http://doi.org/10.26565/1727-6667-2025-1-02
УДК 004.056.5:004.89:651.4/.9

Орлов Олександр Валентинович,
доктор наук з державного управління, професор, професор кафедри
публічної політики навчально-наукового інституту «Інститут державного управління»
Харківського національного університету імені В.Н. Каразіна,
майдан Свободи, 4, м. Харків, 61022, Україна
ORCID ID: https://orcid.org/0000-0001-8995-7383
e-mail: avorlovav@gmail.com

Живило Євген Олександрович,
кандидат наук з державного управління, доцент, докторант кафедри публічної політики
Навчально-наукового інституту «Інститут державного управління»
Харківського національного університету імені В.Н. Каразіна,
майдан Свободи, 4, м. Харків, 61022, Україна
ORCID ID: https://orcid.org/0000-0003-4077-7853
e-mail: zhivilka@i.ua.

Нестеренко Віктор Олександрович,
аспірант кафедри публічної політики
Навчально-наукового інституту «Інститут державного управління»
Харківського національного університету імені В.Н. Каразіна,
майдан Свободи, 4, м. Харків, 61022, Україна
ORCID ID: https://orcid.org/0000-0002-4237-3433
e-mail: svictors_1975@ukr.net

ПРИНЦИПОВІ НАПРЯМИ РОЗГОРТАННЯ ЗАХИЩЕНОГО ДОКУМЕНТООБІГУ НА ПЛАТФОРМАХ ШТУЧНОГО ІНТЕЛЕКТУ

Анотація. У статті досліджено принципові напрями розгортання захищеного документообігу на платформах штучного інтелекту з використанням локальних великих мовних моделей (LLM). Проаналізовано технологічні особливості та функціональні можливості інструментів Ollama та DeepSeek-R1 в контексті забезпечення конфіденційності обробки документів. Розглянуто переваги контейнеризації додатків штучного інтелекту на базі Docker Desktop для підвищення рівня безпеки та ізоляції компонентів системи. Запропоновано концептуальну модель багаторівневої архітектури системи захищеного документообігу, що включає інфраструктурний рівень, рівень обробки документів, рівень захисту даних та рівень моніторингу й реагування. Ідентифіковано ключові безпекові ризики, пов'язані з використанням великих мовних моделей у системах документообігу, та описано методології їх мінімізації, зокрема із застосуванням спеціалізованого інструменту тестування безпеки Garak.

На основі результатів дослідження сформульовано практичні рекомендації щодо розгортання та налаштування систем захищеного документообігу з урахуванням вимог до автономності, масштабованості та регламентованої обробки чутливої інформації. Зокрема, обґрунтовано доцільність використання локальних LLM у середовищах із суворими політиками щодо передачі даних за межі контрольованої зони організації (державні установи, медичні заклади, підприємства оборонного комплексу). Встановлено, що поєднання локальних мовних моделей із контейнеризованою інфраструктурою забезпечує оптимальний баланс між функціональністю, продуктивністю та безпекою.

Результати дослідження демонструють, що інтеграція локальних LLM з контейнеризованою інфраструктурою дозволяє забезпечити оптимальний баланс між функціональністю, продуктивністю та безпекою систем документообігу, особливо в галузях з підвищеними вимогами до конфіденційності даних.

Необхідно зазначити, що перспективою подальших досліджень є формування інтелектуальних, безпечних та гнучких систем публічного управління нового покоління, здатних ефективно функціонувати в умовах складного інформаційного середовища.

Ключові слова: *захищений документообіг, штучний інтелект, великі мовні моделі, Ollama, DeepSeek-R1, Docker Desktop, контейнеризація, інформаційна безпека, інтелектуальна обробка документів.*

Як цитувати: Орлов О. В., Живило Є. О., Нестеренко В. О. Принципові напрями розгортання захищеного документообігу на платформах штучного інтелекту. *Теорія та практика державного управління.* 2025. Вип. 1 (80). С. 24–41. http://doi.org/10.26565/1727-6667-2025-1-02

Вступ. Сучасні тенденції цифрової трансформації документообігу в організаціях різних типів зумовлюють необхідність пошуку інноваційних рішень, які б забезпечували одночасно високу ефективність обробки документів та надійний захист конфіденційної інформації. Використання технологій штучного інтелекту, зокрема великих мовних моделей (LLM), відкриває принципово нові можливості для автоматизації процесів документообігу, проте водночас створює додаткові виклики в контексті інформаційної безпеки [1].

Ключовою проблемою при розгортанні документообігу на базі штучного інтелекту залишається пошук оптимального балансу між функціональністю, продуктивністю та безпекою системи. Традиційні хмарні рішення на основі ШІ забезпечують високу обчислювальну потужність, але призводять до потенційних ризиків витоку конфіденційних даних через необхідність їх передачі на сторонні сервери [2]. Локальні ж рішення, хоча й мінімізують ризики компрометації даних, часто стикаються з обмеженнями в обчислювальних ресурсах та функціональних можливостях.

У цьому контексті інтеграція інструментів локального розгортання LLM, таких як Ollama та DeepSeek-R1, в контейнеризовану інфраструктуру на базі Docker Desktop представляється перспективним напрямом для створення захищених систем документообігу. Такий підхід потенційно дозволяє поєднати переваги хмарних та локальних рішень, забезпечуючи ефективну обробку, аналіз та зберігання документів з дотриманням високих стандартів інформаційної безпеки.

Огляд літератури. Проблематика застосування технологій штучного інтелекту в системах документообігу привертає значну увагу наукової спільноти. Аналіз публікацій останніх років демонструє зростання інтересу до використання великих мовних моделей для оптимізації процесів обробки документів.

Дослідження [3] визначають ключові тенденції в галузі інтелектуальної обробки документів (IDP) та підкреслюють трансформаційний потенціал технологій ШІ для автоматизації робочих процесів, прискорення процедур затвердження та вдосконалення механізмів виявлення шахрайства. Доведено, що впровадження IDP у фінансовому секторі дозволяє скоротити час обробки документів на 70% та підвищити ефективність виявлення «шахрайських схем» на 50%.

Так, колектив авторів DeepSeek AI [3] представив комплексне дослідження моделі DeepSeek-R1, яка демонструє значний потенціал у сферах творчого мислення, генерації програмного коду, вирішення математичних задач та автоматизованого рефакторингу програм. Особливу увагу дослідники приділили функціональності самостійного обмірковування (self-reflection), реалізованої за допомогою технології навчання з підкріпленням, що дозволяє моделі ефективно опрацьовувати складні документи.

Zhang з співавторами [5] провели порівняльний аналіз різних великих мовних моделей для вирішення комплексних завдань, включаючи обробку документів. Їх дослідження підтверджує, що DeepSeek-R1, хоча й потребує більшої кількості обчислювальних ресурсів порівняно з аналогами, забезпечує високу точність навіть у випадку складних або слабо структурованих текстів, що робить її перспективною для застосування в системах документообігу, які вимагають високої якості обробки та надійного функціонування в умовах підвищених вимог до надійності та безпеки.

Derczynski, та Galinkin [6] акцентують увагу на проблематиці безпеки великих мовних моделей і представляють фреймворк Garak для тестування вразливостей LLM. Дослідники виявили, що деякі моделі, включаючи DeepSeek-R1, потенційно можуть генерувати небезпечні відповіді за певних умов, що підкреслює необхідність впровадження комплексних стратегій мінімізації ризиків та проведення регулярного тестування безпеки. Вони також розробили методологію автоматизованого виявлення вразливостей в LLM з використанням утиліти Garak. Їхнє дослідження демонструє ефективність даного підходу для своєчасної ідентифікації та нейтралізації потенційних загроз, пов'язаних з використанням великих мовних моделей у критичних системах, включаючи документообіг.

Sahana Upadhya [6] досліджує особливості використання Docker для розгортання додатків штучного інтелекту та машинного навчання. Автори аналізують аспекти продуктивності та безпеки контейнеризованих рішень, надаючи рекомендації щодо оптимальної конфігурації для різних типів AI/ML додатків. Поряд із цим, деякі технологічні компанії [8] вже сформували комплекс перевірених практик для підвищення безпеки Docker-контейнерів. Серед них – методи верифікації образів, аудит можливостей контейнерів, застосування політик обмеження привілеїв та використання надійних сховищ. Зазначені підходи мають особливу значущість для реалізації захищених систем електронного до-

кументообігу, побудованих на базі контейнеризованої інфраструктури, де забезпечення конфіденційності, цілісності та контроль доступу до ресурсів відіграють критично важливу роль.

Методологічну основу для безпечного управління документами в регульованих галузях було представлено Harshavardhan Reddy Penubadi у співавторстві в науковій роботі [9]. Авторами детально проаналізований взаємозв'язок між технологічними рішеннями, організаційними процесами та нормативними вимогами, що є критично важливим для впровадження систем захищеного документообігу в секторах з підвищеними вимогами до конфіденційності даних.

Незважаючи на значний обсяг досліджень у галузі застосування технологій ШІ в документообігу, питання ефективної інтеграції локальних LLM з контейнеризованою інфраструктурою залишається недостатньо вивченим [8]. Зокрема, існує потреба в комплексному аналізі переваг, ризиків та перспектив такої інтеграції, а також у розробці практичних рекомендацій щодо розгортання захищених систем документообігу на базі штучного інтелекту.

При цьому необхідно зазначити, що проведений аналіз наукової літератури свідчить про наявність кількох суттєвих прогалин у дослідженнях захищеного документообігу на платформах штучного інтелекту. Так, попри активне впровадження великих мовних моделей (LLM) у процеси автоматизованої обробки текстових даних, переважна більшість наявних досліджень зосереджується на функціональних або лінгвістичних аспектах їх використання, залишаючи поза увагою питання безпеки, конфіденційності та контролю над обробкою чутливої інформації.

Зокрема, недостатньо вивченими залишаються проблеми:

– Недостатнього дослідження ефективності інтеграції локальних LLM, зокрема Ollama та DeepSeek-R1, з контейнеризованою інфраструктурою на базі Docker Desktop для забезпечення захищеного документообігу.

– Відсутності комплексного аналізу безпекових ризиків, пов'язаних з використанням великих мовних моделей у системах документообігу, та методологій їх мінімізації.

– Відсутності чітких рекомендацій щодо архітектурної організації багаторівневих систем захищеного документообігу на базі ШІ, що враховували б специфіку локально розгорнутих LLM.

– Недостатнього вивченя питання оптимального балансу між функціональністю, продуктивністю та безпекою в системах документообігу на базі локальних LLM.

– Нестачі емпіричних досліджень, присвячених оцінці ефективності застосування моделі DeepSeek-R1 для специфічних задач документообігу у порівнянні з іншими великими мовними моделями.

Таким чином, виникає потреба в системному дослідженні принципів побудови безпечних, ізольованих і контрольованих середовищ для розгортання LLM у задачах корпоративного чи державного документообігу. У цьому контексті стаття спрямована на заповнення виявлених прогалин через аналіз сучасних інструментів, проектування архітектури захищених систем та розробку рекомендацій для їх практичного впровадження.

Метою даного дослідження є комплексний аналіз принципових напрямів розгортання захищеного документообігу на платформах штучного інтелекту з використанням локальних LLM та контейнеризованої інфраструктури, а також розробка практичних рекомендацій щодо впровадження таких систем.

Для досягнення поставленої мети необхідно було вирішити наступні завдання:

1. Дослідити технологічні особливості та функціональні можливості інструментів Ollama та DeepSeek-R1 в контексті їх застосування в системах документообігу.

2. Проаналізувати переваги, обмеження та ризики контейнеризації додатків ШІ на базі Docker Desktop для забезпечення захищеного документообігу.

3. Розробити концептуальну модель багаторівневої архітектури системи захищеного документообігу на базі локальних LLM та контейнеризованої інфраструктури.

4. Ідентифікувати ключові безпекові ризики, пов'язані з використанням великих мовних моделей у системах документообігу, та запропонувати методологію їх мінімізації.

5. Сформулювати практичні рекомендації щодо розгортання та налаштування систем захищеного документообігу на базі Ollama, DeepSeek-R1 та Docker Desktop.

6. Визначити перспективні напрями подальших досліджень у галузі застосування технологій ШІ для захищеного документообігу.

Вирішення цих завдань дозволило сформувати комплексне розуміння проблематики розгортання захищеного документообігу на платформах штучного інтелекту та запропонувати ефективні підходи до впровадження таких систем у практичну діяльність організацій різного типу.

Методологія дослідження. Методологічною основою даного дослідження є системний підхід до вивчення технологічних, архітектурних та безпекових аспектів розгортання захищеного документообігу з використанням платформ штучного інтелекту, зокрема локальних великих мовних моделей та контейнеризованої інфраструктури.

Дослідження має прикладну спрямованість і передбачає комплексне поєднання теоретичного аналізу, експериментального моделювання та практичного узагальнення кращих інженерних практик у сфері штучного інтелекту та кібербезпеки.

Таким чином, саме комплексний підхід до реалізації дослідницьких завдань, включаючи аналіз інструментальних засобів, архітектурне моделювання, оцінку ризиків, тестування безпеки та формування практичних рекомендацій – забезпечив цілісність і систематизацію знань у сфері побудови захищеного документообігу на основі локальних великих мовних моделей. Це, у свою чергу, дозволило розробити практично релевантну багаторівневу архітектуру, яка враховує вимоги до безпеки, ізольованості та автономності, а також сформулювати узагальнені підходи до її впровадження в організаційних середовищах із підвищеними вимогами до конфіденційності, цілісності та контролю за обробкою даних.

Виклад основного матеріалу дослідження. З моменту появи генеративних моделей штучного інтелекту, зокрема ChatGPT, постало питання щодо забезпечення конфіденційності користувацьких даних при застосуванні таких технологій у документообігу. Хмарні сервіси генеративного ШІ викликають обґрунтовані занепокоєння щодо безпеки даних, оскільки інформація передається на сторонні сервери для обробки [8]. Альтернативним підходом стало розгортання великих мовних моделей локально, що суттєво підвищує рівень захисту даних при збереженні функціональних можливостей ШІ для роботи з документами.

Одним із перспективних інструментів для локального використання LLM є Ollama – програмне забезпечення з відкритим вихідним кодом, що дозволяє запускати великі мовні моделі безпосередньо на персональному комп'ютері користувача [11]. Основними перевагами Ollama є:

– Приватність даних – вся обробка інформації відбувається локально, без передачі на зовнішні сервери, що забезпечує високий рівень конфіденційності.

– Різноманітність підтримуваних моделей – платформа сумісна з широким спектром LLM, включаючи Llama 2, Mistral, Gemma та інші [12].

– Простота використання – процес встановлення та налаштування є інтуїтивно зрозумілим і не потребує спеціальних знань.

– Гнучкість конфігурації – наявність інструменту Modfile дозволяє тонко налаштовувати параметри моделі відповідно до специфічних завдань документообігу [13].

Можливість автономної роботи – після завантаження моделі її використання не потребує підключення до мережі Інтернет.

На сучасному етапі розвитку технологій існує кілька альтернативних інструментів для роботи з локальними LLM. Серед них: LM Studio [16], llamafile [17], Jan [18], NextChat [19], Faraday.dev [20], Backyard AI [21] та GPT4All [22]. Кожен з цих інструментів має свої особливості щодо підтримуваних моделей, зручності користувацького інтерфейсу, сумісності з різними платформами та ліцензійні умови.

Нами було зроблено порівняння характеристик та особливостей існуючого програмного забезпечення для роботи з локальними LLM, зокрема таких інструментів, як Ollama, LM Studio, llamafile, Jan, NextChat, Faraday.dev, Backyard AI та GPT4All. Результати цього порівняння представлені у табл. 1.

Проведений порівняльний аналіз характеристик та особливостей існуючого програмного забезпечення для роботи з локальними LLM вказує на наявність кількох інструментів, які можуть бути впроваджені в державне управління, а саме:

1. Ollama. Вирізняється простою інтеграцією з інструментами розробки та підтримки Docker, що робить її зручною і ефективною для впровадження в системи державного управління, де критично важливими є локальна обробка даних і забезпечення конфіденційності.

2. LM Studio та Faraday.dev пропонують зручні графічні інтерфейси, що підходять для нетехнічних користувачів, що може бути корисним для держслужбовців, які не мають глибоких технічних знань, але потребують ефективних інструментів для управління документообігом.

Таблиця 1

Порівняння характеристик та особливостей різних програм для роботи з локальними LLM

Table 1

Comparison of the characteristics and features of various programs for working with local LLMs

Характеристика / Features	**Ollama**	**LM Studio**	**llamafile**	**Jan**	**NextChat**	**Faraday.dev**	**Backyard AI**	**GPT4All**
Підтримка моделей / Model Support	Власні моделі та інтеграція з популярними LLM (LLaMA)	LLaMA, GPT-J, GPT-NeoX, Mistral та інші	LLaMA, Mistral та інші через єдиний файл	LLaMA, GPT-J, Mistral та інші	LLaMA, Mistral, GPT-J та інші	LLaMA, GPT-J, Mistral та інші	LLaMA, GPT-J, Mistral та інші	LLaMA, GPT-J, Mistral та інші
Зручність використання / Ease of Use	Простий інтерфейс, орієнтований на розробників	ГрІнт (GUI)	Використовує єдиний файл для запуску моделей	ГрІнт (GUI)	ГрІнт (GUI), орієнтований на чат-додатки	ГрІнт (GUI)	ГрІнт (GUI)	ГрІнт (GUI)
Платформи / Platforms	macOS, Linux (Windows у розробці)	Windows, macOS	Windows, macOS, Linux	Windows, macOS, Linux	Windows, macOS, Linux	Windows, macOS, Linux	Windows, macOS, Linux	Windows, macOS, Linux
Ліцензія / License	Open Source	Власницька ліцензія	Open Source	Open Source	Open Source	Open Source	Open Source	Open Source
Інтегр-я з API / API Integration	Так (REST API)	Ні	Ні	Так (REST API)	Так (REST API)	Ні	Ні	Ні
Особливості / Features	Легка інтеграція з інструментами розробки, підтримка Docker	Зручний інт. для вибору та налаштування моделей	Моделі у вигляді єдиного файлу для простоти використання	Підтримка лок-го чату, інтеграція з API	Орієнтований на чат-додатки, підтримка REST API	Зручний інт. для нетехн. корист.	Зручний інт. для нетехн. корист.	Легка інстал., підтримка лок-го використ.
Мова програмування / Programming Language	Go	C++	C++	JavaScript/ TypeScript	JavaScript/ TypeScript	JavaScript/ TypeScript	JavaScript/ TypeScript	Python

3. Llamafile пропонує унікальну можливість упаковки моделей у єдиний файл для простоти використання, що може бути корисним для швидкої інтеграції в різні державні платформи.

4. Jan та NextChat орієнтовані на створення чат-додатків з підтримкою API, що дозволяє інтегрувати ці інструменти в чат-боти та інші платформи для автоматизації взаємодії з громадянами, що може сприяти підвищенню прозорості та оперативності в управлінських процесах.

5. GPT4All є популярним вибором для локального використання моделей через свою простоту та відкриту ліцензію, що робить його доступним інструментом для використання в публічному секторі.

Особливість використання Ollama у публічному управлінні полягає в його здатності інтегрувати різноманітні інструменти, такі як bioreca̧p та rollama, які забезпечують ефективну роботу з текстовими та графічними даними через API. Це дозволяє автоматизувати процеси документообігу, а також анотування даних, що, в свою чергу, сприяє підвищенню рівня управлінської прозорості. Завдяки можливості зберігати конфіденційність даних і налаштовувати моделі для специфічних завдань, Ollama може стати потужним інструментом для оптимізації та вдосконалення управлінських процесів у державному секторі.

Ollama вирізняється серед аналогів можливістю інтеграції з різними інструментами аналізу даних, зокрема з R-пакетом rollama, що дозволяє використовувати API Ollama для анотування текстових або графічних даних, а також для вбудовування документів у власні інформаційні системи [23].

Яскравим прикладом технічної реалізації сучасних інструментів є модель DeepSeek-R1, обрана для локального використання на платформі Ollama. Ця модель виявляється особливо ефективною для вирішення складних завдань, де критично важлива висока точність, навіть за рахунок більших обчислювальних ресурсів. Впровадження DeepSeek-R1 у державні інституції може значно покращити ефективність аналізу даних, підвищити точність ухвалення рішень, а також сприяти зниженню адміністративних бар'єрів. Це, у свою чергу, дозволить підвищити ефективність і оперативність функціонування публічного управління, сприяючи більш прозорому та ефективному процесу прийняття рішень.

Таким чином, впровадження локальних LLM у документообіг забезпечує підвищений рівень захисту даних порівняно з хмарними сервісами ШІ. Інструменти на кшталт Ollama надають можливість організаціям використовувати передові технології обробки природної мови без компромісів щодо конфіденційності інформації. Водночас необхідно враховувати апаратні вимоги та особливості функціонування таких моделей для забезпечення оптимальної продуктивності системи документообігу.

Важливим елементом в забезпечені ефективного функціонування цих інструментів є використання Docker Desktop, який дозволяє швидко розгортати додатки та їх компоненти в стандартизованому середовищі. У контексті державного управління, Docker може бути використаний для безпечного розгортання додатків і модулів ШІ, забезпечуючи масштабованість і безпеку, що є критичними для обробки великих обсягів даних і збереження конфіденційності інформації.

Загалом, впровадження інтелектуальної обробки документів на основі ШІ, зокрема через Docker та Ollama, може змінити підходи до автоматизації документообігу в державному управлінні. Використання таких систем може значно зменшити час обробки документів, підвищити точність у виявленні шахрайства, а також забезпечити більшу прозорість і підзвітність в управлінських процесах. Як показує досвід банківського сектору, автоматизація документів на основі ШІ може знизити витрати, прискорити процеси та покращити моніторинг відповідності [24].

Підсумовуючи, можна констатувати, що інтеграція інструментів ШІ, таких як Ollama, DeepSeek-R1, і Docker, у публічний сектор має великий потенціал для оптимізації управлінських процесів і покращення ефективності функціонування державних інституцій.

В цілому захищений документообіг з використанням Ollama, DeepSeek-R1 та Docker Desktop передбачає створення системи, де конфіденційність, цілісність і доступність документів забезпечуються на кожному етапі їх обробки. Нижче наведено структурну схему такої системи, оформлену у вигляді ієрархічної конструкції (рис. 1.). Ця схема дозволяє візуалізувати всі ключові компоненти системи та їх взаємодію.

Представлена схема описує систему захищеного документообігу, яка використовує Ollama, DeepSeek-R1 та Docker Desktop для забезпечення безпеки та ефективності. Вона складається з чотирьох основних рівнів, кожен з яких виконує певні функції для забезпечення цілісності та конфіденційності даних.

Інфраструктурний рівень (Docker Desktop). Цей рівень формує технологічну базу всієї системи. Docker Desktop служить основою для ізоляції середовищ, де обробляються та зберігаються документи. Завдяки використанню контейнеризації, можна створювати окремі, захищені простори для кожного процесу документообігу. Це дає можливість ефективно керувати ресурсами, зменшити ризик витоків даних та забезпечити високу надійність роботи всієї системи.

– Віртуалізація через Docker – дозволяє запускати кожен сервіс як окремий контейнер, що гарантує незалежність, легкість у розгортанні й безпеку.

– Ізоляція – забезпечує чітке розмежування ресурсів і доступів між модулями.

– Контейнеризація – прискорює процес впровадження нових сервісів та мінімізує конфлікти в системі.

Інституційна стійкість. Забезпечує безперебійну роботу критичних державних сервісів завдяки гнучкості та надійності інфраструктури [25].

– Міжвідомча інтеграція. Через мікросервісну архітектуру державні відомства можуть ефективно інтегруватися, зберігаючи контроль над власними сервісами.

– Гнучке розгортання е-сервісів. У рамках державних програм цифровізації, контейнеризація дозволяє оперативно масштабувати сервіси.

Рівень обробки документів (Ollama & DeepSeek-R1). Цей рівень відповідає за інтелектуальну обробку документів, аналіз, анотації, резюмування, машинне навчання документів за допомогою моделей штучного інтелекту. Ollama відповідає за безпеку на початковому етапі обробки даних, зокрема за захист

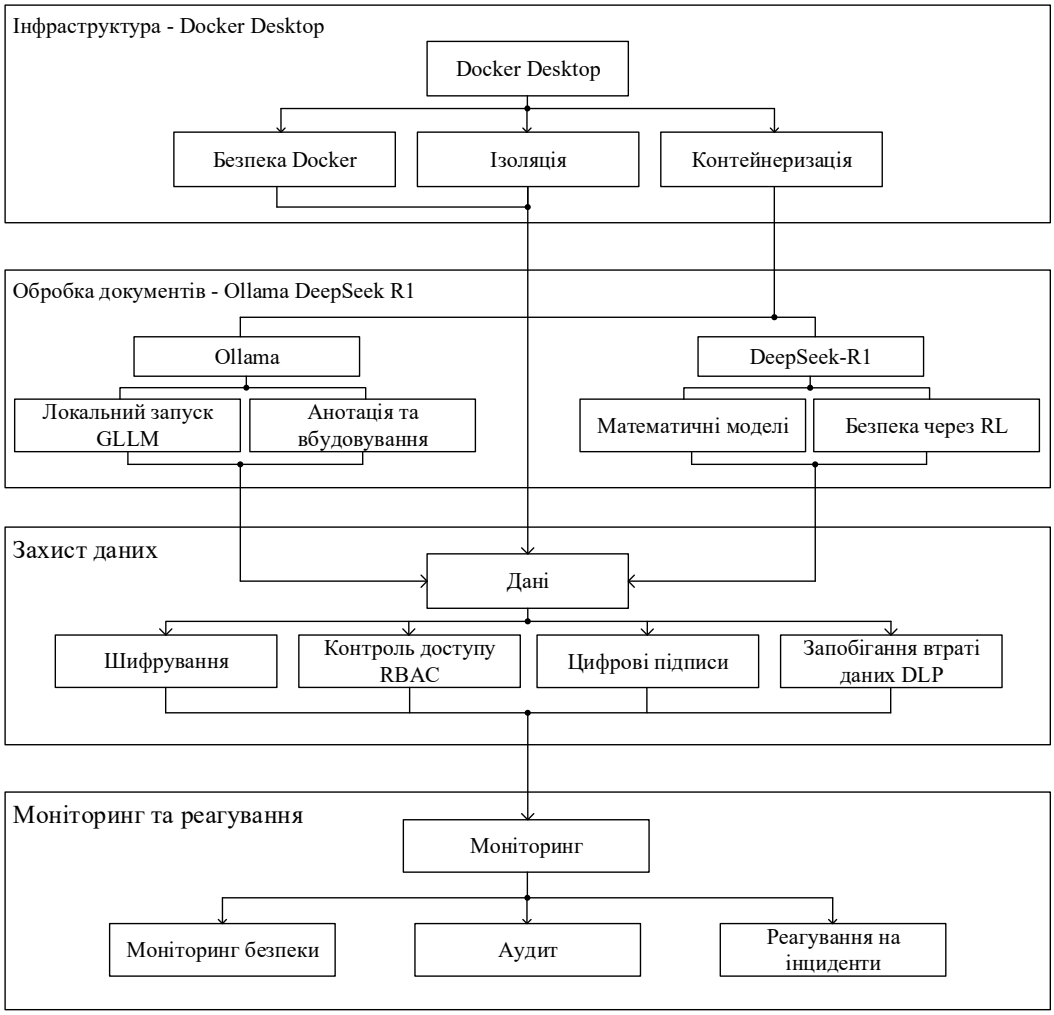

Рис. 1. Методологія документообігу в контексті механізмів державного управління
Fig. 1. Document processing methodology in the context of public administration mechanisms

документів від несанкціонованого доступу та втручання. Вона використовує передові алгоритми шифрування для забезпечення конфіденційності та аутентифікації користувачів. Цей рівень є критично важливим для публічних органів, оскільки дозволяє контролювати доступ до документів на етапі їх створення або редагування. DeepSeek-R1 виступає інструментом для аналізу та перевірки змісту документів на наявність потенційних загроз або вразливих місць. Вона здатна виявляти несанкціоновані зміни, аналізуючи структуру і зміст документів. Це важливо для публічного управління, адже дозволяє

оперативно виявляти можливі маніпуляції з документами та забезпечує їх достовірність на всіх етапах обігу.

– Локальний запуск GLLM – дає змогу обробляти документи без зовнішнього з'єднання (критично для чутливої державної інформації).

– Анотація та обговорення – автоматичне створення тез, рекомендацій, політик.

– Математичні моделі – використовуються для прогнозування наслідків рішень, обґрунтувань регуляторних актів.

– Вставка через RL (Reinforcement Learning) – навчання на практиках реального врядування (поліпшення релевантності висновків).

– Підтримка процесу прийняття рішень. LLM можуть готувати довідки, пояснювальні записки, аналіз впливу регуляцій (RIA).

– Уніфікація державних документів. ШІ-моделі стандартизують термінологію, структуру документів, що сприяє правовій визначеності.

– Аналітична підтримка. Прогнозування ефективності політик, оцінка ризиків, текстовий аналіз суспільної думки тощо.

Рівень захисту даних. На цьому рівні реалізуються механізми захисту даних, забезпечується їх конфіденційність, цілісність та доступність. Застосовуються провідні механізми інформаційної безпеки.

– Шифрування – передача та зберігання інформації відбуваються в зашифрованому вигляді.

– RBAC (Role-Based Access Control) – дозволяє гнучко налаштовувати права доступу до інформації, зокрема між різними органами влади.

– Цифрові підписи – підтвердження автентичності документів, важливо в міжвідомчій комунікації та електронному документообігу.

– DLP (Data Loss Prevention) – захист від витоку критичної інформації зсередини чи зовні.

– Законодавча відповідність. Відповідає вимогам ЗУ «Про захист інформації», КСЗІ, GDPR.

– Інформаційна безпека органів влади. Захист службових і персональних даних, державних інформаційних ресурсів та інформації з обмеженим доступом.

– Прозорість і контроль доступу. RBAC дозволяє забезпечити відповідальність осіб за свої дії з документами (audit trail).

Рівень моніторингу та реагування. Цей рівень відповідає за постійний контроль моніторинг системи та реагування на інциденти безпеки в реальному часі. 24/7 здійснюється перевірка відповідності системи політикам безпеки та виявлення порушень.

– Моніторинг безпеки – збір і аналіз журналів подій, виявлення підозрілої активності.

– Аудит – фіксація дій користувачів (важливо для службових розслідувань, перевірок НАЗК, СБ України).

– Реагування на інциденти – включає в себе автоматизовані сповіщення, блокування доступу, зберігання слідів.

– Підвищення підзвітності. Аудит дозволяє відслідковувати дії службовців і перевіряти їхню доброчесність.

Кібергігієна державного апарату. Активне реагування на інциденти (частина кіберзахисту держави) [26].

Інституційна прозорість. Моніторинг демонструє відповідність стандартам ІТ-управління, знижує ризики корупції.

Запропонована архітектура системи передбачає багаторівневу інтеграцію інструментів інфраструктури, штучного інтелекту, захисту даних та моніторингу в рамках єдиної екосистеми.

Аналізуючи загальні принципи взаємодії між рівнями запропонованої архітектури, доцільно підкреслити, що кожен з них виконує чітко визначені функціональні завдання, водночас перебуваючи у стані постійної взаємодії з іншими компонентами системи.

Така структурна організація забезпечує формування цілісного, узгодженого інформаційного та захисного потоку, що охоплює весь життєвий цикл документа – від його завантаження до безпечного зберігання або передачі визначеному суб'єкту. Це, у свою чергу, створює умови для реалізації комплексної цифрової трансформації процесів документопотоку з урахуванням актуальних ризиків, нормативних вимог у сфері інформаційної безпеки, а також впровадження механізмів інтелектуальної обробки даних.

Представлена система руху документів у межах публічного управління відповідно до адміністративно-правових регламентів, функціональних повноважень суб'єктів владних повноважень та організаційної структури органів державної влади. У запропонованій архітектурній схемі обробки документів реалізовано комплексний підхід до забезпечення уніфікованого, прозорого та контрольованого електронного документообігу, з урахуванням нормативних вимог, ризик-орієнтованого управління та інформаційної безпеки.

Умовно документопотік охоплює такі функціональні стадії:

1. Ініціація документа – створення первинного документа внаслідок управлінського запиту, подання звернення або формування службової інформації. На цьому етапі система автоматично забезпечує первинну ідентифікацію користувача, валідацію структури документа, призначення метаданих та застосування криптографічних механізмів початкового рівня (електронний підпис, хешування тощо).

2. Маршрутизація та передача на розгляд – документ передається відповідним структурним елементам органу публічної влади згідно з внутрішніми маршрутами, що визначаються функціональною компетенцією. Здійснюється контроль доступу, реєстрація усіх транзакцій у системі журналювання та забезпечується цілісність інформації під час передачі.

3. Інформації та ухвалення рішень – на основі наявних даних проводиться їх аналітичне опрацювання, за потреби, доповнення чи узгодження, після чого формується проект управлінського рішення. У цьому процесі застосовуються засоби інтелектуальної обробки даних, автоматизованого аналізу та підтримки прийняття рішень, що сприяють підвищенню обґрунтованості та об'єктивності результатів.

4. Фіналізація документа та його подальше призначення – остаточно верифікований документ або зберігається у захищеній електронній архівній системі, або передається адресату через сертифіковані захищені канали. Передача супроводжується повторною криптографічною обробкою з дотриманням вимог щодо конфіденційності, автентичності та доступності даних.

5. Аудит та контроль процесів – усі дії в рамках життєвого циклу документа підлягають обліку, що дозволяє здійснювати системний моніторинг, аудит відповідності, а також формування аналітичної звітності для потреб внутрішнього та зовнішнього контролю, у тому числі з боку органів державного фінансового чи адміністративного нагляду.

Таким чином, запропонована схема руху документів забезпечує не лише автоматизовану обробку інформації, але й створює основу для цифрової трансформації управлінських процесів, підвищення прозорості державного управління, зміцнення інформаційної безпеки та підвищення загальної ефективності публічного адміністрування.

Висновки з проведеного дослідження і перспективи подальших розвідок. У межах проведеного дослідження було здійснено комплексний аналіз можливостей використання інтелектуальних технологій у сфері організації електронного документообігу в системах державного управління. Обґрунтовано концептуальні положення щодо впровадження захищених архітектурних рішень, які поєднують можливості штучного інтелекту з сучасними вимогами інформаційної безпеки, нормативно-правової відповідності та міжвідомчої інтероперабельності.

Встановлено, що ефективне розгортання систем електронного документообігу з використанням ШІ дозволяє:

– забезпечити високий рівень автоматизації процесів створення, класифікації, маршрутизації та зберігання документів;

– істотно зменшити навантаження на людські ресурси в системах публічного адміністрування;

– підвищити достовірність прийнятих рішень через аналітичну обробку великих масивів управлінських даних;

– реалізувати гнучкі, адаптивні моделі безпеки з урахуванням змінного ризикового профілю.

Окрім того, запропонована архітектурна модель демонструє перспективність побудови багаторівневої системи обробки документів, що включає модулі машинного навчання, криптографічного захисту, цифрової автентифікації та журналювання дій користувачів. Такий підхід формує основу для нової генерації систем публічного управління – прозорих, безпечних і здатних до самонавчання.

З огляду на актуальність цифрової трансформації державного управління, подальші наукові розвідки можуть бути зосереджені на таких напрямах:

– *розробка моделей довірчих взаємодій* у системах державного електронного документообігу з використанням цифрових ідентичностей та смарт-контрактів;

– *оптимізація алгоритмів інтелектуального аналізу документів* для підвищення релевантності класифікації, виявлення аномалій та прогнозування управлінських сценаріїв;

– *поглиблене вивчення етичних і правових аспектів* застосування ШІ в автоматизованих управлінських процесах, зокрема у сфері персональних даних та прийняття рішень;

– *інтеграція із національними інформаційними системами* через єдині шлюзи обміну даними, що дозволить масштабувати рішення в межах єдиного цифрового простору держави;

– *моделювання стійкості систем електронного документообігу до гібридних загроз,* включно з інформаційними атаками, втручанням у ланцюжки прийняття рішень та маніпуляціями даними.

Таким чином, перспективи подальших досліджень відкривають широкі можливості для формування інтелектуальних, безпечних та гнучких систем публічного управління нового покоління, здатних ефективно функціонувати в умовах складного інформаційного середовища.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Feretzakis G., Papaspyridis K., Gkoulalas-Divanis A., Verykios V.S. Privacy-Preserving Techniques in Generative AI and Large Language Models: A Narrative Review. *Information*. 2024, 15(11), 697 p. DOI: https://doi.org/10.3390/info15110697

2. Das B.C., Amini M.H., Wu Y. Security and privacy challenges of large language models: A survey. *ACM Computing Surveys*, 2025, 57(6), Pp. 1-39.

3. The «Big 8» Trends in Document Management in 2025. 2025б January 27. URL: https://www.adlibsoftware.com/news/the-big-8-trends-in-document-management-in-2025

4. Guo D., Zhu Q., Yang D., Xie Z., Dong K., Zhang W., Liang W. DeepSeek-Coder: When the Large Language Model Meets Programming-The Rise of Code Intelligence. *arXiv preprint arXiv*: 2401, 14196. URL: https://arxiv.org/abs/2401.14196

5. Zhihan Zhang, Yixin Cao, Chenchen Ye, Yunshan Ma, Lizi Liao, and Tat-Seng Chua. Analyzing Temporal Complex Events with Large Language Models? A Benchmark towards Temporal, Long Context Understanding. *In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, 2024, (Volume 1: Long Papers), Pp. 1588-1606, Bangkok, Thailand. Association for Computational Linguistics.

6. Derczynski L., Galinkin E., Martin J., Majumdar S., Inie N. Garak: A framework for security probing large language models. *arXiv preprint arXiv*: 2024, 2406.11036.

7. Kolchenko V., Khoma V., Sabodashko D., Perepelytsia P. Exploring large language models' security threats with automated tools. Social Development and Security, 2024, 14(6), Pp. 81-96. DOI: https://doi.org/10.33445/sds.2024.14.6.9. URL: https://arxiv.org/abs/2406.11036

8. Sahana Upadhya et al., A State-of-Art Review of Docker Container Security Issues and Solutions. *American International Journal of Research in Science, Technology, Engineering & Mathematics*, 2017, 17(1), December 2016-February 2017, pp. 33-36. URL: https://www.researchgate.net/publication/315823494

9. Docker Security: 5 Risks and 5 Best Practices for Securing Your Containers. URL: https://www.tigera.io/learn/guides/container-security-best-practices/docker-security/

10. Penubadi H. R. Sustainable electronic document security: A comprehensive framework integrating encryption, digital signature and watermarking algorithms. *Heritage and*

Sustainable Development, 2023, vol. 5, no. 2, pp. 391-404, URL: https://hsd.ardascience.com/index.php/journal/article/view/359

11. Jakesch M. Assessing the Effects and Risks of Large Language Models in Ai-Mediated Communication : Cornell University, 2022. URL: https://www.jakesch-lab.org/assets/pdf/thesis_jakesch_cornell_phd.pdf

12. Ollama. Run Llama 2, Mistral, Gemma and other large language models locally. 2024. URL: https://ollama.com

13. Das Badhan Chandra Amini M. Hadi, Wu Yanzhao. Security and privacy challenges of large language models: A survey. *ACM Computing Surveys*, 2025, 57.6, Pp. 1-39. Ollama Models Library. URL: https://ollama.com/library

14. Ollama Documentation. (2024). Modelfile reference. URL: https://github.com/ollama/ollama/blob/main/docs/modelfile.md

15. Hoffmann J., Borgeaud S., Mensch A., et al. Training Compute-Optimal Large Language Models. *arXiv*: 2022, 2203.15556. URL: https://arxiv.org/abs/2203.15556

16. Chen M., Tworek J., Jun H., et al. Evaluating Large Language Models Trained on Code. *arXiv*, 2021, 2107.03374. URL: https://arxiv.org/abs/2107.03374

17. LM Studio. Run local large language models. 2024. URL: https://lmstudio.ai

18. Mozilla. llamafile: Distribute and run LLMs with a single file. 2024. URL: https://github.com/Mozilla-Ocho/llamafile

19 Jan AI. Run open-source LLMs locally. 2024. URL: https://jan.ai

20. NextChat. Local ChatGPT-like assistant. 2024. URL: https://github.com/next-chat/next-chat

21. Faraday.dev. Local AI Development Environment. 2024. URL: https://faraday-dev.en.softonic.com/web-apps

22. Backyard AI. Self-hosted AI assistant platform. 2024. URL: https://backyard.ai/

23. GPT4All. Run open-source large language models locally. 2024. URL: https://gpt4all.io

24. Gruber Johannes B., Maximilian Weber. Rollama: An R package for using generative large language models through Ollama. 2024. DOI: doi.org/10.21105/joss.05321. URL: https://arxiv.org/abs/2404.07654

25. Shyshatskyi A. (ed.) Information and control systems: modelling and optimizations: collective monograph. Kharkiv: Technology Center PC, 2024, 180. DOI: http://doi.org/10.15587/978-617-8360-04-7. URL: http://monograph.com.ua/pctc/catalog/book/978-617-8360-04-7

26. Живило Є.О., Орлов О.В. Сутність кібербезпеки національного сегменту кіберпростору держави в умовах кризового управління. *Публічне управління XXI століття в умовах гібридних загроз* : зб. тез XXII Міжнар. наук. конгресу, 27 квітня 2022 р. Харків : ХарРІ НАДУ «Магістр», 2022. С. 248–254.

27. Живило Є., Шевченко Д. Оцінка ризиків кібербезпеки та контролю конфіденційності в інформаційних системах державного управління. *Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка*, 2022, (75), С. 66–77. https://doi.org/10.17721/2519-481X/2022/75-07. URL: https://miljournals.knu.ua/index.php/zbirnuk/article/view/957

Конфлікт інтересів: Автори заявляють про відсутність конфлікту інтересів.

Стаття надійшла до редакції 03.03.2025
Стаття рекомендована до друку 03.04.2025

Alexander Orlov, Doctor of Sciences in Public Administration, Professor,
Professor at the Department of Public Policy of the Educational and Scientific Institute «Institute of Public Administration» of V. N. Karazin Kharkiv National University,
4 Svobody Sq., Kharkiv, 61022, Ukraine
ORCID ID: https://orcid.org/0000-0001-8995-7383 e-mail: avorlovav@gmail.com

Yevhen Zhyvylo, PhD in Public Administration, Associate Professor, Doctoral Candidate at the Department of Public Policy of the Educational and Scientific Institute
«Institute of Public Administration» of V. N. Karazin Kharkiv National University,
4 Svobody Sq., Kharkiv, 61022, Ukraine.
ORCID ID: https://orcid.org/0000-0003-4077-7853 e-mail: zhivilka@i.ua

Viktor Nesterenko, post-graduate student of the Department of Public Policy
of the Educational and Scientific Institute «Institute of Public Administration»
of V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine.
ORCID ID: https://orcid.org/0000-0002-4237-3433 e-mail: svictors_1975@ukr.net

PRINCIPAL DIRECTIONS FOR DEPLOYING SECURE DOCUMENT WORKFLOW ON ARTIFICIAL INTELLIGENCE PLATFORMS

Abstract. The article explores the principal directions for deploying secure document workflow on artificial intelligence platforms using local large language models (LLMs). It analyzes the technological features and functional capabilities of tools such as Ollama and DeepSeek-R1 in the context of ensuring confidentiality in document processing. The advantages of containerizing AI applications based on Docker Desktop are considered, particularly for enhancing system security and component isolation. A conceptual model of a multi-level architecture for a secure document workflow system is proposed, including the infrastructure layer, document processing layer, data protection layer, and monitoring and response layer. Key security risks associated with the use of large language models in document workflow systems are identified, and methodologies for their mitigation are described, notably through the application of the specialized security testing tool Garak.

Based on the research findings, practical recommendations have been formulated for the deployment and configuration of secure document workflow systems, taking into account the requirements for autonomy, scalability, and regulated handling of sensitive information. In particular, the feasibility of using local LLMs in environments with strict data transfer policies beyond the organization's controlled zone (such as government institutions, medical facilities, and defense industry enterprises) has been substantiated. It is established that combining local language models with a containerized infrastructure provides an optimal balance between functionality, performance, and security.

The research results demonstrate that integrating local LLMs with containerized infrastructure ensures an optimal balance among functionality, productivity, and security of document workflow systems, especially in sectors with high confidentiality requirements. It is worth noting that a promising direction for further research lies in the development of intelligent, secure, and flexible next-generation public administration systems capable of functioning effectively in complex information environments.

Key words: *secure document workflow, artificial intelligence, large language models, Ollama, DeepSeek-R1, Docker Desktop, containerization, information security, intelligent document processing.*

In cites: Orlov, A. V., & Zhyvylo, Ye. O., Nesterenko V. O. (2025). Principal directions for deploying secure document workflow on artificial intelligence platforms. *Theory and Practice of Public Administration*, 1 (80), 24–41. http://doi.org/10.26565/1727-6667-2025-1-02 [in Ukrainian].

REFERENCES:

1. Feretzakis, G., Papaspyridis, K., Gkoulalas-Divanis, A., & Verykios, V.S. (2024). Privacy-Preserving Techniques in Generative AI and Large Language Models: A Narrative Review. *Information*. 15(11):697. DOI: https://doi.org/10.3390/info15110697

2. Das, B. C., Amini, M. H., & Wu, Y. (2025). Security and privacy challenges of large language models: A survey. ACM Computing Surveys, 57(6), R. 1-39.

3. The «Big 8» Trends in Document Management in 2025. (2025, January 27). URL: https://www.adlibsoftware.com/news/the-big-8-trends-in-document-management-in-2025

4. Guo, D., Zhu, Q., Yang, D., Xie, Z., Dong, K., Zhang, W., ... & Liang, W. (2024). DeepSeek-Coder: When the Large Language Model Meets Programming-The Rise of Code Intelligence. arXiv preprint arXiv:2401.14196. URL: https://arxiv.org/abs/2401.14196

5. Zhihan, Zhang, Yixin, Cao, Chenchen, Ye, Yunshan, Ma, Lizi, Liao, & Tat-Seng, Chua. (2024). Analyzing Temporal Complex Events with Large Language Models? A Benchmark towards Temporal, Long Context Understanding. *In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), 1588-1606, Bangkok, Thailand. Association for Computational Linguistics.

6. Derczynski, L., Galinkin, E., Martin, J., Majumdar, S., & Inie, N. (2024). garak: A framework for security probing large language models. arXiv preprint arXiv:2406.11036.

7. Kolchenko, V., Khoma, V., Sabodashko, D., & Perepelytsia, P. (2024). Exploring large language models security threats with automated tools. *Social Development and Security*, 14(6), 81-96. DOI; https://doi.org/10.33445/sds.2024.14.6.9. URL: https://arxiv.org/abs/2406.11036

8. Sahana Upadhya et al. A State-of-Art Review of Docker Container Security Issues and Solutions. (2017). *American International Journal of Research in Science, Technology, Engineering & Mathematics*, 17(1), (December 2016-February 2017), 33-36. URL: https://www.researchgate.net/publication/315823494

9. Docker Security: 5 Risks and 5 Best Practices for Securing Your Containers URL: https://www.tigera.io/learn/guides/container-security-best-practices/docker-security/

10. Penubadi, H.R. (2023). Sustainable electronic document security: A comprehensive framework integrating encryption, digital signature and watermarking algorithms, *Heritage and Sustainable Development*, vol. 5, no. 2, pp. 391-404, URL: https://hsd.ardascience.com/index.php/journal/article/view/359

11. Jakesch, M. (2022). Assessing the Effects and Risks jf Large Language Models in Ai-Mediated Communication : Cornell University. URL: https://www.jakesch-lab.org/assets/pdf/thesis_jakesch_cornell_phd.pdf

12. Ollama. (2024). Run Llama 2, Mistral, Gemma and other large language models locally. URL: https://ollama.com

13. Das, Badhan Chandra, Amini, M. Hadi, & Wu, Yanzhao. (2025). Security and privacy challenges of large language models: A survey. *ACM Computing Surveys*, 57.6: 1-39. Ollama Models Library. URL: https://ollama.com/library

14. Ollama Documentation. (2024). Modelfile reference. URL: https://github.com/ollama/ollama/blob/main/docs/modelfile.md

15. Hoffmann, J., Borgeaud, S., Mensch, A., et al. (2022). Training Compute-Optimal Large Language Models. arXiv:2203.15556. URL: https://arxiv.org/abs/2203.15556

16. Chen, M., Tworek, J., Jun, H., et al. (2021). Evaluating Large Language Models Trained on Code. arXiv:2107.03374. URL: https://arxiv.org/abs/2107.03374

17. LM Studio. (2024). Run local large language models. URL: https://lmstudio.ai

18. Mozilla. (2024). llamafile: Distribute and run LLMs with a single file. URL: https://github.com/Mozilla-Ocho/llamafile

19. Jan AI. (2024). Run open-source LLMs locally. URL: https://jan.ai

20. NextChat. (2024). Local ChatGPT-like assistant. URL: https://github.com/next-chat/next-chat

21. Faraday.dev. (2024). Local AI Development Environment. URL: https://faraday-dev.en.softonic.com/web-apps

22. Backyard AI. (2024). Self-hosted AI assistant platform. URL: https://backyard.ai/

23. GPT4All. (2024). Run open-source large language models locally. URL: https://gpt4all.io

24. Gruber, Johannes B. & Weber, Maximilian. (2024). Rollama: An R package for using generative large language models through Ollama. DOI: doi.org/10.21105/joss.05321. URL: https://arxiv.org/abs/2404.07654

25. Shyshatskyi, A. (Ed.) (2024). Information and control systems: modelling and optimizations: collective monograph. Kharkiv: Teshnology Center PC, 180. DOI: http://doi.org/10.15587/978-617-8360-04-7. URL: http://monograph.com.ua/pctc/catalog/book/978-617-8360-04-7

26. Zhyvylo, Ye.O., & Orlov, O.V. (2022, April). The essence of cyber security of the national segment of the state's cyberspace in the context of crisis management. In: *Collection of scientific materials of the 22nd International Scientific Congress «Public administration of the 21st century in the context of hybrid threats»,* 248-254.

27. Zhyvylo, Ye., & Shevchenko, D. (2022). Cybersecurity Risk Assessment and Privacy Control in Public Administration Information Systems. *Collection of scientific works of the Military Institute of Kyiv National Taras Shevchenko University*, (75), 66–77. DOI: https://doi.org/10.17721/2519-481X/2022/75-07. URL: https://miljournals.knu.ua/index.php/zbirnuk/article/view/957

Conflict of Interest: The authors declare no conflict of interest.

The article was received by the editors 03.03.2025
The article is recommended for printing 03.04.2025