

СУЧАСНІ ТЕНДЕЦІЇ РОЗВИТКУ ОБРОБКИ ПРИРОДНОЇ МОВИ: ЕВОЛЮЦІЯ, АРХІТЕКТУРИ ТА СФЕРИ ЗАСТОСУВАННЯ

Анотація. Актуальність теми дослідження зумовлена безпрецедентним зростанням обсягів неструктурованих текстових даних. Постановка проблеми полягає у вичерпанні можливостей традиційних комп'ютерних систем, заснованих на жорстких правилах, та необхідності переходу до складних архітектур для забезпечення семантичного розуміння тексту. Метою дослідження є проведення вичерпного аналітичного огляду методів обробки природної мови (NLP), відстеження їхньої архітектурної еволюції від статистичних моделей до сучасних парадигм глибокого навчання та великих мовних моделей (LLMs). Застосовано методи системного, порівняльного та розширеного бібліометричного аналізу. У результаті дослідження систематизовано перехід до трансформерних архітектур та великих моделей міркування (LRMs), а також проаналізовано інноваційні нейроevolюційні методи оптимізації. Ключовими висновками є підтвердження ефективності інтеграції сучасних методів NLP у сфери охорони здоров'я, інженерії програмного забезпечення та бізнес-аналітики, а також артикуляція критичних викликів: обчислювальної складності, алгоритмічної упередженості та необхідності переходу до нейросимволічного штучного інтелекту.

Ключові слова: обробка природної мови, великі мовні моделі, трансформерні архітектури, нейроevolюція, аналіз тональності, бібліометричний аналіз, штучний інтелект, машинне навчання, автоматизація досліджень.

Постановка проблеми. Сучасний етап розвитку глобальних інформаційних систем характеризується безпрецедентним і неконтрольованим зростанням обсягів неструктурованих текстових даних. В епоху тотальної цифровізації, розвитку соціальних мереж та інтелектуалізації бізнес-процесів, обробка природної мови перетворилася з вузькоспеціалізованої галузі комп'ютерної лінгвістики на один із фундаментальних, системоутворюючих напрямків штучного інтелекту. Високі вимоги до систем взаємодії між людиною та машиною, необхідність автоматизації складних аналітичних процесів та потреба в ефективному управлінні великими даними спонукають наукову спільноту до постійного вдосконалення методів NLP. По-перше, людська природна мова за своєю суттю містить величезну кількість неоднозначностей, омонімії, синонімії та характеризується глибокою, нелінійною залежністю від контексту. Відмінності між різними мовами вносять додаткову комплексну розмірність у процеси автоматизованої обробки. Традиційні комп'ютерні системи, засновані на жорстких правилах, виявилися принципово неспроможними забезпечити необхідний рівень семантичного розуміння в умовах

реальної експлуатації з її зашумленими та непередбачуваними даними. По-друге, інтеграція систем штучного інтелекту в критично важливі для суспільства сфери, такі як клінічна медицина, фінансовий аналіз, юриспруденція та управління державними ресурсами, вимагає не лише статистичної точності, але й здатності моделей до складного логічного висновку, контекстуальної адаптивності та багатомовності.

Аналіз останніх досліджень і публікацій. Згідно з макроекономічними прогнозами та статистичними звітами провідних аналітичних агентств [1], глобальний ринок технологій обробки природної мови демонструє стрімке, майже експоненційне зростання. Очікується, що сукупний середньорічний темп зростання (CAGR) складе 24,76% у період з 2026 по 2031 роки, наблизивши капіталізацію ринку до понад 201,49 мільярдів доларів США. Ці показники підкреслюють не просто величезний комерційний інтерес, але й тотальну залежність майбутньої цифрової економіки від ефективності NLP-алгоритмів. Разом з тим, стрімке масштабування параметрів сучасних нейронних мереж, від мільйонів до сотень мільярдів та навіть трильйонів, призвело до виникнення нових, раніше невідомих наукових проблем. Витрати обчислювальних потужностей на навчання та подальше використання таких систем стали надзвичайно високими, створюючи бар'єри для їх широкого впровадження та викликаючи обґрунтоване занепокоєння щодо екологічних наслідків дата-центрів. Крім того, гостро постають етичні питання: моделі неминуче успадковують соціальні, расові та гендерні упередження, присутні у навчальних даних, що призводить до проблем із справедливістю та безпекою ШІ. Виникає нагальна потреба в розробці принципово нових, більш ефективних архітектур, прозорих алгоритмів та методів оптимізації, які б дозволили подолати існуючі обмеження.

Мета дослідження. Головною метою цього наукового дослідження є проведення вичерпного, глибокого та деталізованого аналітичного огляду методів обробки природної мови, відстеження їхньої архітектурної еволюції та систематизація сфер практичного застосування на сучасному етапі технологічного розвитку (з особливим фокусом на інновації останніх років). Для досягнення цієї мети у роботі поставлено та вирішено такі завдання: систематизувати історичний перехід від класичних статистичних моделей машинного навчання до сучасних трансформерних архітектур та новітніх великих моделей міркування; проаналізувавши математичні та концептуальні причини цих зсувів; проаналізувати нетрадиційні та перспективні підходи до оптимізації великих мовних моделей, зокрема застосування нейроеволюційних алгоритмів та еволюційного пошуку архітектур (Neural Architecture Search, NAS); розробити вичерпну класифікацію ключових функціональних завдань та прикладних сфер застосування методів машинного навчання для аналізу неструктурованих текстів; провести комплексний аналіз глобальної наукової літератури для виявлення прихованих тенденцій.

Виклад основного матеріалу. Обробка природної мови (NLP) – це комплексна, глибоко міждисциплінарна галузь науки, що знаходиться на перетині інформатики (зокрема, штучного інтелекту) та структурної лінгвістики. Її головне завдання полягає у розробці обчислювальних алгоритмів та математичних моделей, здатних автоматично

сприймати, розуміти, аналізувати та генерувати людську мову у формі тексту або мовлення. В історичній ретроспективі методи та алгоритми NLP пройшли довгий та складний шлях трансформації, який концептуально поділяється на три фундаментальні парадигми.

Символьна парадигма (Rule-Based NLP)[2]: На початкових етапах розвитку штучного інтелекту, починаючи з середини двадцятого століття, домінували системи, засновані на правилах. Ці підходи вимагали колосальної ручної праці фахівців-лінгвістів, які створювали вичерпні словники, жорсткі граматичні дерева та складні набори лексичних і синтаксичних шаблонів. Хоча цей підхід забезпечував високу точність та передбачуваність у жорстко обмежених, вузькоспеціалізованих доменах, він зіткнувся з непереборною перешкодою: природна мова є занадто гнучкою, неоднозначною та динамічною. Символьні системи виявилися абсолютно немасштабованими.

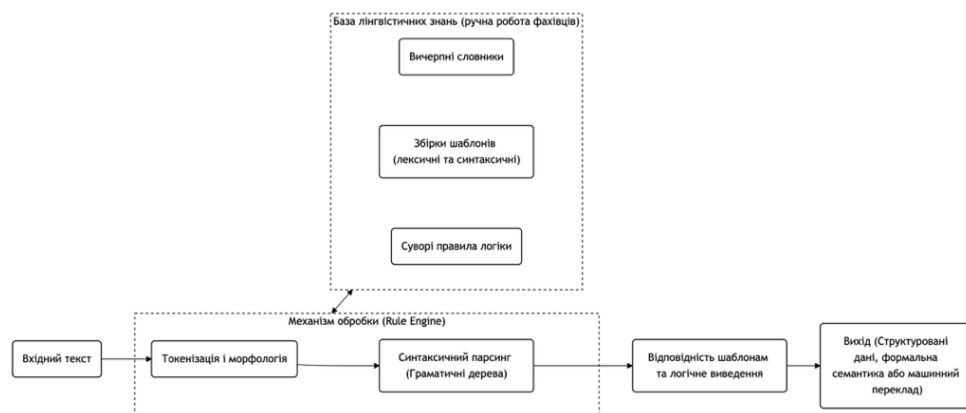


Рисунок 1 – Архітектура Rule-Based NLP

Статистична парадигма: З розвитком обчислювальних потужностей та появою великих оцифрованих текстових корпусів наприкінці 1980-х та у 1990-х роках, дослідники здійснили парадигмальний зсув у бік імовірнісних моделей та класичного машинного навчання (Machine Learning). У цей період набули поширення методи, що аналізували мову не через призму граматики, а через частотність появи тих чи інших слів та їх комбінацій (n-gram)[3]. Популярними інструментами стали наївний байєсівський класифікатор (Naive Bayes), метод опорних векторів (Support Vector Machines, SVM), алгоритми k-найближчих сусідів (k-Nearest Neighbor), дерева рішень, ансамблеві методи (Random Forest) та приховані марковські моделі (Hidden Markov Models). Для перетворення тексту в математичний вигляд масово використовувалися методи на кшталт TF-IDF (Term Frequency-Inverse Document Frequency):

$$TF - IDF(t, d, D) = TF(t, d) \cdot \log \left(\frac{|D|}{|\{d \in D : t \in d\}|} \right) \quad (1)$$

де t – термін (слово), вага якого обчислюється, d – конкретний документ із колекції тексту, D – загальна множина всіх документів (корпус), $TF(t, d)$ – частота терміну t у документі d , $|D|$ – загальна кількість документів у корпусі, $|\{d \in D : t \in d\}|$ – кількість документів, у яких зустрічається термін t .

Статистичні методи дозволили алгоритмам виявляти патерни у великих обсягах даних без необхідності програмування кожного правила, що суттєво покращило показники у завданнях класифікації текстів та базового аналізу тональності. Проте цей підхід мав фундаментальний недолік, векторизація на основі частотності ігнорувала порядок слів у реченні та була нездатна вловлювати глибокі семантичні зв'язки між словами, що знаходяться на великій відстані одне від одного.

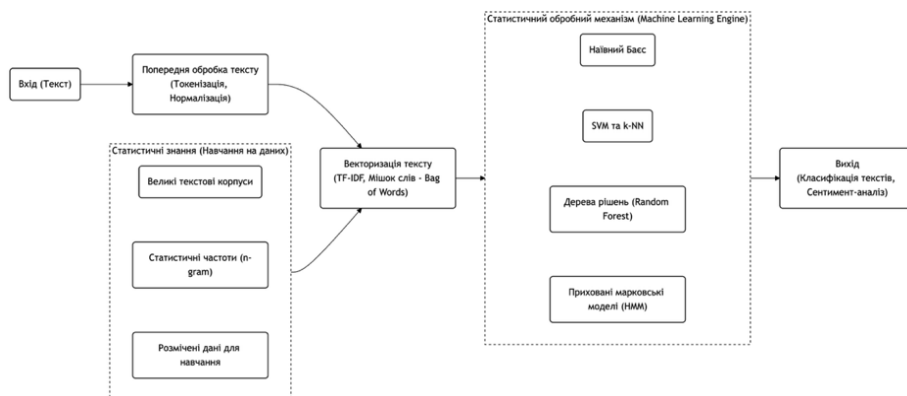


Рисунок 2 – Архітектура статистичної парадигми

Парадигма глибокого навчання (Deep Learning): Справжня наукова революція відбулася з масовим впровадженням штучних нейронних мереж (ШНМ) із багатьма прихованими шарами. Ключовим проривом стала поява методів щільного векторного представлення слів, таких як Word2Vec та GloVe [4], які дозволили відображати слова як вектори у багатовимірному семантичному просторі, де відстань між векторами відображала смислову близькість слів. Для роботи з послідовною природою тексту дослідники почали масово використовувати рекурентні нейронні мережі (Recurrent Neural Networks, RNN) та їх вдосконалену архітектуру, мережі довгої короткочасної пам'яті (Long Short-Term Memory, LSTM)[5]. Ці архітектури зберігали "пам'ять" про попередні слова при обробці наступних, що дозволило значно краще моделювати контекст. Однак, послідовний характер обчислень у RNN унеможлилював ефективне розпаралелювання процесів на сучасних графічних процесорах (GPU), що стало критичним вузьким місцем для подальшого збільшення обсягів навчальних даних та розмірів моделей.

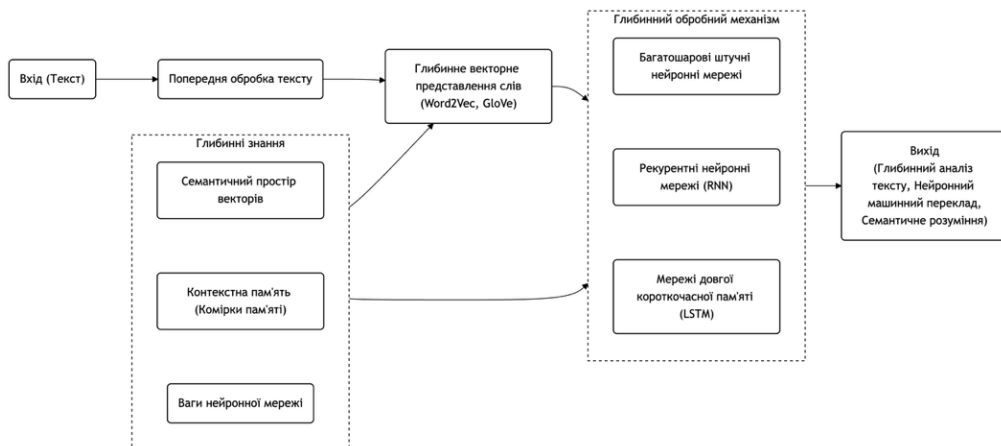


Рисунок 3 – Архітектура глибокого навчання

Фундаментальний тектонічний зсув у технологіях обробки природної мови відбувся у 2017 році після публікації знаковою праці дослідників Google під назвою "Attention is All You Need" [6]. Ця робота представила світу архітектуру Transformer, яка повністю відмовилася від рекурентних та згорткових блоків на користь інноваційного механізму так званої "внутрішньої уваги" (Self-Attention).

Архітектура трансформера дозволила обробляти всі слова у вхідній послідовності паралельно. Математично механізм самостійної уваги обчислює зв'язки між токенами за допомогою матриць запитів, ключів та значень за такою формулою:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

де Q – матриця запитів, K – матриця ключів, V – матриця значень, d_k – розмірність векторів ключів, K^T – транспонована матриця ключів.

Позиційне кодування [7] та залишкові з'єднання [8] з нормалізацією шарів вирішили проблеми збереження порядку слів та зникнення градієнта у глибоких мережах. Еволюція архітектур охопила двонаправлені (BERT [9]), генеративні (GPT-3 [10]) та сучасні моделі міркування (DeepSeek-R1 [11], Qwen [12]).

Для подолання квадратичної складності механізму уваги при обробці наддовгих контекстів сьогодні впроваджуються архітектури "суміші експертів" (MoE) [13] та системи генерації, доповненої пошуком (RAG), які оновлюють знання бази без ресурсомісткого перенавчання. Оскільки стандартні градієнтні оптимізатори (зокрема Adam) схильні до збіжності в субоптимальних локальних мінімумах і не здатні самостійно шукати топологію мережі [14], активно досліджується нейроеволюція. Цей метод оптимізує ваги та архітектури через ітеративні мутації та кросинговер. Довівши ефективність на ранніх моделях RNN/LSTM (алгоритми NEAT), сьогодні нейроеволюція фокусується на автоматизації створення трансформерів. Зокрема, еволюційний пошук архітектур (NAS) дозволяє алгоритмам самостійно конструювати оптимальні блоки трансформера [15] (порівняльний аналіз методів наведено у Таблиці 1).

Таблиця 1

Ефективність методів еволюційного пошуку архітектур

Метод / Архітектура	Особливості оптимізації моделі	Основний результат експериментів
Primer [16]	Використання квадратичної функції активації та шару глибинної згортки після кожної Q, K, V проекції.	Зниження потреби в обчислювальних ресурсах до 25-33% порівняно зі стандартними архітектурами масштабу GPT-3.
T-Razor	Використання підходу "нульових проксі" (метрика DSS++) для вибору точок алгоритмічної мутації.	Екстремальне скорочення часу пошуку ідеального трансформера з 24 діб до 0,4 доби.
SAINT+ [17]	Автоматичне впровадження операцій згортки у структуру трансформера для моделювання відстеження знань.	Досягнення ідеального балансу між моделюванням глобального (довгострокові навички) та локального (забування) контексту.

На фундаментальному рівні сучасні дослідники відзначають глибокі концептуальні паралелі між роботою великих мовних моделей та еволюційними алгоритмами: репрезентація окремого токена нагадує репрезентацію генетичного індивіда, вбудовування позиції відповідає функції формування пристосованості популяції, механізм уваги діє подібно до біологічного відбору, а циклічне навчання трансформерних блоків імітує відтворення та адаптацію параметрів у процесі боротьби за виживання [18]. Це взаємопроникнення відкриває величезні перспективи для створення стійких до збоїв штучних агентів, здатних еволюціонувати в реальному часі під час розв'язання нетипових лінгвістичних чи логічних завдань.

Алгоритми штучного інтелекту демонструють високу ефективність у подоланні інформаційного перевантаження під час наукових досліджень [19, 20, 21], а також у немедичних сферах: інженерії програмного забезпечення [22] та бізнес-аналітиці [23, 24]. Узагальнення експериментальних результатів у цих галузях наведено у Таблиці 2.

Таблиця 2

Ефективність NLP-методів у автоматизації досліджень та аналітиці

Сфера застосування	Використані моделі / Інструменти	Ключові результати та статистика експериментів
Автоматизація SLR (Огляди літератури)	TF-IDF, Sentence-BERT, Mirror-BERT	Досягнення стабільної точності класифікації на рівні 69% на відкритих масивах.
Генеративна підтримка SLR	Великі мовні моделі	Охоплення етапів створення огляду: стратегія пошуку (41%), первинний скринінг (38%), пряма екстракція (30%).
Інженерія вимог до ПЗ	Моделі глибокого навчання та LLM	Кластер "Системи NLP" формує 42,07% усього масиву світової літератури у цій галузі.
Фінансова класифікація (GICS)	Модель RoBERTa	Радикальне зниження вартості роботи; досягнуто показника якості класифікації 0,81.

Аналіз результатів експериментів базових мовних моделей. Еволюція трансформерних архітектур характеризується експоненційним зростанням обсягів параметрів та переходом від статистичної імітації мови до логічного міркування. Для об'єктивної оцінки цього прогресу ключові характеристики фундаментальних моделей зведено до строгих кількісних показників, отриманих на еталонних бенчмарках (Таблиця 3).

Як свідчать кількісні результати експериментів, сучасний етап розвитку ШІ (2025 рік) ознаменувався розв'язанням проблеми фактологічних галюцинацій. Якщо ранні моделі поклалися на статистичне передбачення наступного токена, то інноваційні архітектури (наприклад, лінійка DeepSeek), навчені виключно через масштабне навчання з підкріпленням (RL), здатні самостійно верифікувати логіку перед генерацією відпові-

ді. Застосування цієї парадигми дозволило моделі DeepSeek-R1 досягти точності 79.8% на складному математичному бенчмарку AIME 2024, зрівнявшись із найпотужнішими закритими системами. Водночас експериментальні метрики Qwen2.5-VL (96.4% точності на DocVQA) доводять успішну адаптацію когнітивного міркування до мультимодальних просторово-часових даних.

Таблиця 3

Хронологічна еволюція та кількісні показники продуктивності базових архітектур

Рік	Модель	Архітектурна інновація	К-сть параметрів
2018	BERT	Двонаправлене розуміння контексту (маскування слів).	110–340 млн
2020	GPT-3	Авторегресивна генерація, навчання без прикладів (zero-shot).	175 млрд
2025	DeepSeek-R1	Міркування через RL ("ланцюжки міркувань").	671 млрд
2025	Qwen2.5-VL	Уніфікована мультимодальність (візуально-мовний аналіз).	до 72 млрд

Аналіз та глобальні статистичні тенденції розвитку NLP. Для формування цілісного та об'єктивного розуміння вектора розвитку методів обробки природної мови (NLP) було інтегровано результати комплексного бібліометричного аналізу[25], проведеного на основі даних Scopus за період з 1992 по 2025 роки. Аналіз підтверджує, що NLP та класифікація тексту є одними з найбільш динамічних напрямів у галузі штучного інтелекту, що демонструють стійке зростання наукового інтересу та цитування.

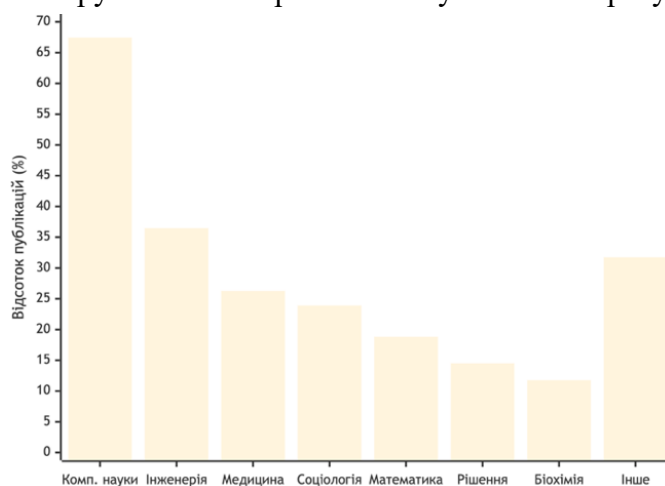


Рисунок 4 – Статистика по галузям

Згідно з отриманими даними, за останні роки спостерігається значне зростання кількості досліджень, причому пікові показники припадають на 2019, 2021 та 2022 роки. Це свідчить про перехід технологій NLP від теоретичних розробок до широкого практичного впровадження, зокрема завдяки розвитку моделей глибокого навчання та архітектури трансформерів.

З точки зору міждисциплінарності, NLP продовжує домінувати в технічних науках, проте його методи швидко інтегруються в інші галузі.

У географічному розрізі світове лідерство у виробництві наукових знань з NLP належить Сполученим Штатам, Китаю та Індії, які виступають центральними хабами глобальної мережі цитування. Також активну позицію займають Німеччина, Південна Корея та Велика Британія. Така концентрація зумовлена наявністю значних обчислювальних потужностей та доступом до капіталів великих технологічних корпорацій.

Важливим аспектом сучасного етапу розвитку є виявлення етичних та технологічних дисбалансів. Бібліометричний аналіз ключових слів вказує на те, що поряд із технічними термінами (deep learning, transformers), все частіше з'являються теми, пов'язані з упередженістю та справедливістю (bias and fairness).

Сучасні алгоритми машинного навчання вимагають колосальних обсягів даних, що створює проблему "цифрового вимирання" для мов із обмеженими ресурсами (low-resource languages). Наукова спільнота наголошує на необхідності підвищення прозорості та підзвітності ШІ-моделей, щоб забезпечити рівний доступ до технологій незалежно від мовної групи.

Висновки. Проведений огляд доводить, що методи обробки природної мови (NLP) стали найбільш динамічним сегментом штучного інтелекту. Еволюція від символічних систем до трансформерних архітектур повністю змінила ландшафт обчислювальної лінгвістики та розширила межі взаємодії людини й машини.

По-перше, спостерігається фундаментальний зсув від статистичної імітації до алгоритмічного міркування. Поява великих моделей міркування (LRMs, наприклад, DeepSeek-R1), здатних генерувати внутрішні ланцюжки роздумів через навчання з підкріпленням, свідчить про зародження прото-когнітивних систем. Їхня здатність до самоверифікації логіки потенційно вирішує хронічну проблему фактологічних галюцинацій у точних науках.

По-друге, оскільки масштаби моделей наблизилися до меж кремнієвої архітектури, класичні методи оптимізації стають недостатніми. Дослідження доводять, що майбутнє належить поєднанню нейромереж з еволюційними алгоритмами (Neuroevolution). Використання методів NAS (T-Razor) та нових математичних примітивів уваги (Primer) здатне радикально зменшити ресурсомісткість і відкрити шлях до розгортання потужного ШІ на локальних пристроях користувачів.

По-третє, аналіз прикладних сфер (зокрема медицини та юриспруденції) вказує на необхідність переходу від універсальних "мега-моделей" до систем, доповнених пошуком (RAG), та експертних архітектур. Варіативність точності базових LLM у медицині (від 3% до 90%) доводить, що статистична пам'ять мережі не здатна повністю замінити фахову доказову базу.

Подальші дослідження мають зосередитися на подоланні етичних викликів: усуненні алгоритмічних упереджень та забезпеченні цифрової рівності для малоресурсних мов. Технологічний прорив очікується у розвитку нейросимволічного ШІ, який поєднає

гнучкість розпізнавання образів з доказовою строгістю логіки, а також в адаптації NLP до квантових обчислень, що остаточно зніме бар'єри складності механізмів уваги.

ЛІТЕРАТУРА

1. ElectroIQ. Natural Language Processing Statistics By Market, Revenue And Trends (2025). 2025. URL: <https://electroi.com/stats/natural-language-processing-statistics/>
2. Strengths and Weaknesses of LLM-Based and Rule-Based NLP Technologies and Their Potential Synergies. *Electronics*. 2025. Vol. 14, № 15. P. 3064. URL: <https://www.mdpi.com/2079-9292/14/15/3064>
3. Golec J., Hachaj T. Ten Natural Language Processing Tasks with Generative Artificial Intelligence. *Applied Sciences*. 2025. Vol. 15, № 16. P. 9057. URL: <https://www.mdpi.com/2076-3417/15/16/9057>
4. Towards an evolutionary-based approach for natural language processing / L. Manzoni та ін. *arXiv preprint arXiv:2004.13832*. 2020. URL: <https://arxiv.org/abs/2004.13832>
5. Natural language processing for analyzing online customer reviews: a survey, taxonomy, and open research challenges. *PeerJ Comput. Sci.* 2024. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11323031/>
6. Attention Is All You Need / A. Vaswani та ін. *arXiv preprint arXiv:1706.03762*. 2017. URL: <https://arxiv.org/abs/1706.03762>
7. A 2D Semantic-Aware Position Encoding for Vision Transformers. *arXiv preprint arXiv:2505.09466*. 2025. URL: <https://arxiv.org/html/2505.09466v1>
8. ReViT: Enhancing Vision Transformers Feature Diversity with Attention Residual Connections / A. Diko та ін. *arXiv preprint arXiv:2402.11301*. 2024. URL: <https://arxiv.org/abs/2402.11301>
9. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin та ін. *arXiv preprint arXiv:1810.04805*. 2018. URL: <https://arxiv.org/abs/1810.04805>
10. Generative Pre-trained Transformer: A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions / G. Yenduri та ін. *arXiv preprint arXiv:2305.10435*. 2023. URL: <https://arxiv.org/abs/2305.10435>
11. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning / DeepSeek-AI та ін. *arXiv preprint arXiv:2501.12948*. 2025. URL: <https://arxiv.org/abs/2501.12948>
12. Qwen2.5-VL Technical Report / S. Bai та ін. *arXiv preprint arXiv:2502.13923*. 2025. URL: <https://arxiv.org/abs/2502.13923>
13. Speed Always Wins: A Survey on Efficient Architectures for Large Language Models / W. Sun та ін. *arXiv preprint arXiv:2508.09834*. 2025. URL: <https://arxiv.org/abs/2508.09834>
14. Enhancing Neural Network Training Through Neuroevolutionary Models: A Hybrid Approach to Classification Optimization. *Mathematics*. 2025. Vol. 13, № 7. P. 1114. URL: <https://www.mdpi.com/2227-7390/13/7/1114>
15. PubMed Article. PMID: 38502633. 2024. URL: <https://pubmed.ncbi.nlm.nih.gov/38502633/>

16. Searching for Efficient Transformers for Language Modeling. *Advances in Neural Information Processing Systems (NeurIPS)*. 2021. URL: https://papers.neurips.cc/paper_files/paper/2021/file/2f3c6a4cd8af177f6456e7e51a916ff3-Paper.pdf
17. Evolutionary Neural Architecture Search for Transformer in Knowledge Tracing. *OpenReview*. 2024. URL: <https://openreview.net/pdf?id=G14N38AjpU>
18. When Large Language Models Meet Evolutionary Algorithms: Potential Enhancements and Challenges / C. Wang та ін. *Science and Technology Review*. 2025. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11948732/>
19. Automated Research Article Classification and Recommendation Using NLP and ML / S. Rahman та ін. *arXiv preprint arXiv:2510.05495*. 2025. URL: <https://arxiv.org/html/2510.05495v1>
20. Improving Systematic Review Updates With Natural Language Processing Through Abstract Component Classification and Selection: Algorithm Development and Validation. 2025. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11986382/>
21. Large language models for conducting systematic reviews: on the rise, but not yet ready for use – a scoping review / Lieberum та ін. *medRxiv*. 2024. URL: <https://www.medrxiv.org/content/10.1101/2024.12.19.24319326v1.full.pdf>
22. A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering. *Electronics*. 2024. Vol. 13, № 11. P. 2055. URL: <https://www.mdpi.com/2079-9292/13/11/2055>
23. Comparative Analysis of NLP-Based Models for Company Classification. *Information*. 2024. Vol. 15, № 2. P. 77. URL: <https://www.mdpi.com/2078-2489/15/2/77>
24. Natural language processing in impact assessment: a review of applications and concerns / C. Hung, S. Kumar. *Impact Assessment and Project Appraisal*. 2025. URL: <https://www.tandfonline.com/doi/full/10.1080/14615517.2026.2620186>
25. Scientific Journal of Engineering Research (SJER). URL: <https://journal.futuristech.co.id/index.php/sjer/article/view/6>

REFERENCES

1. ElectroIQ. (2025). Natural Language Processing Statistics By Market, Revenue And Trends (2025). Retrieved from <https://electroi.com/stats/natural-language-processing-statistics/>
2. Strengths and Weaknesses of LLM-Based and Rule-Based NLP Technologies and Their Potential Synergies. (2025). *Electronics*, 14(15), 3064. Retrieved from <https://www.mdpi.com/2079-9292/14/15/3064>
3. Golec, J., & Hachaj, T. (2025). Ten Natural Language Processing Tasks with Generative Artificial Intelligence. *Applied Sciences*, 15(16), 9057. Retrieved from <https://www.mdpi.com/2076-3417/15/16/9057>
4. Manzoni, L., Jakobovic, D., Mariot, L., Picek, S., & Castelli, M. (2020). Towards an evolutionary-based approach for natural language processing. *arXiv preprint arXiv:2004.13832*. Retrieved from <https://arxiv.org/abs/2004.13832>

5. Natural language processing for analyzing online customer reviews: a survey, taxonomy, and open research challenges. (2024). *PeerJ Comput. Sci.* Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11323031/>
6. Vaswani, A., et al. (2017). Attention Is All You Need. *arXiv preprint arXiv:1706.03762*. Retrieved from <https://arxiv.org/abs/1706.03762>
7. A 2D Semantic-Aware Position Encoding for Vision Transformers. (2025). *arXiv preprint arXiv:2505.09466*. Retrieved from <https://arxiv.org/html/2505.09466v1>
8. Diko, A., Avola, D., Cascio, M., & Cinque, L. (2024). ReViT: Enhancing Vision Transformers Feature Diversity with Attention Residual Connections. *arXiv preprint arXiv:2402.11301*. Retrieved from <https://arxiv.org/abs/2402.11301>
9. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*. Retrieved from <https://arxiv.org/abs/1810.04805>
10. Yenduri, G., et al. (2023). Generative Pre-trained Transformer: A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions. *arXiv preprint arXiv:2305.10435*. Retrieved from <https://arxiv.org/abs/2305.10435>
11. DeepSeek-AI, et al. (2025). DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. *arXiv preprint arXiv:2501.12948*. Retrieved from <https://arxiv.org/abs/2501.12948>
12. Bai, S., et al. (2025). Qwen2.5-VL Technical Report. *arXiv preprint arXiv:2502.13923*. Retrieved from <https://arxiv.org/abs/2502.13923>
13. Sun, W., et al. (2025). Speed Always Wins: A Survey on Efficient Architectures for Large Language Models. *arXiv preprint arXiv:2508.09834*. Retrieved from <https://arxiv.org/abs/2508.09834>
14. Enhancing Neural Network Training Through Neuroevolutionary Models: A Hybrid Approach to Classification Optimization. (2025). *Mathematics*, 13(7), 1114. Retrieved from <https://www.mdpi.com/2227-7390/13/7/1114>
15. PubMed Article. (2024). PMID: 38502633. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/38502633/>
16. Searching for Efficient Transformers for Language Modeling. (2021). *Advances in Neural Information Processing Systems (NeurIPS)*. Retrieved from https://papers.neurips.cc/paper_files/paper/2021/file/2f3c6a4cd8af177f6456e7e51a916ff3-Paper.pdf
17. Evolutionary Neural Architecture Search for Transformer in Knowledge Tracing. (2024). *OpenReview*. Retrieved from <https://openreview.net/pdf?id=G14N38AjpU>
18. Wang, C., et al. (2025). When Large Language Models Meet Evolutionary Algorithms: Potential Enhancements and Challenges. *Science and Technology Review*. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11948732/>
19. Rahman, S., Shanto, H. K., Koana, U. A., & Danish, S. M. (2025). Automated Research Article Classification and Recommendation Using NLP and ML. *arXiv preprint arXiv:2510.05495*. Retrieved from <https://arxiv.org/html/2510.05495v1>

20. Improving Systematic Review Updates With Natural Language Processing Through Abstract Component Classification and Selection: Algorithm Development and Validation. (2025). Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11986382/>
21. Lieberum, et al. (2024). Large language models for conducting systematic reviews: on the rise, but not yet ready for use – a scoping review. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/10.1101/2024.12.19.24319326v1.full.pdf>
22. A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering. (2024). *Electronics*, 13(11), 2055. Retrieved from <https://www.mdpi.com/2079-9292/13/11/2055>
23. Comparative Analysis of NLP-Based Models for Company Classification. (2024). *Information*, 15(2), 77. Retrieved from <https://www.mdpi.com/2078-2489/15/2/77>
24. Hung, C., & Kumar, S. (2025). Natural language processing in impact assessment: a review of applications and concerns. *Impact Assessment and Project Appraisal*. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/14615517.2026.2620186>
25. Scientific Journal of Engineering Research (SJER). Retrieved from <https://journal.futuristech.co.id/index.php/sjer/article/view/6>

Received 16.04.2026.
Accepted 20.04.2026.
Published 30.04.2026

***Review of natural language processing methods:
evolution, architectures, and application domains***

The global Natural Language Processing (NLP) market demonstrates rapid, almost exponential growth, reflecting immense commercial interest and the digital economy's reliance on AI algorithms. However, the rapid scaling of neural network parameters—from millions to hundreds of billions—has introduced new computational and ecological challenges. Furthermore, models inevitably inherit biases present in training data, raising critical ethical and safety concerns. There is an urgent need for novel, highly efficient architectures and transparent optimization methods to overcome these limitations.

The primary goal of this research is to conduct a comprehensive analytical review of NLP methods, trace their architectural evolution from traditional statistical machine learning models to modern Large Language Models (LLMs) and Large Reasoning Models (LRMs), and systematize their practical application domains at the current stage of technological development.

The evolution of NLP is conceptually divided into three fundamental paradigms: rule-based symbolic systems, statistical methods (e.g., TF-IDF, SVM, n-grams), and deep learning. A tectonic shift occurred with the introduction of the Transformer architecture, which replaced recurrent blocks with self-attention mechanisms, enabling parallel processing. To address the quadratic computational complexity of processing long contexts, the study analyzes non-traditional optimization approaches, primarily Neuroevolution and Neural Architecture Search (NAS), which automatically construct optimal transformer blocks. The paper also systematizes the practical application of NLP in healthcare, software engineering, and business analytics, demonstrating varying degrees of accuracy and highlighting the necessity of Retrieval-Augmented Generation (RAG) technologies.

NLP has become the most dynamic segment of artificial intelligence. A fundamental shift from statistical imitation to algorithmic reasoning is observed, evidenced by LRMs like DeepSeek-R1, which are capable of self-verifying logic and mitigating factual hallucinations. As models approach the limits of silicon architecture, the future lies in combining neural networks with evolutionary algorithms. Future research must focus on overcoming ethical challenges, ensuring digital equality for low-resource languages, and advancing neurosymbolic AI and quantum computing adaptations.

Keywords: artificial intelligence, natural language processing, large language models, transformer architectures, neuroevolution, sentiment analysis, bibliometric analysis, large reasoning models, machine learning, research automation.

Ющенко Артур Сергійович – аспірант кафедри інженерії програмного забезпечення, Харківського національного університету радіоелектроніки, Україна.

ORCID: <https://orcid.org/0009-0007-7995-987X>

Смеляков Кирило Сергійович – д.т.н., професор кафедри інженерії програмного забезпечення, Харківського національного університету радіоелектроніки, Україна.

ORCID: <https://orcid.org/0000-0001-9938-5489>

Чуприна Анастасія Сергіївна – к.т.н., доцент кафедри інженерії програмного забезпечення, Харківського національного університету радіоелектроніки, Україна.

ORCID: <https://orcid.org/0000-0003-0394-9900>

Yushchenko Artur Serhiyovych – Postgraduate student at the Department of Software Engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0009-0007-7995-987X>

Smelyakov Kirill Sergeevich – Doctor of Technical Sciences, Professor at the Department of Software Engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0000-0001-9938-5489>

Chupryna Anastasia Sergeevna – Candidate of Technical Sciences, Associate Professor at the Department of Software Engineering, Kharkiv National University of Radio Electronics, Ukraine.

ORCID: <https://orcid.org/0000-0003-0394-9900>