

Oil and Gas Research Institute, Russian Academy of Sciences, Moscow, Russia

**Abstract.** *Background.* The access of Russian researchers to Scopus and Web of Science has become restricted, so the use of open reference databases becomes relevant. *Objective.* Identification of topical problems of energy transition in publications presented in Scilit, a content aggregator for scientific publications with free access. *Materials and methods.* The study utilized 10,121 bibliometric records of articles from 2019–2023. Publications were systematized using Gibbs sampling algorithm for Dirichlet mixture model. The topics of publications within the obtained clusters were analyzed using the demo version of the Carrot2 program. Publications were ranked using the sumy utility with the lex-rank algorithm. *Results.* The identified topical topics are devoted to systemic problems of energy complexes, including integration of different sources of energy generation, energy storage in “accumulators” or “green hydrogen” and optimization of their operation. Much attention is paid to the social aspects of the energy transition, especially relevant for rural areas and regions with a low level of economic development. *Conclusions.* Without financial support and appropriate infrastructure for local energy communities, the energy transition may be rejected by them. Households should be encouraged to use cleaner energy sources that are less harmful to health and the environment.

**Keywords:** energy transition, relevant research topics, Scilit content aggregator, bibliometric analysis

**Funding:** the work was funded by the Ministry of Science and Higher Education of the Russian Federation (State Assignment No. 122022800270-0).

**For citation:** Chigarev B.N. Identification of topical issues of the energy transition in publications of the scientific content aggregator Scilit. Part 2. Clustering of publications. *Actual Problems of Oil and Gas*. 2024. Vol. 15, No. 2. P. 174–199. (In Russ.). <https://doi.org/10.29222/ipng.2078-5712.2024-15-2.art5>

---

✉ Boris N. Chigarev, e-mail: [bchigarev@ipng.ru](mailto:bchigarev@ipng.ru)

© Chigarev B.N., 2024



The content is available under Creative Commons Attribution 4.0 License.

### Conflict of interests

The author declares no conflict of interests.

### References

1. Chigarev B.N. Identification of topical issues of the energy transition in publications of the scientific content aggregator Scilit. Part 1. Keyword clustering. *Actual Problems of Oil and Gas*. 2024. Vol. 15, No. 1. P. 72–94. (In Russ.). <https://doi.org/10.29222/ipng.2078-5712.2024-15-1.art6>
2. Mazarura J., de Waal A. A comparison of the performance of latent Dirichlet allocation and the Dirichlet multinomial mixture model on short text. In: *2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*, Stellenbosch, South Africa, 30 November – 2 December 2016. <https://doi.org/10.1109/RoboMech.2016.7813155>
3. Abdelmotaleb H., Wojtys M., McNeile C. A comparison of a novel optimized GSDMM Model with K-means clustering for topic modelling of free text. *Journal of Machine Intelligence and Data Science*. 2023. Vol. 4. P. 52–62. <https://doi.org/10.11159/jmids.2023.07>
4. Agarwal N., Sikka G., Awasthi L.K. Evaluation of web service clustering using Dirichlet Multinomial Mixture model based approach for Dimensionality Reduction in service representation. *Information Processing & Management*. 2020. Vol. 57, No. 4. P. 102238. <https://doi.org/10.1016/j.ipm.2020.102238>
5. Chigarev B. Social aspects in Energy Research & Social Science journal publications for 2019–2023. Bibliometric analysis. *Preprints*. 2023. <https://doi.org/10.20944/preprints202309.0744.v1>
6. Osiński S., Stefanowski J., Weiss D. Lingo: Search results clustering algorithm based on singular value decomposition. *Advances in Soft Computing*. 2004. Vol. 25. P. 359–368. [https://doi.org/10.1007/978-3-540-39985-8\\_37](https://doi.org/10.1007/978-3-540-39985-8_37)
7. Mustikasari D. Analisis tema skripsi mahasiswa menggunakan document clustering dengan algoritma Lingo [Analysis of student thesis themes using document clustering with Lingo algorithm] // *KINETIK*. 2017. Vol. 2, No. 2. P. 131–140. <https://doi.org/10.22219/kinetik.v2i2.180>
8. Mustika D., Adji T.B., Kadir A. Analisis potensi daerah melalui metode document clustering Laporan pelaksanaan kegiatan Kuliah kerja nyata-pembelajaran pemberdayaan masyarakat [Analyzing regional potential through document clustering method: Report on the implementation of learning activities. Lecture on the real work – community empowerment]. *Jurnal Edukasi dan Penelitian Informatika*. 2015. Vol. 1, No. 1. <https://doi.org/10.26418/jp.v1i1.9978>
9. Rama Rao G., Avanikeeth, Prudvini Ch. et al. News text summarization based on multi-feature and fuzzy logic. *International Journal of Advanced Research in Science, Communication and Technology*. 2020. Vol. 3, No. 1. P. 47–51. <https://doi.org/10.48175/IJARST-6213>
10. Yin J., Wang J. A Dirichlet multinomial mixture model-based approach for short text clustering. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computer Machinery, 2014. P. 233–242. <https://doi.org/10.1145/2623330.2623715>
11. Cunningham H., Tablan V., Roberts A., Bontcheva K. Getting more out of biomedical documents with GATE's full lifecycle open source text analytics. *PLoS Computational Biology*. 2013. Vol. 9, No. 2. P. e1002854. <https://doi.org/10.1371/journal.pcbi.1002854>



12. Ali M., Seraj M. Nexus between energy consumption and carbon dioxide emission: evidence from 10 highest fossil fuel and 10 highest renewable energy-using economies. *Environmental Science and Pollution Research*. 2022. Vol. 29, No. 58. P. 87901–87922. <https://doi.org/10.1007/s11356-022-21900-9>
13. Das N., Murshed M., Rej S. et al. Can clean energy adoption and international trade contribute to the achievement of India's 2070 carbon neutrality agenda? Evidence using quantile ARDL measures. *International Journal of Sustainable Development & World Ecology*. 2023. Vol. 30, No. 3. P. 262–277. <https://doi.org/10.1080/13504509.2022.2139780>
14. Sovacool B.K., Brugger H., Brunzema I. et al. Social innovation supports inclusive and accelerated energy transitions with appropriate governance. *Communications Earth & Environment*. 2023. Vol. 4, No. 1. P. 289. <https://doi.org/10.1038/s43247-023-00952-w>
15. Bresciani S., Rizzo F., Deserti A. Toward a comprehensive framework of social innovation for climate neutrality: A systematic literature review from business/production, public policy, environmental sciences, energy, sustainability and related fields. *Sustainability*. 2022. Vol. 14, No. 21. P. 13793. <https://doi.org/10.3390/su142113793>
16. Ma W., Vatsa P., Zheng H. Cooking fuel choices and subjective well-being in rural China: Implications for a complete energy transition. *Energy Policy*. 2022. Vol. 165. P. 112992. <https://doi.org/10.1016/j.enpol.2022.112992>
17. Guta D., Baumgartner J., Jack D. et al. A systematic review of household energy transition in low and middle income countries. *Energy Research & Social Science*. 2022. Vol. 86. P. 102463. <https://doi.org/10.1016/j.erss.2021.102463>
18. Kabeyi M.J.B., Olanrewaju O.A. Sustainable energy transition for renewable and low carbon grid electricity generation and supply. *Frontiers in Energy Research*. 2022. Vol. 9. P. 743114. <https://doi.org/10.3389/fenrg.2021.743114>
19. Oyewo A.S., Aghahosseini A., Ram M. et al. Pathway towards achieving 100% renewable electricity by 2050 for South Africa. *Solar Energy*. 2019. Vol. 191. P. 549–565. <https://doi.org/10.1016/j.solener.2019.09.039>
20. Li M., Trencher G., Asuka J. The clean energy claims of BP, Chevron, ExxonMobil and Shell: A mismatch between discourse, actions and investments. *PLoS ONE*. 2022. Vol. 17, No. 2. P. e0263596. <https://doi.org/10.1371/journal.pone.0263596>
21. Overland I., Loginova J. The Russian coal industry in an uncertain world: Finally pivoting to Asia? *Energy Research & Social Science*. 2023. Vol. 102. P. 103150. <https://doi.org/10.1016/j.erss.2023.103150>
22. Drożdż W., Mróz-Malik O., Kopiczko M. The future of the Polish energy mix in the context of social expectations. *Energies*. 2021. Vol. 14, No. 17. P. 5341. <https://doi.org/10.3390/en14175341>
23. Manfren M., Nastasi B., Groppi D., Astiaso Garcia D. Open data and energy analytics – An analysis of essential information for energy system planning, design and operation. *Energy*. 2020. Vol. 213. P. 118803. <https://doi.org/10.1016/j.energy.2020.118803>
24. Meng H., Jia H., Xu T. et al. Battery storage configuration of AC/DC hybrid distribution networks. *CSEE Journal of Power and Energy Systems*. 2023. Vol. 9, No. 3. P. 859–872. <https://doi.org/10.17775/CSEEJPES.2021.07630>

25. Citalingam K., Go Y.I. Hybrid energy storage design and dispatch strategy evaluation with sensitivity analysis: Techno-economic-environmental assessment. *Energy Storage*. 2022. Vol. 4, No. 5. P. e353. <https://doi.org/10.1002/est2.353>

26. Romero-Castro N., Ángeles López-Cabarcos M., Miramontes-Viña V., Ribeiro-Soriano D. Sustainable energy transition and circular economy: The heterogeneity of potential investors in rural community renewable energy projects. *Environment, Development and Sustainability*. 2023. <https://doi.org/10.1007/s10668-022-02898-z>

27. Broska L.H., Vögele S., Shamon H., Wittenberg I. On the future(s) of energy communities in the German energy transition: A derivation of transformation pathways. *Sustainability*. 2022. Vol. 14, No 6. P. 3169. <https://doi.org/10.3390/su14063169>

28. Moulebe L.P., Touati A., Akpoviro E.O., Rabbah N. Study on performance of a green hydrogen production system integrated with the thermally activated cooling. *Acta Innovations*. 2022. No. 47. P. 5–19. <https://doi.org/10.32933/ActaInnovations.47.1>

29. Rixhon X., Limpens G., Jeanmart H., Contino F. Taxonomy of the fuels in a whole-energy system. *Frontiers in Energy Research*. 2021. Vol. 9. P. 660073. <https://doi.org/10.3389/fenrg.2021.660073>

#### ***Information about the author***

*Boris N. Chigarev* – Cand. Sci. (Phys.-Math.), Leading Engineer on Scientific and Technical Information, Oil and Gas Research Institute, Russian Academy of Sciences, Moscow, Russia; <https://orcid.org/0000-0001-9903-2800>; e-mail: [bchigarev@ipng.ru](mailto:bchigarev@ipng.ru)

**Received 26.02.2024**