

<https://doi.org/10.31516/2410-5333.065.04><sup>1</sup>

УДК 004.03:658.15(510)

**Го Чжилян**

аспірант, Харківська державна академія культури, спеціальність «Інформаційна, бібліотечна та архівна справа», м. Харків, Україна

guzhli@qq.com

<https://orcid.org/0000-0002-0939-0530>

## **ІННОВАЦІЙНИЙ ФУНКЦІОНАЛ СИСТЕМИ НАУКОВО-ТЕХНІЧНОЇ ІНФОРМАЦІЇ КИТАЮ**

Охарактеризовано реалізовані в КНР національні та галузеві платформи агрегації великих даних, доведено необхідність інтеграції їх ресурсів і налагодження обміну результатами аналітичної обробки великих даних, що є умовою ефективної розбудови цифрової економіки країни. Вартим запозичення визнано досвід Центра документації та інформації Академії наук Китаю, який запускає серію нових інформаційних сервісів для науковців, що розроблені на основі запровадження та функціонування «Науково-технічної платформи великих даних». Зазначено, що, окрім наукових великих даних, важливою складовою інформаційної індустрії КНР є створення Національного центру промислових великих даних з філіями та субцентрами в регіонах. Обґрунтовано необхідність об'єднання ресурсів та налагодження координації діяльності агрегаторів наукових і промислових великих даних, що підвищить ефективність виробництва та впровадження інновацій, збільшить конкурентоспроможність китайської науки та бізнесу.

**Ключові слова:** *інформаційні ресурси, наукові великі дані, промислові великі дані, галузеві платформи великих даних, інноваційна інформаційна індустрія, система науково-технічної інформації, Китай, цифровізація, хмарні технології, штучний інтелект.*

**Guo Zhiliang**

Postgraduate Student, Department of Information, Library and Archival Affairs, Kharkiv State Academy of Culture, Kharkiv, Ukraine

## **INNOVATIVE FUNCTIONALITY OF THE SCIENTIFIC AND TECHNICAL INFORMATION SYSTEM OF CHINA**

**The purpose of this article** is to reveal the structure and innovative potential of the big data industry as a component of China's scientific and technical information system.

**The methodology.** The research was conducted using a complex of theoretical and empirical methods of scientific knowledge, which made it possible to identify the main components of China's scientific and technical information system, establish its innovative functionality through the identification of the capabilities of national and industry platforms of big data, their aggregators and users. The following scientific approaches were used as a cognitive toolkit for the realization of the set goal: systemic, structural-functional, comparative, as well as research methods: source studying, statistical, modeling, content analysis of the sites of the Documentation and Information Center of the Chinese Academy of Sciences.

---

1 This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

**The results.** Summarizing the achievements of the People's Republic of China in the implementation of the "National Strategy for the Formation of Big Data" made it possible to establish that the key aggregators of this information resource are the Ministry of Industry and Information Technology of the People's Republic of China, the Center for Documentation and Information of the Chinese Academy of Sciences, the National Center for Innovative Technologies of Engineering Big Data on Earth and Space, the Institute of Aerospace Information Innovations of the Chinese Academy of Sciences, the National Astronomical Observatory of China, the Bureau of Geological Survey of the Ministry of Natural Resources of China, the Institute of Remote Sensing and Digital Earth of the Chinese Academy of Sciences, the Institute of Agricultural Resources and Agricultural Zoning of the Chinese Academy of Agricultural Sciences, which form powerful industrial big data platforms. It is substantiated that the adoption of unified standards and data management protocols and their exchange, guaranteeing data security through the development of powerful technical tools for the protection of digital resources will contribute to increasing the efficiency of the interaction of big data platforms regarding their corporate use. An important organizational measure to accelerate the aggregation and integration of big data is the creation of a national industrial Internet center for big data processing, among the functions of which will be the collection of industrial data, support for their monitoring and analysis, opening of the Center's resources for small and medium-size enterprises, popularization of the possibilities of corporate application of big data in all aspects of the life cycle of production of products and services as a condition for introducing innovations and increasing business competitiveness. The most effective model of the Center's work is a multi-level structure that combines the resources of regional and industry sub-centers of industrial and scientific big data, which will ensure their integration and corporate cooperation of big data aggregators on the creation, processing and use of this powerful innovative resource.

**The scientific novelty.** For the first time, an idea of the structure, functionality and technologies of China's big data industry was introduced into the scientific circulation, the qualitative features and subjects of the aggregation of scientific and industrial big data were identified, their possibilities for the production of intelligent information products and services were determined, tools for the integration of regional and industry resources were proposed platforms of big data, the prospects of corporate processing and use of big data as a strategic resource for the development of the digital economy were outlined.

**The practical significance.** The experience of China's development of an innovative infrastructure of powerful scientific and industrial big data platforms, which are analyzed using modern cloud technologies, blockchain and artificial intelligence, and its transformation into a knowledge industry is useful for borrowing by the Ukrainian state system of scientific and technical information.

**Keywords:** *information resources, scientific big data, industrial big data, sectoral big data platforms, innovative information industry, scientific and technical information system, China, digitalization, cloud technologies, artificial intelligence.*

**Постановка проблеми.** Стрімке вдосконалення цифрових технологій обробки інформації сприяло започаткуванню якісно нового етапу розвитку системи науково-технічної інформації КНР, який ознаменувався виникненням інноваційних інфраструктурних складових, пов'язаних з накопичен-

ням та обробкою великих даних. За результатами контент-аналізу потоку наукових статей, з 2008 р. такі авторитетні журнали, як “Nature”, “Science”, “Economist” і такі організації, як “Computing Community Consortium”, ввели в науковий обіг поняття «великі дані», застосовуючи його до різних галузей науки (Гао Фен & Ван Цзянь, 2018). Під «великими даними» фахівці розуміють сукупність даних, обсяги яких швидко зростають і не можуть бути зібрані, оброблені, збережені та обчислені протягом певного періоду часу звичайними інструментами обробки даних.

Досвід промислово розвинених країн світу свідчить, що індустрія аналітичної обробки великих даних стає ключовим фактором конкурентної переваги в галузі розбудови цифрової економіки та суспільства знань. У зв'язку із цим, у 2015 р. Державна рада КНР ухвалила «Стратегію дій щодо сприяння розвитку великих даних», наголошуючи, що ці дані стають нині основним національним економічним ресурсом (国务院关于印发促进大数据发展行动纲要的通知\_信息产业). У грудні 2016 р. Міністерство промисловості та інформаційних технологій Китаю затвердило «План розвитку індустрії великих даних (2016–2020)», у якому було чітко зазначено, що до 2020 р. прибутки від продуктів і послуг, пов'язаних із великими даними, перевищать 1 трлн юанів, а середньорічні темпи зростання обсягів великих даних залишатимуться на рівні 30% (大数据产业发展规划 (2016–2020年), 2016). План передбачав прискорення будівництва техніко-технологічної та організаційної інфраструктури, яка забезпечить потужну промислову підтримку виробництва та аналізу великих даних. За останні п'ять років було створено Національну платформу великих даних та її галузеві й регіональні філії, але не досягнуто потрібних ступеня інтеграції їх ресурсів та консолідації зусиль щодо обміну результатами діяльності, що знижує ефективність застосування великих даних у всіх галузях народного господарства.

**Аналіз останніх досліджень і публікацій.** Теоретичну базу розвитку національних та галузевих платформ агрегації великих даних найактивніше розробляли фахівці Центру документації та інформації Китайської академії наук, зокрема Цянь Лі, Лю Сівень, Чжан Чжисюн та ін. Важливою є оглядова праця «Дослідження можливості підтримки інновацій науково-технічних інформаційних ресурсів на фоні великих даних» (Гао Фен & Ван Цзянь, 2018), яка розкриває методологію екстракції прихованих знань на основі багатоаспектної аналітичної обробки потужних платформ наукових даних. Стаття наукових співробітників факультету управління бібліотеками, інформацією та архівами Університету Китайської академії наук висвітлює методи екстракції нових знань на основі аналізу великих даних науково-технічної літератури (Ду Юе et al., 2022). З досягненнями щодо результатів дослідження семантичних даних наукових публікацій та можливостей їх

раціонального застосування ознайомлюють колеґ публікації Сюй Лей, Пан Джун (2018), Tian Junfeng, Wang Yanji, He Xinfeng, Zhang Juntao, Yang Wanhe, Pang Yanan (2020), Лі Цзяо (2021).

Функції та завдання регіонального рівня розбудови платформ великих даних розкрито в статтях Ду Юнле, Чжао Цзіцян, Лю Хуей, Чжу Є (2023) та Цянь Лі, Лю Сівень, Чжан Чжисюн (2021). Про необхідність системного підходу в розбудові платформ промислових та наукових великих даних як складової індустрії науково-технічної інформації КНР наголошують Гуань Цзялінь, Чжан Чао (2007), X. Xinwen, Z. Hu, I. Davydova, O. Marina, S. Marin (2023) та ін. Але узагальнюючого дослідження проблеми корпоративної співпраці агрегаторів великих даних Китаю немає, що актуалізує значущість цієї наукової праці.

**Мета статті** — розкрити структуру та інноваційний потенціал індустрії великих даних як складової системи науково-технічної інформації Китаю.

**Виклад основного матеріалу дослідження.** Уряд Китаю вважає, що регіональні та галузеві центри обробки великих даних відіграватимуть важливу роль в економічній трансформації й інноваційній модернізації країни. На виконання рішень Уряду Міністерство промисловості та інформаційних технологій КНР розробило проект будівництва промислового Інтернет-центру великих даних, підкреслюючи, що ресурси великих даних є ключовими елементами розвитку промислового Інтернету. Передбачалося, що промисловий Інтернет-центр великих даних матиме багаторівневу структуру та розподілену конструкцію по типу «національний центр + регіональні та галузеві підцентри». Національний центр промислових інтернет-ресурсів має здійснювати агрегацію, інтеграцію, аналіз та застосування великих даних, які будуть збирати регіональні й галузеві промислові Інтернет-субцентри. Наступним етапом розгортання діяльності центрів обробки великих даних є сприяння співпраці та спільному використанню корпоративних промислових ресурсів великих даних у промисловому Інтернеті, прискорення розробки їх класифікації, алгоритму підтримки повного життєвого циклу, технологій управління даними та інших стандартів для внутрішніх даних підприємств, а також впровадження експериментальної перевірки отриманих результатів. Важливим етапом розбудови промислового Інтернет-центру великих даних є заохочення ключових підприємств у регіонах отримувати доступ до його ресурсів і послуг, до корпоративної взаємодії щодо створення та використання промислового безпечного й надійного простору промислових даних (科技大数据知识发现平台).

Галузеві центри обробки великих даних теж є важливою складовою інформаційної інфраструктури КНР. Так, функціонування китайської системи безперервного спостереження планети Земля дозволяє отримувати

величезні обсяги інформації про повітря та космос. Завдяки широкому предметно-тематичному спектру й різноманітності інформації, отримуваної китайськими науковцями в результаті безперервних спостережень за планетою Земля, Світовим Океаном та Всесвітом, обсяг даних швидко зростає, набуваючи значень у сотні петабайт, які потрібно вчасно обробляти та аналізувати для підтримки прийняття управлінських рішень, здобуття нових наукових знань. Застосування технологій обробки великих даних, штучного інтелекту (далі — ШІ), глибокого навчання дозволяє інтегрувати цю інформацію в процес прийняття управлінських рішень у різних галузях народного господарства, пришвидшуючи цифровізацію промисловості, економіки та суспільства. Але для вирішення цих важливих завдань збір та аналіз величезних обсягів інформації потребує налагодження корпоративної співпраці галузевих і регіональних платформ великих даних як складових національної інфраструктури науково-технічної інформації.

У 2018 р. у КНР було прийнято Національну стратегію розвитку наукових великих даних, яка передбачала побудову сучасної цифрової інфраструктури з метою кумуляції, збереження, захисту та аналітичного опрацювання великих даних у різних галузях наукового знання. Головна мета реалізації цієї національної стратегії — застосування передових інформаційних технологій, таких як хмарне сховище, Інтернет речей, ШІ, глибокі обчислення, інтерактивна візуалізація та змішана реальність для побудови ефективної системи аналізу великих даних та перетворення отриманої інформації в знання (科学大数据——国家大数据战略的基石, 2018).

Приклади організації плідної співпраці в цьому напрямі демонструють, зокрема, Національна астрономічна обсерваторія Китаю “China Sky Eye”, що здійснює безперервні спостереження за космічними об’єктами, та Університет Гуйчжоу, які створили спільну «Лабораторію великих даних в галузі астрономії». Основними завданнями цієї лабораторії є вирішення проблем накопичення, збереження, захисту та аналітичного опрацювання великих даних, що постійно генеруються Національною астрономічною лабораторією й містять цінне «приховане знання» у сфері аерокосмічної безпеки та великомасштабної структуризації Всесвіту. Важливим напрямом діяльності новоствореної Лабораторії великих даних у галузі астрономії буде налагодження корпоративної співпраці з Державною лабораторією великих даних КНР щодо збільшення інвестицій у будівництво цифрової інфраструктури, такої як потужні сервери для зберігання астрономічних великих даних і суперкомп’ютери для їх обробки, залучення науковців для створення на основі екстракції нових знань інноваційних продуктів, що сприятимуть розвитку цифрової економіки провінції Гуйчжоу.

Однією зі складових національної інформаційної інфраструктури є Китайський центр інженерних наук і технологій (далі «Центр знань») — проект створення платформи суспільного добробуту, інтеграції відкритих ресурсів знань і сервісної платформи в галузі національної інженерії. Через акумулювання та інтеграцію великих даних у галузях, пов'язаних з інженерною наукою та технологіями, центр формує потужні інформаційні ресурси, які підтримуються сучасними цифровими технологіями обробки великих даних та експертами, що на основі екстракції знань здатні виробляти інформаційні послуги для централізованого управління економічним та технологічним розвитком країни. Центр знань має на меті надавати інформаційну підтримку й інформаційні послуги для ухвалення важливих рішень, управління ключовими інженерно-науковими і технологічними заходами, корпоративними інноваціями й навчанням персоналу в галузі національної інженерної науки та технологій, і зрештою має бути вбудованим у провідну на міжнародному та національному рівнях і широко впливову інформаційну систему, яка об'єднує центр збору інформації в галузі інженерних наук і технологій, центр інтелектуального аналізу даних і центр обслуговування знань.

Про важливість збирання та своєчасної обробки великих даних свідчить той факт, що нині всі розвинуті країни світу активно інвестують у будівництво «цифрового океану» та закладають основу для подальшої розбудови платформ «розумного океану», таких як проект «Нептун», створений США, Канадою та Японією, проект «ARANA», китайська платформа «iOcean», а також проект, започаткований 25 прибережними африканськими країнами «Платформа даних та мережевої інформації про африканські морські ресурси», тощо.

Вельми продуктивним є моніторинг океанічних великих даних з погляду підтримки ухвалення рішень для раннього попередження змін клімату, прогнозування погодних катаклізмів. Наприклад, Управління атмосфери та океанографії США (NOAA) планує запустити систему WoF (Warn — on Forecast), яка може надавати точні прогнози погоди та попередження про стихійні лиха для Сполучених Штатів і прилеглих вод: розмір обчислювальної сітки в Сполучених Штатах становить 3–10 км, а сітка в глобальній зоні — 15 км; обчислювальна спроможність системи становить 10 000 млрд разів (Лі Цзяо, 2021).

У січні 2022 р. в Циндао пройшов «Форум з інновацій та розвитку аерокосмічної інформації та цифрової економіки», ініційований Xingtu Wisdom Technology Co., Ltd. (далі «Xingtu Wisdom»). Організацію та проведення форуму спонсорували Інститут аерокосмічних інформаційних інновацій Академії наук Китаю (далі — «Аерокосмічний інститут»), Шаньдунський

інститут промислових технологій та Науково-технічне бюро Циндао, влада провінції Шаньдун та народний уряд району Лаошань. Учасники форуму підкреслювали, що галузь аерокосмічної інформації є стратегічною галуззю економіки КНР, тому компанія “Xingtu Wisdom”, інвесторами якої є Шаньдунський інститут промисловості та досліджень разом з корпорацією “Zhongke Xingtu”, має стати в майбутньому провідним у світі постачальником інформаційних продуктів і послуг для інтелектуального управління аерокосмічною галуззю (科技大数据知识发现平台).

Президент компанії “Xingtu Wisdom” Чень Вей підкреслив, що найближчим завданням його підприємства є реалізація спільного з Академією наук Китаю проєкту “Space Big Data Platform”. Кумуляція аерокосмічної інформації на єдиній платформі великих даних сприятиме покращенню державного управління країною, будівництву нових розумних міст і розвитку нової індустріалізації, інформатизації, урбанізації та модернізації сільського господарства, відродженню сільської місцевості провінції Шаньдун. На форумі відбулася презентація серії інформаційних продуктів “GEOVIS + Smart Management”, які охоплюють: GEOVIS Smart Agriculture Management Product (GEOVIS Smart Agro), GEOVIS Smart Emergency Management Product (GEOVIS Smart Emer), GEOVIS Low-altitude General Aviation Smart Management Product (GEOVIS Smart ATS), що основані на глибокій інтеграції аерокосмічних великих даних, просторово-часового блокчейну, Інтернету речей і технології ШІ, спираючись на цифрову земну просторово-часову базу. Серія продуктів “GEOVIS + Smart Management” — це потужна інформаційна база для ведення «розумного сільського господарства», «розумного управління діями в надзвичайних ситуаціях», запровадження таких галузевих додатків, як «розумний транспорт», «розумне водозбереження», «розумні ліси та поля», «розумні міста», «лінійки розумних продуктів» та ін. Послуги компанії забезпечують потужну підтримку просторової інформації та інтелектуальну підтримку ухвалення урядових рішень, розумне управління, модернізацію інформаційних сервісів.

Не менш важливою галуззю інформаційної індустрії Китаю є створення Платформи великих даних про ґрунт. 6 квітня 2020 р. Китайська академія сільськогосподарських наук оголосила про завершення науково-дослідного проєкту під назвою «Високоточний цифровий ґрунт», який реалізував Інститут сільськогосподарських ресурсів і сільськогосподарського районування. Проєкт тривав 21 рік і акумулював різноманітні дані про ґрунт, зібрані дослідниками за останні 40 років, а також представив їх на мапі ґрунтів. За результатами проєкту створено найповнішу і найточнішу базу даних наукової інформації про земельні ресурси та якість ґрунту в Китаї. Окрім цього, було впроваджено платформу «Великі дані про ґрунт», яка виконує

функції аналізу та прогнозування змін ґрунту, що має велике значення для точного зонування ґрунтових ресурсів, ефективного екологічного менеджменту та контролю за раціональним використанням земельних ресурсів.

Керівник проекту — дослідник Інституту сільськогосподарських ресурсів і сільськогосподарського районування Китайської академії сільськогосподарських наук Чжан Вейлі підкреслив, що аналітична діяльність платформи великих даних про ґрунт основана на використанні технологій «3S», а саме геоінформаційної системи (GIS), системи глобального позиціонування (GPS) та дистанційного зондування (RS), інших сучасних методах моделювання та відтворення характеристик просторового розподілу властивостей ґрунту, таких як типи ґрунту та його поживні речовини. Проект «Високоточний цифровий ґрунт» може достатньо точно відобразити ресурси орної землі Китаю та якість ґрунту, застосовуючи технології ШІ.

Цінно, що мапа «Високоточний цифровий ґрунт» надає інформацію про дев'ять шарів ґрунту, які відбивають його фізичні та хімічні властивості, ємність ґрунту та особливості навколишнього середовища, а також вміст середніх і мікроелементів. Точність прогнозів про якість ґрунту може досягати 100 м × 100 м, або 1 га. Це означає, що кожен гектар ріллі має десятки даних про ґрунт, які фермери можуть безпосередньо використовувати для управління сільськогосподарськими угіддями. Так, імпортуючи інформаційний ресурс національної платформи «Великі дані про ґрунт» у бортовий інтелектуальний пристрій, фермер точніше вносить добрива і раціональніше засіває ті чи інші сільськогосподарські землі, контролює втрати азоту та фосфору на сільськогосподарських угіддях, запобігає забрудненню й ерозії ґрунту, його рекультивації, що сприяє зменшенню кількості стихійних лих тощо.

За даними Чжан Вейлі, інформаційні продукти й послуги платформи «Великі дані про ґрунт» щорічно використовують понад 60 професійних науково-дослідних установ Китаю для проведення відповідних досліджень, а також сільськогосподарські управління 31 провінції, автономних регіонів та численних муніципалітетів у процесі прийняття рішень щодо охорони навколишнього середовища та природних ресурсів, реалізації національних проектів щодо підвищення родючості ґрунтів. Унікальність інформаційної платформи в тому, що вона дозволяє передбачити зміни в хімічному складі ґрунту, прогнози погоди, тому її дані систематично використовує Міністерство сільського господарства та інші відповідні департаменти (中国科学院知识服务平台).

Щодо акумулювання Платформи великих даних, накопичуваних у результаті спостереження супутників за планетою Земля, то з 2017 р. Китай стає провідною країною у вивченні великих даних у науках про Землю, а



щорічна кількість опублікованих із цієї проблематики наукових статей перевищує показник США. Найвпливовішими установами, що генерують найцитованіші наукові публікації про результати дослідження великих даних, отриманих у процесі космічних спостережень, є Академія наук Китаю, Уханський університет та Пекінський педагогічний університет. Базовим підрозділом Пекінського педагогічного університету, який досліджує великі дані про Землю, є Державна ключова лабораторія дистанційного зондування, створена Інститутом дистанційного зондування та цифрової Землі Китайської академії наук спільно з Пекінським педагогічним університетом. Водночас кафедра географічних наук Пекінського педагогічного університету створила Центр аналізу географічних даних і додатків, головним завданням якого є вивчення геопросторових даних для аналізу географічних процесів. Головні структурні підрозділи Центру — відділ збору та зберігання великих даних; обчислювальна лабораторія; Інститут служби географічних даних та Інститут досліджень великих географічних даних. В Уханському університеті створено Школу дистанційного зондування та інформаційної інженерії, а також Школу геодезії та картографування, які є основними дослідницькими підрозділами щодо проведення досліджень великих даних про Землю.

Основними споживачами наукової інформації, отриманої в результаті досліджень Великих даних про Землю, є геологія, географія, океанографія, метеорологія та наука про атмосферу, екологія, лісове та сільське господарство. Теми дослідження великих даних в основному зосереджені на дистанційному зондуванні, нормалізованому індексі рослинності, хмарах та південних коливаннях, випаровуванні повітря, тропічних циклонах, внутрішніх хвилях, спонтанній географічній інформації, кольорі океану, льодовиках, геостатистиці, зсувах ґрунтів, лісових пожежах та ін. Розроблення цих актуальних проблем вельми важливо для розвитку цифрової економіки та суспільства знань, своєчасного прийняття управлінських рішень у багатьох галузях народного господарства.

Одну з найвисокопродуктивніших інформаційних інфраструктур побудувала в процесі цифрової трансформації геологорозвідувальна галузь Китаю, яка об'єднала виробництвом інтелектуальних продуктів та послуг 43 підрозділи Китайської геологічної служби, що підключені до 13 провінційних, промислових і університетських вузлів. Кожний з учасників цієї мережі має право вільного доступу до різноманітних просторових баз даних геологічних карт і даних оцінки ризиків геологічних катастроф, які мають велике соціальне значення й забезпечують надання персоналізованих сервісів усім категоріям користувачів. Крім того, 26 травня 2021 р. Бюро геологічної служби Міністерства природних ресурсів Китаю оголосило про офіційний запуск Національної платформи служби обміну великими даними про

Землю “Geological Cloud 3.0”, яка надаватиме послуги для реалізації основних національних стратегій у галузі народного господарства. Консолідація на єдиній платформі системи великих даних про Землю створила умови для видобутку авторитетної наукової інформації та розроблення цілісної «Геологічної хмари», яка була оновлена від версії 1.0 до версії 3.0.

“Geology Cloud 3.0” інтегрує багатоелементну та багат шарову систему великих даних «Єдина карта наук про Землю». Відповідно до структурних складових науки про Землю, великі дані охоплюють інформацію про базову та морську геологію, енергію та мінерали, водні, земельні та лісові ресурси, ресурси пасовищ та водно-болотних угідь, підземний простір та інші 11 категорій і майже 100 ключових баз даних. Їх інформаційно-тематичний діапазон охоплює надземні та підземні ресурси, сушу та океан, точність даних стеження становить від 1:12 млн до 1:10000. Національна платформа “Geological Cloud 3.0” передбачає можливість хмарного обміну необробленими даними та великий асортимент хмарних сервісів у режимі реального часу для виконання важливих завдань динамічного моніторингу.

“Geological Cloud 3.0” — це перевірка та вдосконалення можливостей застосування сервісу великих даних у галузі геологорозвідувальних робіт. Для високоякісної розбудови економічного поясу річок Янцзи та Хуанхе, їх екологічного захисту й реалізації інших базових регіональних стратегій, будівництва великих національних проектів було створено 8 серій авторитетних геологічних інформаційних продуктів, оприлюднено 5,8 млн одиниць геологічних даних і 30000 одиниць геологічної інформації. Географічні карти, 228000 родовищ корисних копалин, 25000 продуктів з інформацією про ресурси та навколишнє середовище, 500000 м фізичних зображень важливих економічних об'єктів, 110000 книг і сотні мільйонів статей з геологічної науки, інші унікальні інформаційні продукти, призначені для своєчасного надання комплексних інформаційних послуг, реалізації основних національних стратегій, покращення рівня обслуговування першочергових соціальних потреб. Так, запровадження в “Geological Cloud 3.0” таких інтелектуальних інструментів, як технологія ідентифікації гірських мінералів дозволяє швидко розпізнати понад 70 поширених мінералів і картографічних зображень гірських порід, збагатити методи роботи та зміст мінералогії, сприяти підвищенню якості геологорозвідувальних пошуків. Таким чином, функціональний потенціал «Геологічної хмари 3.0» сприяв суттєвій модернізації традиційної моделі геологічної роботи. Запровадження в дію онлайн-системи геологічної зйомки дозволило підвищити ефективність польових геологічних досліджень, уможливило спільний обмін хмарною інформацією, знаннями та інструментами аналізу даних (Гао Фен & Ван Цзянь, 2018).

На основі подальшого поглиблення розуміння землі «Геологічна хмара 3.0» підсилює підтримку роботи Центру управління природними ресурсами КНР. Нині створено кілька важливих прикладних систем, таких як Національна інформаційна система геологічної небезпеки, Національна інформаційна система моніторингу підземних вод, Міська система геологічної інформації, розроблено інформаційну систему оцінки пропускну здатності ресурсного середовища та його придатності для розвитку земного потенціалу Китаю. Ця інформаційна система може проводити контроль у режимі реального часу, здійснювати багатоаспектний масштабний аналіз та оцінку понад 20 000 даних моніторингу підземних вод по всій країні, а також надавати інформацію про стан навколишнього середовища, екологію, катастрофи і послуги підтримки просторової інформації для міського планування, будівництва та управління.

На думку віцепрезидента Китайської академії наук Чжана Япіна, саме Китайська академія наук має взяти на себе провідну роль методичного центру та координатора активної взаємодії партнерів щодо обміну великими даними, оскільки накопичення, аналітична обробка та обмін науковими даними нині є важливим ресурсом і рушійною силою для розвитку науки і технологій у світі. У 2018 р. було досягнуто суттєвих успіхів у кількох ключових питаннях обміну даними та запровадженні його інституційних механізмів, сформувавши нову екологію обміну даними для Китайської академії наук і навіть на національному рівні. Для того щоб розбудити «сплячі» наукові дані, 1 січня 2018 р. було офіційно запущено “Earth Big Data Science Project”, стратегічний пілотний науково-технічний проект класу «А» Академії наук Китаю з терміном реалізації від 5 років. Як важливе досягнення 2018 р. було офіційно запущено сервісну платформу цього проекту — “Earth Big Data Sharing Service Platform”. Ця платформа надає глобальним користувачам систематизовані, різноманітні, динамічні, безперервні та стандартизовані дані про велику Землю з глобальними унікальними ідентифікаторами, сприяючи формуванню нової моделі обміну даними у сфері наук про Землю через створення системи обміну даними, яка інтегрує дані, обчислення і послуги.

Згідно з даними, наданими компанією “Guo Huadong”, загальний обсяг спільних даних, які нині надає платформа, становить близько 5 ПБ (1 ПБ = 1024 ТБ). Серед них 1,8 ПБ — дані спостереження Землі, 2,6 ПБ — біологічні та екологічні дані, 0,4 ПБ атмосферні і океанографічні дані, 0,2 ПБ основні географічні дані і дані наземних спостережень; 490 000 стратиграфічних і палеонтологічних баз даних, 360 китайських каталогів біологічних видів, 420000 записів у базі даних мікробних ресурсів. Важливо підкреслити, що вже нині користувачі цієї платформи можуть отримувати 40% даних в онлайн-режимі. З постійним удосконаленням засобів апаратного забез-

печення платформи великих даних зможуть запускатися одна за одною та оновлюватися з обсягом даних з ПБ щороку (中国科学院知识服务平台).

Сервісна система обміну даними та система банку даних “CASEarth”, а також регіональна система — система великих даних Землі є складовими потужного національного проекту Китаю «Цифровий Шовковий шлях». На основі спеціальних характеристик система може забезпечувати різні режими виявлення великих даних, такі як класифікація елементів, пошук ключових слів, фільтрація хмари тегів, рекомендації щодо асоціації даних тощо. У той же час вона пропонує різноманітні формати збору даних, такі як онлайн-завантаження, доступ до інтерфейсу прикладного програмування та підтримує різні версії налаштувань системи. Серед її базових сервісів: онлайн-перегляд та попередній перегляд даних, можливість отримання багатоформатних великих даних, відповіді на різні персоналізовані запити, надання статистичних даних, рекомендації щодо завантаження та оцінювання інформаційних послуг. Цінною є можливість корпоративного використання інформаційних послуг на основі дозволів на обмін даними.

Система великих даних “Digital Silk Road Earth Big Data System” містить 94 набори тематичних даних про ресурси, навколишнє середовище, клімат, катастрофи, спадщину тощо в регіоні «Один пояс, один шлях», 57 категорій продуктів даних із незалежними правами інтелектуальної власності та спільні дані, обсяг яких перевищує 120 трлн байт. Нині система вже має програмне та апаратне середовище петабайтного рівня та є першою у світі, яка розробила набір інструментів для вилучення, перетворення та завантаження великих даних Землі в межах загальної платформи великих даних, реалізуючи пошук, обмін, візуалізацію продуктів та їх поширення на міжнародному рівні через китайську, англійську та французьку версії.

Завдяки системі “CASEarth Databank” є можливість спостерігати за динамікою змін із ресурсами планети Земля протягом десятиліть, оскільки система забезпечує довгострокові набори готових до використання даних спостереження Землі з багатьох джерел, включаючи супутникові дані та дані, отримані з моменту побудови Китаєм в 1986 р. наземної станції дистанційного зондування. Водночас система застосовує передові технології, такі як ШІ, щоб забезпечити різні рівні аналізу даних спостереження Землі, а також послуги з виробництва та надання інформації для пересічних громадян, промислових користувачів і вчених, виробляє «індивідуальний» контент на замовлення підприємств і комерційних фірм.

Планується, що в майбутньому користувачі зможуть не лише використовувати свої дані, обчислення та технології для реалізації додатків, а й завантажувати дані з кількох джерел, вбудовувати моделі алгоритмів і об'єднувати системні ресурси для завершення конкретного тематичного видобутку інформації, виявлення знань і підтримки ухвалення рішень.

Важливо також, що Центр документації та інформації Академії наук Китаю запускає серію нових інформаційних сервісів для науковців, які розроблені на основі створення та функціонування «Науково-технічної платформи великих даних» (科技大数据知识发现平台). Ця платформа в змозі підтримувати якісно нову модель відкриття нових знань, що базується на обчисленні великих даних і вирішує проблему інтеграції обчислень та «вивільнення» цінних знань, що раніше ускладнювалося через розпорошеність та ізольованість великих даних у галузі науки і технологій. Нові сервіси суттєво підвищують ефективність наукових досліджень, оскільки базуються на технологіях інтелектуального аналізу знань, точного їх виявлення через автоматизовану обробку великих даних, своєчасного відстеження появи нових міждисциплінарних знань, активного їх надання через академічні обміни на основі принципів відкритості й доступності. Так, серія інтелектуальних сервісних продуктів “Smart”, що представлена “Smart Eye” і “Smart Scientific Research”, втілює нові підходи до розвитку науково-технічного забезпечення та інформаційного обслуговування членів Китайської академії наук, зокрема: точність, своєчасність, проактивність та відкритість. Завдяки таким сервісам дослідники мають можливість використовувати карти знань і технології глибокого навчання, здійснювати масштабні обчислення масивів наукових і технологічних великих даних у ширину та глибину, а також реалізують відкриту інтеграцію та спільне використання ресурсів на основі академічних обмінів і раціонального управління науковими дослідженнями.

**Висновки.** Таким чином, усі новостворювані національні та галузеві платформи великих даних є невіддільною складовою інформаційної інфраструктури Китаю, важливим завданням яких є інтеграція ресурсів і зусиль щодо генерації інновацій на якісно новому рівні цифровізації суспільства. Підвищенню ефективності взаємодії агрегаторів платформ великих даних щодо їх корпоративного використання сприятиме прийняття уніфікованих стандартів та протоколів управління даними й обміну ними, гарантування безпеки даних через розроблення ефективних технічних інструментів захисту цифрових ресурсів. Важливим організаційним заходом для прискорення агрегації та інтеграції великих даних є створення національного промислового Інтернет-центру обробки великих даних, серед функцій якого буде збір промислових даних, підтримка їх моніторингу і аналізу, відкриття ресурсів Центру для малих і середніх підприємств, популяризація можливостей корпоративного застосування великих даних у всіх аспектах життєвого циклу виробництва продуктів і послуг як умови впровадження інновацій та підвищення конкурентоспроможності бізнесу. Найефективнішою моделлю роботи Центру є багаторівнева структура, яка поєднує ресурси регіональних та галузевих субцентрів промислових і наукових великих

даних, що забезпечить їх інтеграцію та корпоративну співпрацю агрегаторів великих даних щодо створення, обробки та використання цього потужного інноваційного ресурсу.

### Список посилань

国务院关于印发促进大数据发展行动纲要的通知\_信息产业(含电信)\_中国政府网(www.gov.cn)

(*Стратегія дій щодо сприяння розвитку великих даних, ухвалена Державною радою КНР у 2015 р.*)

【大数据产业发展规划(2016-2020年)】-国家发展和改革委员会(ndrc.gov.cn)

(*План розвитку індустрії великих даних (2016–2020), затверджений Міністерством промисловості та інформаційних технологій Китаю у 2016 р.*)

科学大数据——国家大数据战略的基石(cas.cn)

(*Національна стратегія розвитку наукових великих даних Китаю, прийнята у 2018 р.*)

中国科学院知识服务平台(las.ac.cn)

(*Офіційний вебсайт Центру документації та інформації Академії наук Китаю*)  
科技大数据知识发现平台(scholareye.cn)

(*Офіційний сайт Науково-технічної платформи великих даних*)

钱力, 刘细文, 张智雄, 等. 智慧知识服务生态体系研究设计与应用实践——以中国科学院文献情报中心智慧服务平台建设为例[J]. 图书情报工作, 2021, 65(15), 78-90. (Цянь Лі, Лю Сівень, Чжан Чжисюн та ін. (2021). AI + Smart Knowledge Service Ecosystem Research Design and Application Practice — На основі рівня Smart Service Центру документації та інформації Китайської академії наук. Будівництво Тайваню як приклад [J]. *Бібліотечно-інформаційна робота*, 65(15), 78–90).

关家麟, 张超. 我国科技信息事业发展的回顾与展望[J]. 情报科学, 期 2007, 25(1), 1–6. (Гуань Цзялін, & Чжан Чао (2007). Огляд і перспективи розвитку науково-технічної інформаційної промисловості моєї країни [J]. *Intelligence Science*, 25(1), 1–6.)

高峰, 王. 大数据背景下科技信息资源创新支撑能力研究[J]. 情报杂志, 2018, 37(10), 183-188. (Гао Фен, & Ван Цзянь (2018). Дослідження можливості підтримки інновацій науково-технічних інформаційних ресурсів на фоні великих даних [J]. *Журнал Intelligence*, 37(10), 183–188).

杜永乐, 赵继强, 刘辉, 朱烨. 省市县不动产登记数据一致性方法与信息平台建设研究——以江苏省为例[J]. 自然资源信息化. 2023(05). (Ду Юнле, Чжао Цзіцян, Лю Хуей, & Чжу С. (2023). Дослідження методу узгодженості та побудови інформаційної платформи провінційних, міських і повітових реєстраційних даних про нерухоме майно — на прикладі провінції Цзянсу [J]. *Інформатизація природних ресурсів*, (05)).

引用格式: 杜悦, 常志军, 董美, 钱力, 王颖. 一种面向海量科技文献数据的大规模知识图谱构建方法[J/OL]. 数据分析与知识发现. 2022

- (07). (Ду Юе, Чан Чжіцзюнь, Дун Мей, Цянь Лі, & Ван Ін (2022). Широкомасштабні знання, орієнтовані на великі дані наукової та технічної літератури. Спосіб побудови карти. *Аналіз даних і відкриття знань*, (07)).
- 田俊峰,王彦彘,何欣枫,张俊涛,杨万贺,庞亚南. 数据因果一致性研究综述[J]. 通信学报. 2020(03). (Tian Junfeng, Wang Yanji, He Xinfeng, Zhang Juntao, Yang Wanhe, & Pang Yanan. (2020). Огляд дослідження причинно-наслідкової узгодженості даних [J]. *Journal of Communications*, (03)).
- 徐雷,潘璐. 科学出版物语义数据及其应用研究[J] 中国科技期刊研究. 2018(07). (Сюй Лей, & Пан Джун (2018). Дослідження семантичних даних наукових публікацій та їх застосування [J]. *Дослідження китайських наукових і технологічних журналів*, (07)).
- 李娇. 基于知识图谱的科研综述生成研究[D] 中国农业科学院 2021. (Лі Цзяо (2021). *Дослідження щодо створення оглядів наукових досліджень на основі графіка знань [D]*. Китайська академія сільськогосподарських наук).
- Xinwen X., Hu Z., Davydova, I., Marina, O., & Marin, S. (2023). *Analysis of Project Activities of Libraries of the People's Republic of China. In Proceedings of the 2nd International Workshop on Social Communication and Information Activity in Digital Humanities (SCIA-2023). Lviv, Ukraine, November 9, 2023*, Pp. 184–195. <https://ceur-ws.org/Vol-3608/paper14.pdf>.
- Peng, Y., Davydova, I., Kunanets, N., & Veretennikova, N. (2021). University Library Projects for the Formation of China's Digital Library and Information Space. In *Computer Science and Information Technologies 22–25 September, 2021. Lviv, 2021*, P. 399–402. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85124327189&origin=resultslist&sort=plf-f&src=s&nlo=&nlr=&nls=&sid=98ae310989c6ffb0c9e6ffb9706f959b&sot=aff&sdt=cl&cluster=scoprefnameuid%2c%22Davydova%2c+I.O.%2357216904009%22%2ct%2c%22Davydova%2c+I.O.%2357216904009%22%2ct&sl=50&s=AF-ID%28%22Kharkiv+State+Academy+of+Culture%22+60275400%29&relpos=1&citeCnt=0&searchTerm>

## References

- Action Strategy for Promoting Big Data Development, adopted by the State Council of the People's Republic of China in 2015.* www.gov.cn. [In Chinese].
- Big Data Industry Development Plan (2016–2020), approved by the Ministry of Industry and Information Technology of China in 2016.* ndrc.gov.cn. [In Chinese].
- China's National Strategy for the Development of Scientific Big Data, adopted in 2018.* cas.cn. [In Chinese].
- An official website of the Documentation and Information Center of the Chinese Academy of Sciences.* las.ac.cn. [In Chinese].
- Official website of the Big Data Science and Technology Platform.* scholareye.cn. [In Chinese].

- Qian Li, Liu Xiwen, Zhang Zhixiong, et al. (2021). AI+Smart Knowledge Service Ecosystem Research Design and Application Practice-Based on the Smart Service level of the Documentation and Information Center of the Chinese Academy of Sciences. Taiwan's construction as an example [J]. *Library and Information Science*, 65 (15), 78–90. [In Chinese].
- Guan Jialin, & Zhang Chao (2007). Review and development prospects of my country's scientific and technological information industry [J]. *Intelligence Science*, 25 (1), 1–6. [In Chinese].
- Gao Feng, & Wang Jian (2018). Research on the possibility of supporting the innovation of scientific and technical information resources against the background of big data [J]. *Journal of Intelligence*, 37 (10), 183–188. [In Chinese].
- Du Yongle, Zhao Jiqiang, Liu Hui, & Zhu Ye (2023). Research on the method of consistency and construction of the information platform of provincial, city and county real estate registration data — a case study of Jiangsu Province [J]. *Informatization of natural resources*, (05). [In Chinese].
- Du Yue, Chang Zhijun, Dong Mei, Qian Li, & Wang Ying (2022). Large-scale knowledge discovery based on big data from scientific and technical literature. A mapping method. *Data analysis and knowledge discovery*, (07). [In Chinese].
- Tian Junfeng, Wang Yanji, He Xinfeng, Zhang Juntao, Yang Wanhe, & Pang Yanan. (2020). A review of the research on data causal consistency [J]. *Journal of Communications*, (03). [In Chinese].
- Xu Lei, & Pan Jun (2018). Research on semantic data of scientific publications and its application [J]. *Research on Chinese science and technology journals*, (07). [In Chinese].
- Li Jiao (2021). *Research on the establishment of scientific research reviews based on knowledge graph* [D]. Chinese Academy of Agricultural Sciences. [In Chinese].
- Xinwen X., Hu Z., Davydova, I., Marina, O., & Marin, S. (2023). *Analysis of Project Activities of Libraries of the People's Republic of China. In Proceedings of the 2nd International Workshop on Social Communication and Information Activity in Digital Humanities (SCIA-2023). Lviv, Ukraine, November 9, 2023*, Pp. 184–195. <https://ceur-ws.org/Vol-3608/paper14.pdf>. [In English].
- Peng, Y., Davydova, I., Kunanets, N., & Veretennikova, N. (2021). University Library Projects for the Formation of China's Digital Library and Information Space. In *Computer Science and Information Technologies 22–25 September, 2021. Lviv, 2021*, P. 399–402. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85124327189&origin=resultslist&sort=plf-f&src=s&nlo=&nlr=&nls=&sid=98ae310989c6fbb0c9e6ffb9706f959b&sot=aff&sdt=cl&cluster=scoprefnameauid%2c%22Davydova%2c+I.O.%2357216904009%22%2ct%2c%22Davydova%2c+I.O.%2357216904009%22%2ct&sl=50&s=AF-ID%28%22Kharkiv+State+Academy+of+Culture%22+60275400%29&relpos=1&citeCnt=0&searchTerm>. [In English].