

Методически инструментариум за анализ на публикации, свързани с развитието на устойчивите логистични системи

Николай Драгомиров*, Светлана
Александрова-Златанска**, Лиляна
Михова***

Резюме: Устойчивото развитие на логистичните системи се утвърждава в ключов изследователски приоритет в контекста на засилващи се икономически, екологични и социални предизвикателства, регулаторен натиск и бърза дигитална трансформация. В условията на нарастващ обем научни публикации и разширяващи се масиви от структурирани и неструктурирани данни възниква необходимост от усъвършенстване на методическия инструментариум за анализ на литературата в областта на устойчивите логистични системи.

Целта на настоящата разработка е да обхване съвременните технологични и аналитични подходи за изследване на научни публикации, да систематизира водещите методи и да предложи разширен модел за литературен анализ, приложим към големи корпуси от

индексирани трудове. В статията са обобщени редица класически виждания и методи, като те са надградени с възможностите на съвременните информационни системи. Разгледани са решения за извличане на данни от индексирани системи, алгоритми за филтриране и автоматизирано изчистване, техники за предварителна текстообработка, клъстеризация, включително и прилагане на модели за прогнозиране с цел очертаване на бъдещото развитие в отделни изследователски течения. Акцент е поставен и върху ролята на системите с изкуствен интелект и тяхното приложение, както и върху значението на триангулацията между количествени и качествени техники.

Резултатът е извличане и обобщение на информация от научни публикации чрез разработен интегриран методически инструментариум. Решението може да бъде прилагано не само към тематиката на устойчивите логистични системи, но и в други научни

* Николай Драгомиров е доктор, доцент в катедра „Логистика и вериги на доставките“ на УНСС.

** Светлана Александрова-Златанска е доктор, професор в катедра „МИО и бизнес“ на УНСС.

*** Лиляна Михова е доктор, главен асистент в катедра „Логистика и вериги на доставките“ на УНСС.

области, подпомагайки изследователите при откриване на тенденции, дефицити и перспективи за развитие на бъдещи изследвания.

Ключови гуми: устойчиви логистични системи, библиометрия, текстов анализ.

JEL: M11, Q01.

1. Въведение

Логистичните системи се утвърдиха като системи с ключова роля за функционирането на съвременната икономика. Те осигуряват движението на суровини, междинни и крайни продукти по цялата верига на стойността и свързват производители, търговци и крайни потребители в национален и глобален мащаб. Пандемичната криза от 2020 г. и несигурността в международната търговия показва, че проблемът за устойчивостта на доставките и на логистичните системи стана стратегически приоритет за бизнеса. Темата за устойчивото развитие на веригите на доставка и на логистичните системи е изключително актуална в условията на засилващи се социални, екологични и икономически предизвикателства на глобално ниво. Нарастващата обществена чувствителност към проблеми като климатичните промени, изчерпването на природните ресурси и социалната отговорност на бизнеса води до повишени очаквания към компаниите за прилагане на устойчиви и екологосъобразни практики по цялата верига на доставки.

Устойчивото развитие на логистичните операции и веригите на доставка е обсъждана тема за изследователи и

практици, тъй като компаниите от логистичния сектор осъзнават необходимостта от балансиране на икономическия растеж с екологичната и социалната отговорност. Актуалността на проучването на дискутирани теми за устойчивостта на логистичните операции в научните публикации се засилва и от въвеждането на ESG стандартите и изискванията за прозрачност, отчетност и контрол върху екологичните и социалните въздействия на веригите на доставка. Бързият технологичен прогрес, дигитализацията и използването на изкуствен интелект създават нови възможности за оптимизация на логистичните процеси, за намаляване на ресурсната интензивност и ограничаване на негативното въздействие върху околната среда.

Научните публикации и анализи, посветени на устойчивото развитие на логистичните системи и на веригите на доставка, имат съществено значение както за развитието на теорията, така и за усъвършенстването на управленската практика в логистичния сектор. Те дават задълбочено разбиране на взаимовръзките между икономическата ефективност, екологичната отговорност и социалните аспекти на логистичната дейност, като по този начин разширяват концептуалната рамка за ефективността на управлението на веригите на доставка.

В научните изследвания се срещат различни разбирания за същността на устойчивостта на логистичните системи. Устойчивостта започва да се интегрира в логистичните системи и веригите на доставки скоро след появата на концепцията за устойчиво

развитие. Дефиниции за устойчива логистика и устойчиви вериги на доставки се появяват в края на 90-те години и началото на 2000-те години, когато вече съществуват множество изследвания, посветени на тази проблематика. Някои от тях представляват обширни литературни прегледи (Koberg & Longoni, 2019; Seuring & Müller, 2008; Touboulic & Walker, 2015), които се фокусират върху концептуализацията и развитието на устойчивото управление на веригите на доставки (УУВД). Те очертават ключови понятия, свързани с интегрирането на устойчивостта в традиционното управление на веригите на доставки. Други автори подчертават значението на концепцията за “triple bottom line” на устойчивото развитие, която обхваща икономическите, екологичните и социалните измерения (Govindan et al., 2013; Shou et al., 2019; Varriale et al., 2023). Скорошна публикация пък показва, че до наши дни УУВД се е утвърдило като трансформиращо и самостоятелно направление, което интегрира икономически, екологични и социални измерения и се превръща в системен отговор на климатични кризи и регулаторни кризи с цел справяне с ключови предизвикателства за постигане на устойчиво бъдеще (Mohsin et al., 2025). В този контекст на устойчивостта в логистичните системи и веригите на доставки може да се говори за „устойчиви логистични системи“.

Така в съвременните изследвания се забелязва изместване на фокуса от устойчиви логистични практики и устойчиви вериги на доставки към по-големи и устойчиви логистични системи, които обхващат и самата

инфраструктура и свързаните с нея проблеми. Този подход е логично продължение на естеството на съвременната икономика и съответната необходимост логистиката да бъде трансформирана в основен инструмент за постигане на устойчиво развитие.

Значимостта на устойчивите логистични системи се потвърждава както от емпирични изследвания (Ji et al., 2023), така и от мащабни библиометрични обзори, подходи за картографиране и др., базирани на използването на голям брой научни публикации с отворен достъп (Ren et al., 2020; Tabanfeh et al., 2025). Те отчитат ясно и засилено нарастване на научния интерес към технологично развити и устойчиви логистични системи. В допълнение Sun et al. (2022) отчитат, че изследванията върху устойчивите логистични системи отбелязват ускорено развитие, характеризиращо се с интегриране на Industry 4.0 технологии, засилен фокус върху дигитализация и автоматизация за повишаване на икономическата, екологичната и социалната ефективност, ориентация към системни модели на устойчивост.

Комбинирането на термините устойчиви логистични системи и УУВД в мащабни изследвания за текстово извличане помага за идентифициране на основните тенденции и рамки. Този подход предоставя нови прозрения в областта, изследва ключови тенденции и нови теми и разкрива тяхната динамична промяна. В изследването на Mohsin et al. (2025), в което е направен много обстоен систематичен преглед и библиометричен анализ не само на литературата, но и на използваните

методи в областта на устойчивите вериги на доставки, се казва, че се използват разнообразни методи, включително качествени и количествени (анкети, икономически модели, математическо моделиране и оптимизации) и смесени подходи, като напоследък се интегрират технологии като големи данни, машинно обучение и изкуствен интелект за подобряване на анализа и прогнозите в устойчивостта на веригите на доставки. Това от своя страна дава ясен знак за посоката на награждане на методите, както и тяхната трансформация. Например утвърдени методи като провеждане на литературен анализ, каквото са систематичните ревюта (Systematic review), могат да бъдат допълнително наградени. Именно поради това, а и предвид актуалността на темата, **целта на разработката** е да се направят редица обобщения за методическия инструментариум за литературен анализ на публикации, свързани с развитието на устойчивите логистични системи. Това от своя страна е предпоставка за развитие на последващи изследвания в областта и темата за устойчивостта на логистичните системи, както и приложение на идеите в групи научни теми. Основните изследователски въпроси, които могат да бъдат посочени, са:

- Какви технологични решения се прилагат за анализ на научни публикации?
- До какъв вид резултат може да се достигне от прегледа на научни публикации и методи за анализ, които могат да бъдат използвани

В последващи научни и приложни изследвания?

- Какви са перспективите за развитие на научни изследвания в условията на богатство от данни и информация, както и при настоящата динамика на развитие на информационни технологии?

2. Процес и технологични решения за изследване на публикации в контекста на устойчивите логистични системи

Определено фокусът на използваните методи в научните изследвания се измества, като посоката е от използване предимно на количествени данни към това да се разширява включването и на качествени и неструктурирани масиви (Dragomirov, 2024). Това от своя страна предпоставя редица нови възможности за изследователите, включително и в частта на провеждането на литературни анализи. Ако в класическите подходи се обхващат до няколкостотин публикации, то със съвременните технологични възможности може да се работи с хиляди такива. Независимо от това може да се посочи, че в съвременните условия стъпките на подобен вид анализ се дефинират в следната обобщена от практиките рамка (Dragomirov, 2022):

- Разработване на протокол за провеждане на анализа;
- Идентифициране на публикации;
- Оценка и избор на публикации, които да бъдат включени в анализа;
- Извличане на данни от публикации и изграждане на база данни с резултати;

Дигитализация

- Представяне на данните.

Следвайки тази логика и в контекста на устойчивите логистични системи, могат да бъдат направени редица допълнителни уточнения. Определено водещ момент е ясно дефиниране на целите на изследването и на основните изследователски въпроси, правила и пр. Подобен въпрос би могъл да бъде дефинирането на основните изследователски направления за определен период от време или, в контекста на устойчивите логистични системи – кои са основните направления на изследване и какви са съответните тенденции? По отношение на втората стъпка и идентифицирането на публикациите могат да бъдат приложени различни видове подходи, които са релевантни към момента на написването на статията:

- Извличането на записи от индексирани системи посредством използване на ключови думи.
- Извличане на записи от индексирани системи, като се използват системи, базирани на изкуствен интелект. Включително и системи, обхващащи връзките на цитиране между публикациите.

И в двата случая следва да се вземат две допълнителни решения. Първото касае използването на пълнотекстови записи или само на резюмета, а второто е свързано с релевантността на записите. По отношение на първия въпрос може да се посочи, че при работата с големи масиви записи е удачно използването на абстрактите на трудовете, защото те следва да синтезират основните съдържателни

моменти (APA, 2020) и затова могат да бъдат използвани като алтернатива на пълните текстове. Отделна мотивация за това е ограничеността на достъпа до пълния текст на съответната публикация. По отношение на втория въпрос, касаещ релевантността на получените записи, може да се маркира необходимостта от използване на софтуерни инструменти за премахване на дублирания, както и филтриране на публикации (Dragomirov, 2022). Като подобно решение може да бъде посочен проектът ASReview LAB, който използва ефективен алгоритъм за сортиране на базата данни посредством използване на machine learning техники. Това по същество представлява процеса по оценка и избор на публикации, които да бъдат включени в анализа. Подобна е логиката и на утвърдения в научната практика модел PRISMA, който предлага изключване на публикации като част от неговите стъпки (Moher et al., 2009). Допълнителни видове филтриране може да бъдат приложени според целта на изследването, като такива са филтриране според вида на публикациите (статии или доклади в конференци), видове издания, регионални ограничения, видове индексация на публикациите, брой цитирания и пр.

Последващият етап от модела е извличането на данни от публикациите и изграждане на база данни с резултатите. При анализа на текстове възможностите за обработка могат да бъдат класифицирани според степента на използване на ког. В единия случай системите са напълно графично осигурени с GUI (graphical user interface), докато във

Втория се изискват познания по програмиране (Dragomirov, 2024):

- Отгелни прогукти – WordStat, MonkeyLearn, Lexalytics, RapidMiner, KNIME, IBM Watson NLP, Orange, Google Cloud Natural Language API и др.
- Библиотеки, които в повечето случаи са за Python – SpaCy, NLTK (Natural Language Toolkit), TextBlob, Gensim, Scikit-learn, PyTorch/NLP + TensorFlow и пр.

При извличането на записи от индексирани системи могат да бъдат извлечени редица данни, като заглавия, издания, година на публикуване, абстракт, ключови думи, както и различни видове данни за геолокация, брой идентифицирани цитирания, дигитален идентификатор (digital object identifier – DOI) и пр. В настоящия модел се визира предимно използване на Python в среда на Jupyter Lab и използване на среда, предоставена от Anaconda Distribution за инсталация и контрол на библиотеките. Същевременно като входящи данни се използват предимно метаданни – година на публикуване, абстракт, ключови думи, данни за авторите, но могат да бъдат включени данни за налични цитирания, връзки и други метрики. Подобен аутпут се предлага от разпространените бази данни за индексирани публикации, като в повечето случаи има възможност за сваляне на базата в различни формати, като CSV (Comma-separated values) и RIS (Research Information Systems), включително и посредством предоставяне на директен достъп през API (An application programming interface). Конкретният метод зависи от това дали се предвижда

използване на reference manager като междинна стъпка, както и в каква среда ще бъдат проведени анализите.

В така формираната база анализите са в две базови направления – анализ по ключови думи и анализ на текстове от абстракти. По своята същност това са относително различни материали за обработка, въпреки че подходите са донякъде сходни. При ключовите думи съществува висока степен на конкретизация и яснота, докато при абстрактите е необходимо предварително да се извлече смисловото съдържание на текста. Независимо от това и за двата подхода е валидна предварителната обработка на данните – премахване на текстове, специални символи, данни за издатели и пр., които нямат отношение към процеса на анализ.

При работата с ключови думи е възможно процесът на токенизация да не се прилага, а да се използва заместване и свързване на ключови думи с цел подобряване на консистентността, стандартизиране на терминологията, уеднаквяване на синоними и премахване на грешки. Това е допустимо, защото при дефинирането на ключови думи авторите и издателите прецизират във висока степен тяхното използване. В случая се визира премахване на множествени числа, като supply chain и supply chains, замяната на “IoT” с “Internet of Things”, “Blockchain Technology” с “Blockchain”, “ЕС” с “Европейски съюз” и пр. Според конкретния случай е удачно преминаването към малки букви, както и замяната на празните полета с “_”. Неразделна част от този предварителен етап е премахването на цели думи, които имат прекалено висока честота

Дигитализация

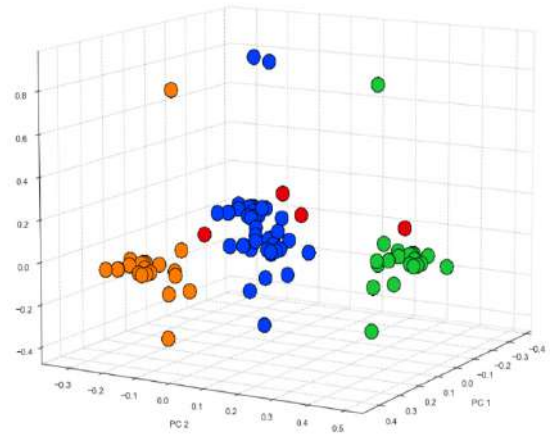
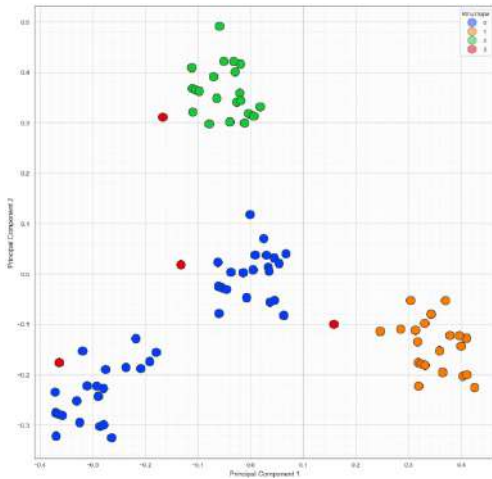
или пък на такива, които имат много ниска. На този етап е изключително важно познаването на проблемите на устойчивостта на логистичните системи с цел по-точното дефиниране на правилните речници и списъци. На последващия етап е удачно прилагането на класически подходи, като извеждане на различни видове разпределения, като едни от най-честите са (според възможния аутпут на индексиранията система):

- По гържави и институции – общ брой гуми, водещи ключови гуми и пр.;
- По години – приоритетни ключови гуми по години;
- Извеждане на n-grams – в повечето случаи bigrams и trigrams;
- Генериране на bag of words, wordcloud и пр.;
- Мрежови връзки и други.

Във всички случаи получените резултати позволяват различни видове обединения и проследяване на тенденции, както и направата на сечения между тях. По-съществен е обаче моментът, при който се преминава към търсене на по-големи обединения и идентифициране на евентуални клъстери. В случая на анализ на ключови гуми могат да бъдат използвани различни видове методи, но сред базовите е този на k-means клъстеризация, след съответната векторизация. Обикновено той е предпочитан поради своята базовост и разпространеност (Tussipov et al., 2025), като същевременно подходът е удачен и при работата с големи масиви от данни (Kulshrestha & Santani, 2024). При него се групират сходни ключови

гуми въз основа на техните векторни представяния в предварително зададен брой клъстери (k). За точния брой клъстери може да се използва цикъл за проследяване на техните силуети или други характеристики, както и да се приложи elbow method посредством съответната библиотека.

При визуализацията на клъстерите, в практиката често се прилагат техники за редукция на измеренията, за да се преодолеят предизвикателствата, свързани с високата размерност на данните, както и да се улесни графичното представяне на резултатите. В този контекст е удачно прилагането на PCA (Principal Component Analysis), което позволява намаляване на броя на характеристиките, като същевременно се запазва максимално голяма част от дисперсията в данните. PCA се използва основно като инструмент за визуализация и интерпретация, без да променя качеството на самата клъстеризация. Въпреки редицата особености решението има удачен характер при анализа на подобни видове данни. Като алтернатива на k-means може да се използва Latent Dirichlet Allocation (LDA), който е сред базовите методи в теорията и практиката на Natural Language Processing (NLP) и други. Независимо от избрания подход е възможна примерна визуализация в 2D, която да бъде представена и в 3D, като примерни визуализации са посочени на фигура 1. В двата случая всяка ос в пространството може да бъде описана от комбинация от гуми с цел извлечение на семантичното значение на измерението. По същество това са едни и същи клъстери, но са представени в



Фигура 1. Илюстративни визуализации на клъстери в 2D и последваща друга перспектива в 3D пространство
 Източник: авторите

Таблица 1. Структура на новообразувана таблица

година\клъстер	0	1	...	n
2010	брой публикации	брой публикации	брой публикации	брой публикации
2011	брой публикации	брой публикации	брой публикации	брой публикации
...	брой публикации	брой публикации	брой публикации	брой публикации
2025	брой публикации	брой публикации	брой публикации	брой публикации

Източник: авторите

различни проекции и измерения и може да се анализира тяхното визуално разположение.

След дефинирането на клъстерите и техния запис в отделна колона в базата данни се създава възможност и за допълнителен анализ. Една от възможностите е извеждане на дескриптивна статистика за съдържанието на клъстерите като ключови думи. Успоредно с това се създава възможност за обобщаване на данните, като брой публикации за всеки клъстер и по години. Това от своя страна дефинира нова база със структура от вида, както е

посочено в таблица 1, която предоставя възможности за напрева и на прогнози.

В случая могат да се използват познатите в практиката примери за различните видове пълзящи средни, както и използването на методи, базирани на експоненциално изглаждане и авторегресионните модели ARIMA (Мишев & Гоев, 2010). Като цяло библиометричните данни имат потенциал за подобен вид прогнози (Song & Cao, 2022). В тази посока разсъжденията могат да се развият и възможностите за прогнози да се обобщят в две направления:

Дигитализация

- Прогнозиране на публикационната активност в съответния клъстер;
- Прогнозиране на появата на отделни гуми, самостоятелно или в рамките на клъстерите.

Принципно резултатите следва да се интерпретират правилно, тъй като прогнозите най-вероятно ще имат редица слабости по отношение на тяхното качество и надеждност според метриците на моделите; Въпреки това те са ценен ориентир за идентифициране на генерализираната посока на движение.

Формирането на клъстер от гуми позволява и други анализи, например изследване на мрежи (networks). При тях отделни гуми или клъстери могат да се разглеждат като елементи от системи, между които съществуват връзки. По този начин, освен подобрената визуализация, могат да се изследват дефицити във връзките между отделните елементи. Например между кои базови елементи на устойчивите логистични системи се наблюдават слаби връзки и съответно това са дефицитите, които мотивират насочването на нови изследвания в тези микрообласти.

Дотук посочените виждания се отнасят за работата с ключови гуми. В така формираните бази данни неразделна част са резюметата на публикациите, от които могат да бъдат извлечени допълнителни опорни точки в анализа. Особеност тук е използването на токенизация (Tokenization). Този процес се определя в теорията и практиката като фундаментална стъпка за Natural Language Processing (NLP). По същество това е раздробяване на текста на по-малки части. Тези части

могат да са гуми, изречения или други текстови сегменти. Този процес се награжда и с два други процеса, а именно Stemming и Lemmatization, които извършват допълнителна обработка на текстовете (Jacob Murel & Eda Kavlakoglu, 2023). В случая визираще това, че токениите са гуми, но не винаги (Watkins, 2024). Като цяло ключовите гуми се подбират много внимателно и при анализа на ключови гуми може да не се прави подобна трансформация, но в случая с резюметата това е необходим процес. След въпросната обработка анализите са сходни с тези, представени до момента – библиометрика, регионални особености, n-grams, клъстеризация, прогнозиране на гуми, прогнозиране на активност по теми и пр. В допълнение могат да се включат идентифициране на частите на речта (Part-of-speech tagging – POS), оценка на цели изречения според това дали са позитивни, негативни или неутрални по даден въпрос (Average sentiment score), степен на разнообразие на използваните гуми и богатството на речника, степента на сходство (Dragomirov, 2024), както и категорийни анализи в частта “Dimension Reduction”, “Logistic Regression” и пр. (Hassani et al., 2020).

Всички посочени методи определено имат своя специфичен принос в изследването, но в случая могат да се включат и възможностите на LLM (Large Language Model) в помощ при интерпретацията на резултатите. Когато говорим за анализ на текстове, освен скоростта, тези вид системи могат да правят и допълнителни интерпретации. Това е възможно поради реди причини, но в основата е това, че LLM

могат да интерпретират контекста в текста (Devlin et al., 2018; Vaswani et al., 2017). В комбинация с това, че тези системи могат да се използват ефективно дори и под формата на Few-Shot Learners (Brown et al., 2020), ги прави незаменим помощник в интерпретацията на резултатите. В случая по никакъв начин не се ограничават значението и ролята на изследователя, а по-скоро се акцентира върху разширяването на неговия инструментариум. Разбира се, при използването на LLM модели особено важен въпрос е изследователската етика, което в последните години започва да става все по-често коментирана тема.

Последната стъпка от анализа следва да бъде представяне на данните. В случая обаче е по-добре да се дефинират възможностите за обогатяване на резултатите. По същество след реализиране на процеса ще бъдат налични множество резултати, показващи проблема от една или друга гледна точка. Въпросът е по какъв начин те да бъдат обединени в посока целта на изследването. Подобни решения са характерни за контент анализа и са представяни под една или друга форма (Krippendorff, 2013). В действителност в случая може да се използват вижданията за търсене на триангулация между изследванията за постигане на по-добри и надеждни резултати (Denzin, 2017; Patton, 1999). При този метод един от най-често срещаните подходи е свързан с интерпретация на количествени и качествени данни и съответно се търсят конвергенции, разлики и комбинации между тях (Creswell, 2009). Всичко това позволява многопластовите

резултати от приложението на множество методи да бъдат аугментирани и да се постигнат по-съществени и значими резултати. В резюме разширеният методически модел за литературен анализ обхваща следните стъпки:

- дефиниране на изследователска рамка и въпроси;
- извличане на публикации и изграждане на база данни;
- подготовка и почистване на данни;
- първична обработка и векторизация;
- тематично моделиране;
- анализ на динамиката по години и насока на прогнози;
- интерпретация чрез триангулация, включително и посредством използване на системи с изкуствен интелект.

Предложеният разширен методически модел не претендира за уникалност, а по-скоро за надграждане на утвърдените практики, като процес на систематизиране на данни и на разширяване на обхвата на информация от научни публикации на основата на прилагане на съвременни аналитични и технологични решения. Ефективността на този инструментариум е доказана чрез успешното му прилагане за анализ на големи масиви от неструктурирани данни в областта на устойчивите логистични системи (Dragomirov et al., 2026). По същество предложената методика служи като устойчива основа, позволяваща гъвкаво надграждане спрямо специфичните изследователски цели, включително чрез интеграция на системи, базирани на изкуствен интелект. Логиката на модела позволява неговото адаптиране към други научни области, изискващи системен анализ

на големи обеми литература, като осигурява надеждна рамка за надграждане на изследователското знание.

3. Изводи и възможности за развитие

Анализът на научни публикации днес е значително по-различен и се основава на развита комбинация от цифрови технологии, методи за автоматизирана обработка, алгоритми за работа с големи масиви от данни и пр. С това донякъде се дава кратък отговор на първия изследователски въпрос – *какви технологични решения се прилагат за анализ на научни публикации?* Технологичните решения включват системи за извличане на библиографска информация, инструменти за предварителна обработка на текст, методи за лингвистичен анализ и алгоритми за разпознаване на смислови структури в съдържанието. Те позволяват обработка не само на ограничен брой публикации, а на хиляди и дори десетки хиляди текстове едновременно, което дава възможност за изграждане на ново знание. В основата на тези подходи стоят автоматизирани процеси за филтриране на релевантни източници, техники за преобразуване на текст в структурирани данни, както и алгоритми за групиране, сравняване, интерпретиране и визуализиране на съдържателни зависимости между научни теми. Изключително важен въпрос е подготовката на данните в частта за премахване на доминиращи термини, както и цялостното почистване от ненужни такива.

Вторият изследователски въпрос търси *вида резултат, до който може да се достигне от подобен тип анализи.*

Подобен вид многопластови анализи предоставят богата основа за последващи научни изследвания. На първо място, те позволяват ясно очертаване на основните тематични направления и концептуални полета в дадена научна област. На второ място, чрез анализ на ключови думи, резюмета и семантични единици и пр. могат да бъдат идентифицирани тенденции, нововъзникващи теми, направления с интензивен растеж и области, в които научната активност е ограничена. Алгоритмите очертават основните тематични направления и разкриват структурата на научната дискусия, като в допълнение може да се проследи и тенденцията в развитието. Тези резултати подпомагат формулирането на нови хипотези, откриването на изследователски дефицити и насочването на бъдещи проучвания.

Третият въпрос касае *перспективите за развитие на научни изследвания* в условията на богатство от данни и информация, както и при настоящата динамика на развитие на информационните технологии. Това развитие се характеризира с нарастващи обеми данни и бързо развитие на информационните технологии. Всичко това се свързва с по-широка автоматизация на процесите по откриване, обработване и интерпретация на научна информация. Увеличаването на наличните данни изисква методи, които позволяват работа с пълнотекстови корпуси и по-дълбока смислова интерпретация. Съчетаването на количествени и качествени техники ще се превърне в основа за изграждане на по-надеждни научни модели, а усъвършенстваните

технологии за разбиране на контекста ще подпомагат формулирането на нови идеи.

Предложеният инструментариум предоставя в обобщен вид допълнителни възможности за разширяване на научното познание и за систематизиране на данни и информация от научни изследвания в контекста на устойчивите логистични системи, но той има много по-широко приложение в научните изследвания. Сред основните ограничения е използването на метаданни от индексирани системи, което може

да се награди, включително и в посока използване на пълнотекстови версии, доколкото това е възможно. Независимо от това, с настоящия материал се представя една допълнителна визия за развитие на методите в контекста на процесите на разширяване на видовете данни, които се използват в изследванията, както и начините на тяхната обработка.

Публикацията съдържа резултати от изследване, финансирано със средства от целева субсидия за НИД на УНСС по договор № НИД НИ - 16/2025/А.

Цитирани източници (References):

1. Мишев, П. & Гоев, В. (2010). Статистически анализ на времеви редове. Изд. „Авангард Прима“.
(Mishev, P. & Goev, V. (2010). *Statisticheski analiz na vremevi redove*. Izd. „Avangard Prima“)
2. American Psychological Association (2020). *Publication Manual of the American Psychological Association* (7th ed.). American Psychological Association.
3. ASReview LAB Developers (2022). *ASReview LAB: A tool for AI-assisted systematic reviews* [Software]. Zenodo. <https://doi.org/10.5281/zenodo.6625304>
4. Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D.M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners (Version 4) [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2005.14165>
5. Creswell, J.W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). SAGE.
6. Denzin, N.K. (2017). *The research act: A theoretical introduction to sociological methods* (1st ed.). Routledge. <https://doi.org/10.4324/9781315134543>
7. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding (Version 2) [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.1810.04805>
8. Dragomirov, N. (2022). Digital tools usage for literature review. In *Logistics in times of crisis: Challenges and solutions* (pp. 158–162). <https://doi.org/10.5281/zenodo.7324287>
9. Dragomirov, N. (2024). A conceptual approach to analyzing quantitative interview data for logistics research. *Innovative Information Technologies for Economy Digitalization (IITED)*, 1, 13–19.

10. Dragomirov, N., Caratù, M. & Mihova, L. (2026). Current Trends and Forecasts of Sustainable Supply Chains: Large-Scale Text Mining and Forecasting. *Sustainability*, 18(8), 3842. <https://doi.org/10.3390/su18083842>.
11. Govindan, K., Khodaverdi, R. & Jafarian, A. (2013). A fuzzy multi-criteria approach for measuring sustainability performance of a supplier based on triple bottom line approach. *Journal of Cleaner Production*, 47, 345–354. <https://doi.org/10.1016/j.jclepro.2012.04.014>
12. Hassani, H., Beneki, C., Unger, S., Mazinani, M.T. & Yeganegi, M.R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1), Article 1. <https://doi.org/10.3390/bdcc4010001>
13. IBM (2023). *What are stemming and lemmatization?* <https://www.ibm.com/topics/stemming-lemmatization>
14. Ji, X., Zhai, Y., Fu, S. & Lu, C. (2023). Towards the sustainable development of logistics system model: A system dynamics approach. *PLOS ONE*, 18(1), e0279687. <https://doi.org/10.1371/journal.pone.0279687>
15. Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). SAGE.
16. Kulshrestha, S. & Santani, D. (2024). Leveraging TF-IDF matrix for document clustering with K-means algorithm. *International Journal of Scientific Research and Modern Technology*, 3(10), 39–44. <https://doi.org/10.38124/ijrsmt.v3i10.61>
17. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
18. Mohsin, A.K.M., Schmidt, V., Aschauer, F., Plasch, M. & Gerschberger, M. (2025). Examining the evolution of sustainable supply chain management: A systematic review and bibliometric analysis. *Sustainable Development*, 33(6), 8213–8238. <https://doi.org/10.1002/sd.70093>
19. Patton, M.Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, 34(5 Pt 2), 1189–1208.
20. Project Jupyter (n.d.). *JupyterLab*. <https://jupyter.org>
21. Ren, R., Hu, W., Dong, J., Sun, B., Chen, Y. & Chen, Z. (2020). A systematic literature review of green and sustainable logistics: Bibliometric analysis, research trend and knowledge taxonomy. *International Journal of Environmental Research and Public Health*, 17(1), Article 261. <https://doi.org/10.3390/ijerph17010261>
22. Scikit-learn (n.d.). *TfidfVectorizer*. <https://scikit-learn.org>
23. Seuring, S. & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production*, 16(15), 1699–1710. <https://doi.org/10.1016/j.jclepro.2008.04.020>
24. Shou, Y., Shao, J., Lai, K., Kang, M. & Park, Y. (2019). The impact of sustainability and operations orientations on sustainable supply management and the triple bottom line. *Journal of Cleaner Production*, 240, 118280. <https://doi.org/10.1016/j.jclepro.2019.118280>

25. Song, Y. & Cao, J. (2022). An ARIMA-based study of bibliometric index prediction. *Aslib Journal of Information Management*, 74(1), 94–109. <https://doi.org/10.1108/AJIM-03-2021-0072>
26. Sun, X., Yu, H., Solvang, W.D. et al. (2022). The application of Industry 4.0 technologies in sustainable logistics: A systematic literature review (2012–2020) to explore future research opportunities. *Environmental Science and Pollution Research*, 29, 9560–9591. <https://doi.org/10.1007/s11356-021-17693-y>
27. Tabanjeh, S.T.A., Pott, C. & Reining, C. (2025). The SuLo framework: A systematic literature review of drivers and barriers to sustainable logistics practices. *Sustainability*, 17(12), Article 5575. <https://doi.org/10.3390/su17125575>
28. Touboulic, A. & Walker, H. (2015). Theories in sustainable supply chain management: A structured literature review. *International Journal of Physical Distribution & Logistics Management*, 45(1–2), 16–42. <https://doi.org/10.1108/IJPDLM-05-2013-0106>
29. Tussupov, J., Kassymova, A., Mukhanova, A., Bissengaliyeva, A., Azhibekova, Z., Yessenova, M. & Abuova, Z. (2025). Analysis of short texts using intelligent clustering methods. *Algorithms*, 18(5), Article 289. <https://doi.org/10.3390/a18050289>
30. Varriale, V., Cammarano, A., Michelino, F. & Caputo, M. (2023). Industry 5.0 and triple bottom line approach in supply chain management: The state-of-the-art. *Sustainability*, 15(7), Article 5712. <https://doi.org/10.3390/su15075712>
31. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). Attention is all you need (Version 7) [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.1706.03762>
32. Watkins, R. (2024). Guidance for researchers and peer-reviewers on the ethical use of large language models (LLMs) in scientific research workflows. *AI and Ethics*, 4(4), 969–974. <https://doi.org/10.1007/s43681-023-00294-5>

Metodicheski instrumentarium za analiz na publikatsii, svarzani s razvitiето na ustoychivite logistichni sistemi

Nikolay Dragomirov, Svetlana Aleksandrova-Zlatanska, Lilyana Mihova

Methodological Instruments for Analyzing Publications on Sustainable Logistics Systems Development

Nikolay Dragomirov, Svetlana Aleksandrova-Zlatanska, Lilyana Mihova

Abstract: The sustainable development of logistics systems has become a key research priority. This pursuit is in the context of growing economic, environmental, and social challenges, increasing regulatory pressure, and fast digital transformation. The volume of scientific publications expands, and datasets are transformed encompassing both structured and unstructured qualitative data. The need to develop the methodological toolkit for literature analysis in the field of sustainable logistics systems becomes evident. The aim of this paper is to examine the existing technological and analytical approaches for analyzing scientific publications, systematize the leading methods, and propose an extended model for literature analysis. The article summarizes a range of classical concepts and methodological approaches while augmenting them with the capabilities of modern information technologies and software. It covers solutions for data extraction from indexing databases, algorithms for filtering and automated cleaning, techniques for processing textual data, clustering methods, and the application of forecasting models to predict developments within specific research trends. Special attention is given to the role and application of artificial intelligence systems, as well as to the importance of triangulating quantitative and qualitative techniques. The resulting outcome is an expanded methodological framework for analyzing publications related to the development of sustainable logistics systems. This model can be applied not only within the field of sustainable logistics systems but also across other scientific fields. This can support researchers in identifying trends, gaps, and future research perspectives.

Key words: sustainable logistics systems, bibliometrics, text analysis.

JEL: M11, Q01.