УДК: 004.4:004.8:004.72 https://doi.org/10.15407/jai2026.01.117

А. М. Горачук
Державна наукова установа «Центр інноваційних технологій охорони здоров'я»
Державного управління справами, Україна
вул. Верхня, 5, Київ, 01014
a.gorachuk@gmail.com
https://orcid.org/0009-0000-6067-9025

ТЕХНІЧНИЙ ПІДХІД ДО КОНВЕРТАЦІЇ АРХІВІВ МЕДИЧНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ (МІС) «ЕСКУЛАП» ДЛЯ ЗАДАЧ ШТУЧНОГО ІНТЕЛЕКТУ: ДОСВІД ДНУ «ЦІТОЗ» ДУС

A. Horachuk
State Institution of Science "Center of Innovative Healthcare Technologies"
State Administrative Department, Ukraine
Verkhnia Street, 5, Kyiv, 01014
a.gorachuk@gmail.com
https://orcid.org/0009-0000-6067-9025

TECHNICAL APPROACH TO CONVERTING MEDICAL INFORMATION SYSTEM (MIS) "ESCULAP" ARCHIVES FOR ARTIFICIAL INTELLIGENCE TASKS: EXPERIENCE OF STATE INSTITUTION OF SCIENCE "CENTER OF INNOVATIVE HEALTHCARE TECHNOLOGIES" STATE ADMINISTRATIVE DEPARTMENT

Анотація. Цифровізація медицини потребує перетворення накопичених даних у формати, придатні для аналізу. Значна частина медичної інформації в Україні зберігається в архівах застарілих систем, зокрема МІС «Ескулап». Ці дані використовують структуру dBase (файли .DBF та .FPT), що унеможливлює їх пряме застосування у сучасних завданнях штучного інтелекту та машинного навчання.

У статті описано технічний підхід до конвертації таких архівів. Розроблено алгоритм на мові Python із використанням бібліотек dbfread, pandas та numpy. Методика забезпечує екстракцію та систематизацію деперсоналізованих даних пацієнтів, діагнозів та історій лікування. Окрему увагу приділено вирішенню проблем із кодуванням тексту та обробці полів memo у файлах .FPT.

Результатом роботи є перетворення реляційних таблиць застарілого формату у файли CSV. Це дозволяє працювати з історичними медичними даними за допомогою сучасних аналітичних методів. Отримані масиви даних стають придатними для машинного навчання, створення нейронних мереж і проведення статистичних досліджень у сфері громадського здоров'я.

Запропонований підхід перевірено на базі архівів ДНУ «ЦІТОЗ» ДУС. Розроблений алгоритм забезпечує коректну підготовку даних для подальшого аналізу захворюваності за тривалі періоди. Використання розробленого програмного інструменту дозволяє щорічно оновлювати бази даних для досліджень, забезпечуючи перетворення нових архівних записів у зручний для обробки формат.

Практична реалізація алгоритму передбачала інвентаризацію 195 архівних файлів, формування повних пар .DBF/.FPT та перевірку цілісності MEMO-полів. У процесі конвертації враховано особливості кодування Windows-1251, наявність пошкоджених або неповних таблиць, а також ризик автоматичної зміни типів даних під час подальшого відкриття CSV у табличних редакторах. Це дозволило мінімізувати втрати інформації та зберегти логічні зв'язки між пацієнтами, клінічними епізодами та текстовими записами.

Ключові слова: медичні інформаційні системи, МІС «Ескулап», конвертація даних, Python, деперсоналізація, формат dBase, громадське здоров'я, штучний інтелект, машинне навчання.

Abstract. The digital transformation of healthcare requires converting accumulated medical data into formats suitable for analysis. A significant portion of medical information in Ukraine is stored in archives of legacy systems, including the medical information system (MIS) "Esculap". These data rely on the dBase structure (.DBF and .FPT files), which prevents their direct use in modern artificial intelligence and machine learning applications.

This paper describes a technical approach to converting such archives. An algorithm was developed in Python using the libraries dbfread, pandas, and numpy. The proposed method enables extraction and systematization of

depersonalized patient data, diagnoses, and treatment histories. Particular attention is given to resolving text encoding issues and processing MEMO fields stored in .FPT files.

The result of the study is the transformation of relational tables in legacy formats into CSV files compatible with contemporary analytical tools. The resulting datasets can be used for machine learning tasks, neural network development, and statistical research in public health.

The proposed approach was tested on archival data of the State Institution of Science "Center of Innovative Healthcare Technologies" of the State Administrative Department. The developed algorithm ensures accurate preparation of longitudinal medical data for further morbidity analysis. The software tool enables annual updates of research databases by converting newly generated archival records into a structured and analysis-ready format.

The practical implementation of the algorithm included the inventory of 195 archival files, identification of complete .DBF/.FPT pairs, and validation of MEMO field integrity. The conversion process accounted for Windows-1251 encoding specifics, the presence of corrupted or incomplete tables, and the risk of automatic data type alteration when opening CSV files in spreadsheet editors. This approach minimized information loss and preserved logical links between patients, clinical episodes, and textual medical records.

Keywords: medical information systems, MIS "Esculap", data conversion, Python, depersonalization, dBase format, public health, artificial intelligence, machine learning.

Дослідження виконане в рамках НДР "Медико-соціальне обґрунтування, розробка та запровадження моделі «Центр інноваційних технологій охорони здоров'я» на основі триєдності науки, освіти та практики в роботу багатопрофільного закладу охорони здоров'я і визначення її ролі у формуванні єдиного медичного простору" (реєстраційний номер 0125U000318).

Вступ

Зростання обсягів медичних даних у закладах охорони здоров'я та активне впровадження аналітичних систем визначають потребу у використанні структурованих і стандартизованих підходів до їх обробки. Це узгоджується з положеннями *Стратегії розвитку системи охорони здоров'я України до 2030 року* [1], де акцентовано на розбудові цифрової інфраструктури та використанні даних для прийняття клінічних рішень. Проте значна частина медичних інформаційних систем, розроблених у попередні десятиліття, зберігає дані у застарілих форматах, що не передбачають прямого експорту у формати, сумісні з сучасними інструментами статистичного аналізу та машинного навчання. Це створює бар'єри для впровадження алгоритмів прогнозування та систем підтримки прийняття рішень.

Як свідчать результати досліджень Kalkatawi та ін. (2025) [2], Chen та ін. (2023) [3], Huang та ін. (2020) [4] і Markus та ін. (2021) [5], основними проблемами використання історичних медичних даних є фрагментація джерел, низька інтероперабельність, відсутність уніфікованих стандартів, неповнота та шум у записах. Такі обмеження ускладнюють формування якісних навчальних вибірок та суттєво впливають на точність моделей штучного інтелекту. Тому ключовим завданням стає забезпечення цілісності даних, збереження логічних зв'язків між клінічними сутностями та дотримання вимог конфіденційності.

У Державній науковій установі «Центр інноваційних технологій охорони здоров'я» Державного управління справами (ДНУ «ЦІТОЗ» ДУС) експлуатується медична інформаційна система «Ескулап», яка зберігає архіви у форматах .DBF та .FPT, характерних для ранніх реалізацій dBASE. Обмежена інтероперабельність, складна внутрішня структура та відсутність технічної документації унеможливлюють пряме використання цих даних у сучасних аналітичних платформах. Тому актуальним є розроблення технічної методики, що дозволить інтегрувати історичні архіви МІС «Ескулап» у процеси побудови моделей штучного інтелекту для медичного прогнозування.

Мета даного дослідження - розробити відтворювану та масштабовану методику трансформації архівних даних медичної інформаційної системи (МІС) «Ескулап» у структурований формат, придатний для статистичного аналізу та навчання моделей машинного опрацювання медичних даних.

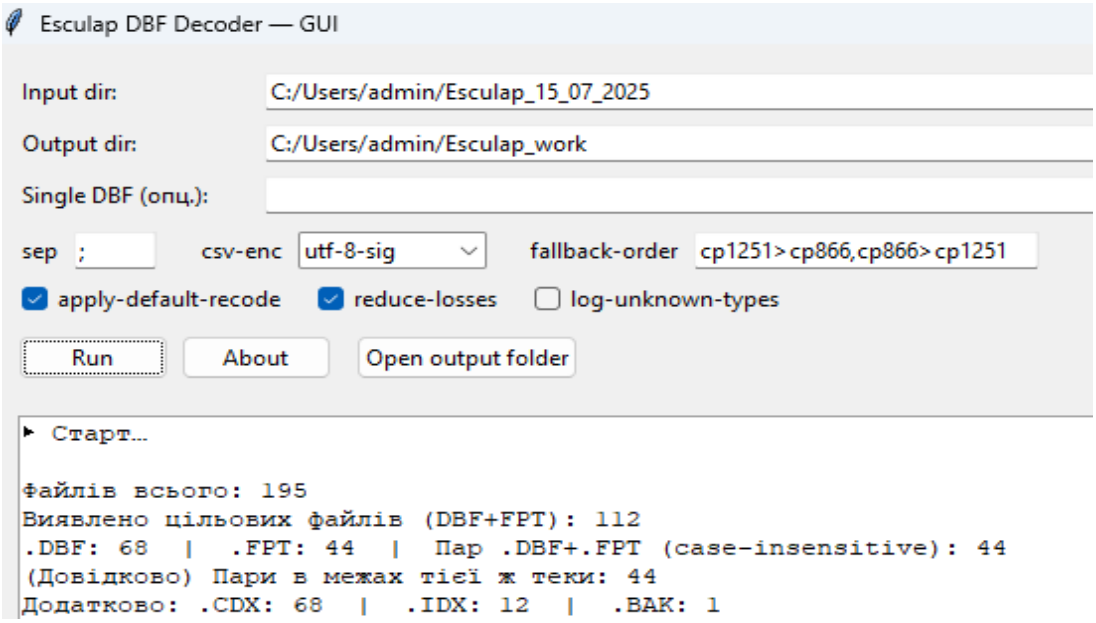

Рис. 1. Вивід скрипта підрахунку кількості файлів та розподілу за розширеннями в архіві МІС «Ескулап»

Особливу увагу приділено етапу первинного очищення (видалення технічних, порожніх та нерелевантних полів) та забезпеченню можливості формування вибірок для медичного прогнозування.

Для досягнення поставленої мети у роботі використано деперсоналізовані архівні дані амбулаторних пацієнтів у xBase-подібних форматах медичної інформаційної системи «Ескулап», що експлуатується у Державній науковій установі «Центр інноваційних технологій охорони здоров'я» Державного управління справами. Система охоплює повний цикл амбулаторної допомоги, включно з реєстрацією звернень, діагностичними процедурами, призначеннями та текстовими клінічними записами.

Архіви зберігаються у форматах, сумісних із сімейством xBase, зокрема .DBF (структуровані таблиці) та .FPT (MEMO-поля). Ці формати є характерними для програмних рішень початку 2000-х років і не забезпечують прямої сумісності з сучасними аналітичними платформами.

Для дослідження було отримано повні копії баз даних, попередньо деперсоналізовані ІТ-відділом шляхом видалення або заміни ідентифікаційних даних (ПІБ, адреси, дати народження та ін.). Усі процедури виконувалися на відокремлених копіях архівів, що унеможливлювало вплив на роботу діючої МІС.

Структура архіву та компоненти даних

Під час інвентаризації архівів МІС «Ескулап» (Рис. 1) виявлено 195 файлів, серед яких 68 таблиць формату .DBF та 44 відповідні файли .FPT, що утворюють 44 повні структурні пари. Коректні пари .DBF +.FPT забезпечують можливість зчитування MEMO-полів та відтворення текстових клінічних записів. Частина таблиць .DBF не має супровідного .FPT, що свідчить про неповноту окремих наборів, однак не перешкоджає опрацюванню структурованої частини таблиці. Решту становили службові файли: .CDX (68), .IDX (12), .BAK (1), а також два технічні файли без розширення типу .DS_Store, які не містили медичних даних.

Архіви включали п'ять основних типів файлів (Рис. 2):

– .DBF - структуровані таблиці даних [6];

– .FPT - MEMO-файли для текстових полів (анамнез, скарги, рекомендації) [7];

– .CDX [8] та .IDX [10] - індекси без клінічного змісту;

– .BAK - резервні копії [9].

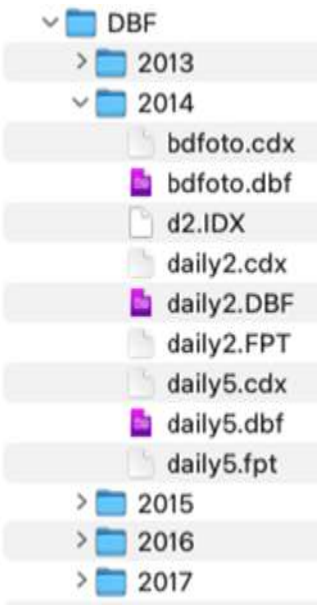

Рис. 2. Типи файлів, виявлені у складі архіву МІС «Ескулап» (фрагмент)

Для таблиць із MEMO-полями необхідна наявність парного .FPT; за його відсутності значення MEMO-полів повертаються як None відповідно до специфікації dbfread [11].

Первинна перевірка охоплювала контроль наявності парних файлів, відповідність структури таблиць стандартам МІС «Ескулап», коректність MEMO-посилань, а також виключення службових або пошкоджених файлів. Результати фіксувалися у зведеному журналі (Рис. 3).

```
=== ПІДСУМОК ===
Усього DBF:              68
Успішно (CSV):           60
Помилка відкриття:       0
CSV-помилка запису:      0
Без валідних рядків:     8
Рядків збережено:        9705495
Рядків відкинуто:        0
Час загалом (сек):       1754.96

Зведені логи:
  /Users/admin/Desktop/Esculap_v1.0/_master_failed_files.log
  /Users/admin/Desktop/Esculap_v1.0/_master_badrows.log
```

Рис. 3. Зведений журнал обробки та виявлених проблемних таблиць (фрагмент)

Файли, що містили зсуви полів, биті рядки або порожню структуру, автоматично позначалися як проблемні та переносилися у «карантинну» директорію для повторної ручної перевірки (Рис. 4).

_master_failed_files.log

```
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/2013/bdfoto.dbf
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/2014/bdfoto.dbf
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/2015/prof.dbf
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/2020/prof.dbf
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/2021/prof.dbf
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/2022/prof.dbf
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/2023/prof.dbf
NO_VALID_ROWS    /Users/admin/Desktop/Esculap_15_07_2025/DBF/1koe.dbf
```

Рис. 4. Структура карантинного архіву з файлами, що не містять валідних записів

Такий підхід забезпечив збереження максимальної кількості валідних записів та формування репрезентативної структури даних для подальшої трансформації у CSV.

Програмне забезпечення та інструменти

Для автоматизованої обробки архівів застосовано мову програмування Python 3.x [12] з такими бібліотеками:

– pandas - робота зі структурованими табличними даними [13];

– numpy - базові математичні операції та векторизація обчислень [14];

– dbfread - зчитування .DBF таблиць, включно з підтримкою MEMO-полів за наявності відповідних .FPT файлів [15];

– chardet - автоматичне визначення кодування текстових полів [16];

– os та logging - обходження файлової системи, контроль структури архівів та ведення журналів обробки [17], [18].

Обчислювальні процедури виконувалися у середовищах командного рядка Windows Terminal [19] та macOS Terminal [20], які забезпечують повну сумісність із Python-скриптами та підтримку обробки великих обсягів даних без участі графічного інтерфейсу.

На основі розробленого ядра алгоритму у подальшому створено десктопний програмний додаток, що забезпечує зручний інтерфейс для запуску процедур конвертації, моніторингу процесу обробки та контролю коректності сформованих файлів. Розроблений програмний інструмент «Esculap DBF Decoder (core) v1.0» зареєстровано як комп'ютерну програму. Свідоцтво про реєстрацію авторського права на твір № 141779 видано 26 січня 2026 року. Автор - Горачук А. М.; майнові права належать Державній науковій установі «Центр інноваційних технологій охорони здоров'я» Державного управління справами.

Як допоміжний засіб точкової верифікації застосовувалося комерційне програмне забезпечення DBF Commander Professional [21], виключно для перегляду проблемних таблиць або усунення одиничних структурних дефектів. У розрахунковий процес воно не входило.

Таким чином, проведений аналіз структури архіву дозволяє ідентифікувати ключові вузли даних та визначити вимоги до програмного інструментарію для їх автоматизованої конвертації.

Алгоритм обробки даних

Розроблений обчислювальний пайплайн реалізує повністю автоматизовану процедуру трансформації архівів МІС «Ескулап» і охоплює такі етапи:

1. Рекурсивний обхід файлової структури. Виконується пошук усіх файлів із розширеннями .DBF, .FPT, .CDX, .IDX та .BAK незалежно від рівня вкладеності каталогів.

2. Формування пар .DBF + .FPT. Для кожної таблиці .DBF перевіряється наявність відповідного MEMO-файлу (.FPT). Якщо .FPT відсутній, таблиця обробляється лише у межах структурованих полів. Ізольовані .FPT визначаються як неповні та виключаються з обробки.

3. Зчитування структурованих і текстових полів. Таблиці .DBF зчитуються з урахуванням кодування Windows-1251, характерного для кириличних xBase-структур. За наявності повної пари виконується прив'язка MEMO-полів із відповідного .FPT.

4. Перевірка цілісності таблиць. Для кожного файлу здійснюється оцінка відповідності типів полів, коректності MEMO-посилань, виявлення порожніх чи нульових файлів (до 1 КБ), а також відповідності найменувань файлів стандартам МІС «Ескулап».

5. Нормалізація структури даних. Назви колонок уніфікуються шляхом приведення до нижнього регістру, видалення службових символів та, за потреби, транслітерації кириличних позначень.

6. Логування обробки. Під час виконання формуються структуровані журнали, які містять інформацію про кількість рядків, наявність парних .FPT-файлів, виявлені дефекти, обсяг даних та

підсумковий статус («успішно», «карантин», «помилка читання»).

7. Обробка дефектних таблиць. Файли, що не відповідають вимогам цілісності, автоматично переміщуються в карантинну директорію з відповідним записом у журналі обробки. Повторна верифікація можлива після ручного усунення дефектів.

8. Експорт результатів. Усі валідні таблиці експортуються у формат CSV (UTF-8) із збереженням структури каталогів та вихідних ідентифікаторів, що забезпечує відтворення зв'язків між пацієнтами, візитами та клінічними епізодами.

Така послідовність дій забезпечує цілісність даних (data integrity), що є критичним для запобігання помилок при подальшому навчанні нейронних мереж.

Синхронізація записів

Уніфікований ідентифікатор id_pt, сформований ІТ-відділом для кожного пацієнта, був ключовим механізмом узгодження між таблицями. За відсутності id_pt використовувалися альтернативні унікальні ідентифікатори, наявні у вихідних даних. Ідентифікатори інтегровано у всі базові таблиці (пацієнти, візити, призначення тощо) ще на етапі деперсоналізації, що забезпечило коректне об'єднання інформації без ризику розкриття персональних даних.

Вибіркова ручна валідація виконувалася шляхом зіставлення записів з різних таблиць для одного й того самого ідентифікатора. Такий підхід дозволив підтвердити повноту синхронізації, виключити дублювання та запобігти втраті інформації.

Результати та їх обговорення

У процесі конвертації архівів МІС «Ескулап» реалізовано повністю автоматизовану трансформацію даних із форматів .DBF/.FPT у CSV (UTF-8), придатний для статистичного та машинного аналізу.

1. Структура вхідних даних. Рекурсивне сканування одинадцятирічного архіву виявило 195 файлів, серед яких 68 структурованих таблиць .DBF і 44 відповідні .FPT, що утворюють 44 повні пари. Службові файли (.CDX, .IDX, .BAK) виключено як технічно нерелевантні.

2. Обробка та верифікація. Таблиці .DBF зчитано з урахуванням кодування Windows-1251; за відсутності .FPT оброблялися лише структуровані поля. Фіксувалися типові дефекти (биті рядки, зсуви полів, порожні файли до 1 КБ), а MEMO-поля перевірялися на можливі персональні дані з подальшим видаленням таких фрагментів.

3. Структура сформованих даних. Сформовано уніфікований набір CSV-файлів, що охоплює основні медичні сутності: пацієнтів (id_pt), діагнози (МКХ-10), текстові клінічні записи, візити, процедури та призначення. Усі поля збережено у вихідному вигляді; приклад наведено на Рис. 5.

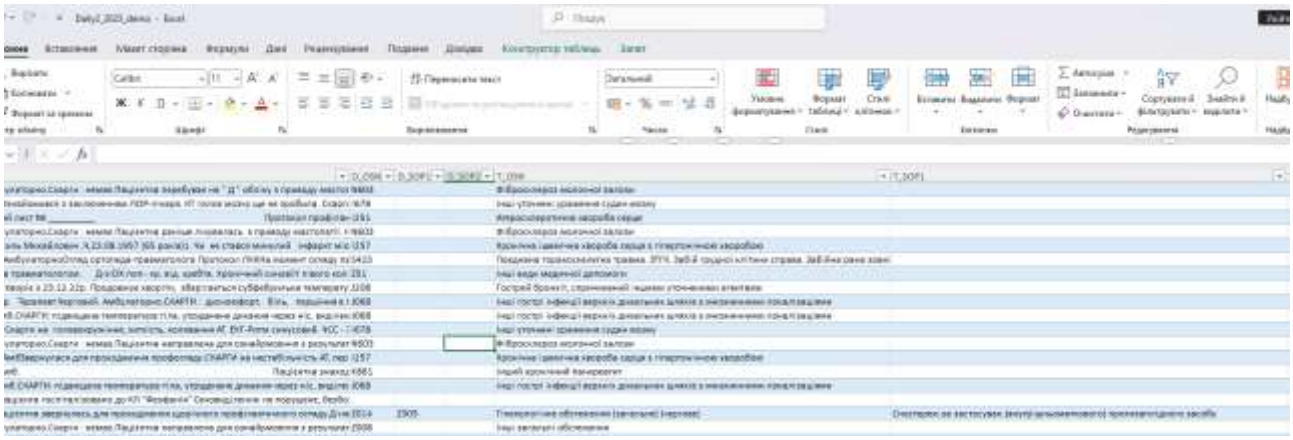

Рис. 5. Фрагмент обробленого результату пари файлів .DBF та .FPT після конвертації у табличний формат

4. Логічні зв’язки. Перевірено відповідність між записами пацієнтів, діагнозами, текстовими блоками та подіями лікування. Збережено метадані про джерело кожного запису.

5. Готовність до подальшого аналізу. Отриманий масив є цілісним та уніфікованим; він готовий до очищення, стандартизації змінних і формування навчальних вибірок для моделей штучного інтелекту.

6. Запропонована методика демонструє практичний спосіб адаптації архівних медичних даних до сучасних аналітичних платформ без модифікації медичної інформаційної системи та без використання спеціалізованих модулів експорту. На відміну від традиційних рішень (експорт через FoxPro, формування PDF або виконання разових запитів до адміністратора), алгоритм забезпечує повну автоматизацію, рекурсивну обробку вкладених структур і відтворюваність процесу.

7. Ключовою перевагою є акцент на змістовних форматах даних: до обробки включаються лише файли зі структурними (.DBF) або текстовими (.FPT) записами, тоді як службові індексні об’єкти (.CDX, .IDX, .BAK) свідомо виключаються. Це мінімізує ризик технічних помилок і спрощує подальшу аналітику.

8. Алгоритм враховує типові проблеми архівних xBase-систем: пошкоджені рядки, розсинхронізацію MEMO-посилань, неповні структури, дублікати. Реалізовано комбінований механізм автоматичного логування та можливості ручної корекції з повторним зчитуванням, що дозволяє відновлювати інформаційно значущі дані.

9. Використання уніфікованого ідентифікатора пацієнта (id_pt), сформованого ІТ-відділом, забезпечує узгодженість між таблицями та дає змогу формувати часові вибірки, необхідні для статистичного аналізу, побудови time-to-event моделей і алгоритмів прогнозування. Це є критично важливим для патологій із вираженою часовою динамікою клінічних подій.

10.Формат CSV забезпечує сумісність із Python, R, SPSS, MedCalc та Excel, а також дозволяє інтегрувати структуровані й текстові поля в єдиний аналітичний контур. Апробація в умовах реального медичного закладу підтвердила можливість масштабування підходу на інші МІС, побудовані на dBASE/FoxPro-архітектурі.

Обмеження дослідження та технічні виклики

У процесі реалізації алгоритму було виявлено низку специфічних проблем, зумовлених застарілістю формату dBASE та особливостями збереження медичної інформації в МІС «Ескулап»:

– Неоднорідність кодування: незважаючи на те, що основна частина архіву використовує кириличне кодування Windows-1251, окремі системні довідники та старі записи потребували додаткової обробки для коректного відображення символів. Використання стандартних засобів експорту часто призводило до появи «артефактів» у тексті, що було вирішено шляхом примусового призначення параметру encoding='cp1251' у скрипті обробки.

– Конфлікти форматів при експорті (CSV Delimiter). Стандартне використання коми як розділювача виявилося непридатним для даного масиву даних. Емпіричним шляхом (методом підбору) було встановлено, що лише використання крапки з комою (;) дозволяє зберегти читабельність структури. Це пов’язано з тим, що MEMO-поля (анамнези, описи діагнозів) містять велику кількість авторських знаків пунктуації та ком, які при стандартному розділювачі «руйнували» таблицю, зміщуючи дані по стовпцях.

– Автоматична зміна типів даних. Виявлено критичну проблему при обробці результатів лабораторних аналізів. Окремі числові показники (наприклад, концентрації з десятковою крапкою «1.25») при відкритті у табличних редакторах помилково інтерпретувалися як календарні дати. Це вимагало впровадження додаткового етапу верифікації та примусового приведення типів даних у

стовпцях, щоб уникнути втрати точності, що є критичним для медичних ШІ-моделей.

Дотримання етичних норм

Усі дані було повністю деперсоналізовано до початку будь-якої обробки. Використаний технічний ідентифікатор id_pt, сформований ІТ-відділом установи, не містить жодної персональної інформації та не дозволяє ідентифікувати пацієнтів.

Опрацювання виконувалося виключно на копіях архівів, відокремлених від робочих клінічних баз, що унеможливлювало будь-який вплив на реальні медичні процеси. Текстові MEMO-поля проходили додаткову перевірку на наявність потенційно чутливих фрагментів; виявлені порушення анонімності вилучалися перед формуванням результатів.

Усі процедури відповідали вимогам конфіденційності, інформаційної безпеки та належного використання медичних даних і узгоджувалися з чинними національними та міжнародними рекомендаціями щодо етичної роботи з медичною інформацією.

Висновки

У роботі реалізовано повний цикл перетворення архівних даних МІС «Ескулап» із форматів .DBF та .FPT у структуровані CSV-файли, придатні для подальшого статистичного та машинного аналізу. Розроблений алгоритм поєднує автоматизоване зчитування з контрольованими етапами ручної валідації, що забезпечує стійкість до дефектних чи неповних структур файлів.

Досягнуто такі результати:

– сформовано відтворюваний та масштабований пайплайн попередньої підготовки архівних медичних даних;

– забезпечено сумісність підготовлених CSV з Python, R, SPSS, MedCalc, Excel та іншими інструментами;

– доведено можливість повноцінної конверсії без втручання у роботу МІС та без використання зовнішніх модулів експорту;

– гарантовано дотримання етичних норм і вимог конфіденційності шляхом застосування унікальних ідентифікаторів, що не розкривають особистість пацієнта;

– підтверджено можливість масштабування методики на інші МІС, побудовані на архітектурі Clipper/FoxPro/dBASE-подібних систем. Отриманий структурований масив даних формує технічну основу для побудови навчальних вибірок та розробки моделей прогнозування захворювань і клінічних подій з можливістю подальшого розширення на інші напрямки аналізу.

Перспектива подальших досліджень

Подальший етап передбачає поглиблену підготовку отриманих даних шляхом усунення технічних та нерелевантних елементів. Планується:

– видалення службових полів, що не несуть аналітичного змісту;

– усунення порожніх або майже порожніх колонок;

– виключення дублюючих або нефункціональних змінних, не релевантних меті дослідження.

Після очищення буде сформовано узгоджений датасет для побудови моделей штучного інтелекту. Першочергово планується розробка нейромережевої моделі прогнозування медичних подій, після чого можливе створення спеціалізованих моделей для окремих нозологічних груп та time-to-event-моделей для оцінки ризиків у динаміці. Передбачена також інтеграція ознак із текстових MEMO-полів із застосуванням методів NLP, що дасть змогу врахувати клінічну інформацію, відсутню у структурованій частині таблиць.

Оскільки МІС «Ескулап» є комерційним продуктом стороннього розробника, у різних медичних закладах можуть використовуватися системи з іншою архітектурою та форматами даних. Для їх обробки алгоритм вимагатиме адаптації, зокрема для рішень на базі 1С, MedSTAR, HealthSystem, а також локальних інсталяцій Clipper/FoxPro/dBASE. Отримані результати формують основу для розробки уніфікованого підходу до роботи з

історичними медичними архівами в Україні.

Література

1. Кабінет Міністрів України. (2025). “Стратегія розвитку системи охорони здоров’я на період до 2030 року” (Розпорядження № 34-р). Retrieved from https://zakon.rada.gov.ua/laws/show/34-2025-%D1%80#Text
2. Kalkatawi, M. Beyond the upgrade: unraveling the complexities of health information system migration. *Discov Health Systems* 4, 7 (2025). https://doi.org/10.1007/s44250-025-00186-x
3. Chen, W., Xie, F., McCarthy, D. P., Reynolds, K. L., Lee, M., Coleman, K. J., Getahun, D., Koebnick, C., & Jacobsen, S. J. (2023). Research data warehouse: using electronic health records to support clinical and translational research. JAMIA Open, 6(2), e2023–039. https://doi.org/10.1093/jamiaopen/ooad039
4. Huang, C., Koppel, R., McGreevey, J. D. III, Craven, C. K., & Schreiber, R. (2020). Transitions from one electronic health record to another: challenges, pitfalls, and recommendations. *Applied Clinical Informatics, 11*(5), 742–754. https://doi.org/10.1055/s-0040-1718535
5. Markus, A. F., Kors, J. A., & Rijnbeek, P. R. (2021). The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. *Journal of Biomedical Informatics, 113*, 103655. https://doi.org/10.1016/j.jbi.2020.103655
6. dBASE LLC. (n.d.). *dBASE Table File Format (DB7)*. Retrieved from https://www.dbase.com/Knowledgebase/INT/db7_file_fmt.htm
7. Microsoft Corporation. (n.d.). MEMO File Structure (.FPT) - Visual FoxPro 9.0. Retrieved from https://www.vfphelp.com/help/_5WN12PC0N.htm
8. Microsoft Corporation. (n.d.). *Compound Index File Structure (CDX)* - Visual FoxPro 9.0 Help. Retrieved from https://www.vfphelp.com/help/_5wn12pbyu.htm
9. Microsoft Corporation. (2024, December 17). Back Up and Restore of SQL Server Databases. Retrieved from https://learn.microsoft.com/en-us/sql/relational-databases/backup-restore/back-up-and-restore-of-sql-server-databases
10. Microsoft Corporation. (n.d.). Index File Structure (.IDX) - Visual FoxPro 9.0 SP2 Help. Retrieved from https://www.vfphelp.com/help/html/fbad54cc-cf7f-4add-a0d9-ddbeec5e00cc.htm
11. Ole Martin Bjørndalen. (n.d.). *DBFread documentation - Introduction*. Retrieved from https://dbfread.readthedocs.io/en/latest/introduction.html
12. Python Software Foundation. (n.d.). Python 3.x Documentation. Retrieved from https://docs.python.org/3/
13. Pandas Development Team. (2025). pandas: powerful Python data analysis toolkit. Retrieved from https://pandas.pydata.org/docs/
14. NumPy Developers. (n.d.). NumPy Documentation. Retrieved from https://numpy.org/doc/
15. Ole Martin Bjørndalen. (n.d.). DBFread - Read DBF Files with Python. Retrieved from https://dbfread.readthedocs.io/en/latest/
16. Chardet Developers. (n.d.). chardet 5.2.0 documentation. Retrieved from https://chardet.readthedocs.io/
17. Python Software Foundation. (n.d.). os - Miscellaneous operating system interfaces. In Python 3.x documentation. Retrieved from https://docs.python.org/3/library/os.html
18. Python Software Foundation. (n.d.). logging - Logging facility for Python. In Python 3.x documentation. Retrieved from https://docs.python.org/3/library/logging.html
19. Microsoft. (n.d.). Windows Terminal Documentation. Retrieved from https://learn.microsoft.com/en-us/windows/terminal/
20. Apple Inc. (n.d.). What is Terminal on Mac? Retrieved from https://support.apple.com/guide/terminal/what-is-terminal-trmld4c92d55/mac
21. Elphsoft. (n.d.). *DBF Commander Professional Online Help*. Retrieved from https://dbf-software.com/help
22. Горачук А. М. Комп’ютерна програма «Esculap DBF Decoder (core) v1.0» : свідоцтво про реєстрацію авторського права на твір № 141779 від 26.01.2026. Київ : Український національний офіс інтелектуальної власності та інновацій, 2026.

References

1. Cabinet of Ministers of Ukraine. (2025). “Strategy for the Development of the Healthcare System until 2030 ” (Order № 34-p). Retrieved from https://zakon.rada.gov.ua/laws/show/34-2025-%D1%80?lang=en#Text
2. Kalkatawi, M. Beyond the upgrade: unraveling the complexities of health information system migration. *Discov Health Systems* 4, 7 (2025). https://doi.org/10.1007/s44250-025-00186-x
3. Chen, W., Xie, F., McCarthy, D. P., Reynolds, K. L., Lee, M., Coleman, K. J., Getahun, D., Koebnick, C., & Jacobsen, S. J. (2023). Research data warehouse: using electronic health records to support clinical and translational research. JAMIA Open, 6(2), e2023–039. https://doi.org/10.1093/jamiaopen/ooad039
4. Huang, C., Koppel, R., McGreevey, J. D. III, Craven, C. K., & Schreiber, R. (2020). Transitions from one electronic health record to another: challenges, pitfalls, and recommendations. *Applied Clinical Informatics, 11*(5), 742–754. https://doi.org/10.1055/s-0040-1718535
5. Markus, A. F., Kors, J. A., & Rijnbeek, P. R. (2021). The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and

evaluation strategies. *Journal of Biomedical Informatics, 113*, 103655. https://doi.org/10.1016/j.jbi.2020.103655

6. dBASE LLC. (n.d.). *dBASE Table File Format (DB7)*. Retrieved from https://www.dbase.com/Knowledgebase/INT/db7_file_fmt.htm
7. Microsoft Corporation. (n.d.). MEMO File Structure (.FPT) - Visual FoxPro 9.0. Retrieved from https://www.vfphelp.com/help/_5WN12PC0N.htm
8. Microsoft Corporation. (n.d.). *Compound Index File Structure (CDX)* - Visual FoxPro 9.0 Help. Retrieved from https://www.vfphelp.com/help/_5wn12pbyu.htm
9. Microsoft Corporation. (2024, December 17). Back Up and Restore of SQL Server Databases. Retrieved from https://learn.microsoft.com/en-us/sql/relational-databases/backup-restore/back-up-and-restore-of-sql-server-databases
10. Microsoft Corporation. (n.d.). Index File Structure (.IDX) - Visual FoxPro 9.0 SP2 Help. Retrieved from https://www.vfphelp.com/help/html/fbad54cc-cf7f-4add-a0d9-ddbeec5e00cc.htm
11. Ole Martin Bjørndalen. (n.d.). *DBFread documentation - Introduction*. Retrieved from https://dbfread.readthedocs.io/en/latest/introduction.html
12. Python Software Foundation. (n.d.). Python 3.x Documentation. Retrieved from https://docs.python.org/3/
13. Pandas Development Team. (2025). pandas: powerful Python data analysis toolkit. Retrieved from https://pandas.pydata.org/docs/
14. NumPy Developers. (n.d.). NumPy Documentation. Retrieved from https://numpy.org/doc/
15. Ole Martin Bjørndalen. (n.d.). DBFread - Read DBF Files with Python. Retrieved from https://dbfread.readthedocs.io/en/latest/
16. Chardet Developers. (n.d.). chardet 5.2.0 documentation. Retrieved from https://chardet.readthedocs.io/
17. Python Software Foundation. (n.d.). os - Miscellaneous operating system interfaces. In Python 3.x documentation. Retrieved from https://docs.python.org/3/library/os.html
18. Python Software Foundation. (n.d.). logging - Logging facility for Python. In Python 3.x documentation. Retrieved from https://docs.python.org/3/library/logging.html
19. Microsoft. (n.d.). Windows Terminal Documentation. Retrieved from https://learn.microsoft.com/en-us/windows/terminal/
20. Apple Inc. (n.d.). What is Terminal on Mac? Retrieved from https://support.apple.com/guide/terminal/what-is-terminal-trmld4c92d55/mac
21. Elphsoft. (n.d.). *DBF Commander Professional Online Help*. Retrieved from https://dbf-software.com/help
22. Horachuk A. M. Esculap DBF Decoder (core) v1.0: computer program. Certificate of copyright registration for a work No. 141779, January 26, 2026. Kyiv: Ukrainian National Office of Intellectual Property and Innovations, 2026.

The article has been sent to the editors 04.02.26.
After processing 24.02.26.
Submitted for printing 30.03.26.

Copyright under license CCBY-SA4.0.