

## ІНФОРМАЦІЙНА ПІДТРИМКА ПРОЦЕСІВ СОЦІАЛІЗАЦІЇ ОСОБИСТОСТІ НА ОСНОВІ СПІЛЬНИХ ІНТЕРЕСІВ

Тарас Батюк<sup>1</sup>, Вікторія Висоцька<sup>2</sup>

Національний університет “Львівська політехніка”, кафедра інформаційних систем та мереж,

<sup>1</sup> e-mail: taras.batiuk.mnsa.2020@lpnu.ua, ORCID: 0000-0001-5797-594X,

<sup>2</sup> e-mail: Victoria.A.Vysotska@lpnu.ua, ORCID: 0000-0001-6417-3689

© Батюк Т., Висоцька В., 2022

Створено проєкт інформаційної системи для соціалізації за особистими інтересами на основі SEO-технологій та методів машинного навчання. Основна мета цієї інформаційної системи – ідентифікація користувача в системі за допомогою нейронних мереж і вибір подібних користувачів на підставі аналізу поточної інформації користувача. Створено інформаційну систему, яка за допомогою токенів Identity та JWT забезпечує оптимізовані та безпечні функції авторизації, реєстрації та підтримки поточного сеансу користувача системи. Пошук обличчя на фотографії користувача та перевірку наявності подібного користувача в базі даних реалізовано за допомогою згорткових і сіамських нейронних мереж. Аналіз та формування подібних гудків користувачів реалізовано за допомогою алгоритмів нечіткого пошуку, алгоритму Левенштейна та моделі Noisy Channel, що дало змогу максимально автоматизувати вибір користувача та оптимізувати час, витрачений на цей процес. Також створено інструменти для перегляду профілів інших користувачів, уподобань та особистого листування. Уся приватна кореспонденція та інформація про неї зберігаються в поточній базі даних. Кожен користувач інформаційної системи може переглянути всю інформацію про надіслані та отримані повідомлення. Створена інформаційна система реалізує ідентифікацію користувачів, аналіз, відбір та подальшу соціалізацію користувачів.

**Ключові слова:** проєкт; нечіткий пошук; згорткова нейронна мережа; сіамська нейронна мережа; відстань Левенштейна; шумовий канал.

### Вступ

Низка публікацій стосується актуальної проблеми, а саме дослідження та аналізу взаємодії користувачів між собою всередині соціальних мереж [1–3]. Тема актуальна, оскільки в сучасному світі більшість людей використовують соціальні мережі, проводячи в них величезну частину свого вільного часу, відповідно аналіз поведінки користувачів у тих чи інших ситуаціях під час користування соціальними мережами дуже важливий і дає змогу одержувати інформацію про користувача на основі його діяльності. В цих роботах розглянуто основні аспекти взаємодії користувачів між собою, що передбачає аналіз певних користувацьких повідомлень та реакцію на ці повідомлення інших користувачів. Це дає змогу визначити певні закономірності та параметри користувачів системи, а це, своєю чергою, додатково автоматизувати роботу інформаційної системи. Також висвітлено реалізацію механізмів аналізу та перевірки поточного перебування у мережі користувачів, що є дуже важливим моментом для динамічних соціальних мереж, які відстежують основні дії користувача інформаційної системи заради оптимізації роботи тих чи інших механізмів системи. Але існують важливі актуальні проблеми/завдання модерації інтернет-

контенту всередині соціальних мереж, що даватиме змогу ефективно аналізувати тексти користувачів та реагувати на ті чи інші порушення правил.

Мета цієї праці – розроблення інформаційної системи для соціалізації за особистими інтересами на основі SEO-технологій та методів машинного навчання.

Основні завдання наукового дослідження:

1. Визначення та порівняльний аналіз основних метрик функціонування інформаційної системи для соціалізації за особистими інтересами.
2. Розроблення функціональних схем типової інформаційної системи для соціалізації за особистими інтересами.
3. Реалізація процесів ідентифікації користувача в системі за допомогою нейронних мереж і вибору подібних користувачів на підставі аналізу поточної інформації користувача.
4. Розроблення інформаційної системи для соціалізації за особистими інтересами, яка за допомогою токенів Identity та JWT має забезпечувати оптимізовані та безпечні функції авторизації, реєстрації та підтримки поточного сеансу користувача системи.
5. Реалізація процесів пошуку обличчя на фотографії користувача та перевірка наявності подібного користувача в базі даних за допомогою згорткових і сіамських нейронних мереж.
6. Розроблення модуля аналізу формування подібних гудків користувачів за допомогою алгоритмів нечіткого пошуку, алгоритму Левенштейна та моделі Noisy Channel, що дасть змогу максимально автоматизувати вибір користувача та оптимізувати час, витрачений на цей процес.
7. Створення інструментів для перегляду профілів інших користувачів, уподобань та особистого листування.
8. Проведення експериментальної апробації розробленої інформаційної системи для соціалізації за особистими інтересами на основі SEO-технологій та методів машинного навчання.
9. Аналіз отриманих результатів експериментальної апробації розробленої інформаційної системи для соціалізації за особистими інтересами на основі SEO-технологій та методів машинного навчання.

*Об'єктом дослідження* є процес соціалізації користувачів на основі SEO-технологій та методів машинного навчання. В наш час більшість соціальних мереж намагаються максимально спростити та оптимізувати цей процес, відповідно, щоб конкурувати з іншими соціальними мережами.

*Предмет дослідження* – методи та засоби соціалізації за особистими інтересами користувача інформаційної системи. Оскільки інформаційна система створюється передусім для кінцевого користувача, який її використовуватиме відповідно до потреб, то необхідно дослідити і звернути увагу на потенційних користувачів інформаційних систем на те, як застосовують системи-аналоги, що саме людей приваблює та відштовхує в інформаційних системах певного виду, як саме краще покроково здійснити роботу системи, щоб користувачу все було очевидно та зрозуміло, щоб користувач міг легко знаходити інших користувачів, взаємодіяти з ними. Відповідно виявлені користувачі повинні мати близький індекс схожості, що дасть змогу користувачам зі спільними інтересами знаходити один одного та взаємодіяти за допомогою спілкування, або особистого, або у вигляді коментарів до публікацій інших користувачів. Також кожен користувач відповідно зможе створювати власні публікації на певні теми і знаходити користувачів зі спільними інтересами, які зможуть ці теми порушувати та обговорювати. Чим оптимальніше продумані всі такі процеси, тим довше в системі затримається предмет цього дослідження, а саме користувач системи, і чим довше він затримається у системі, тим більший шанс, що відбудеться процес соціалізації.

*Наукова новизна одержаних результатів* полягає у визначенні множини метрик функціонування подібних ІС та виконанні їх порівняльного аналізу, розробленні відповідних функціональних схем функціонування таких систем та реалізації типової інформаційної системи.

Поки що не існує аналогічних інформаційних систем для соціалізації користувачів, які б виконували повноцінний аналіз всієї текстової інформації щодо користувача і формували кінцевий ідентифікатор схожості поточного користувача системи з уже наявними користувачами. Відповідно вперше буде створено інформаційну систему, яка практично повністю автоматизує процес соціалізації користувачів. Варто зазначити, що удосконалено алгоритм Левенштейна для нечіткого пошуку, який в цій системі використано не у базовому вигляді, а оптимізовано для паралельного опрацювання одразу декількох наборів текстової інформації, що відповідно дає змогу набагато швидше і зручніше опрацьовувати всю необхідну інформацію і формувати результат. Також вперше застосовано не лише якийсь один певний алгоритм, а одразу чотири різних, а саме алгоритм Левенштейна, алгоритм N-грам, алгоритм розширення вибірки та реалізацію моделі Noisy Channel. Ці алгоритми схожі та ставлять перед собою одну мету – аналіз слів у тексті та формування оцінок, але кожен них підходить для різних видів тексту та типів інформації, яка поділена неструктуровано, відповідно спочатку відбувається поділ тексту на чотири різні вибірки текстових даних, після чого кожен алгоритм здійснює опрацювання тих даних, для яких він найбільше підходить, тим самим оптимізуючи увесь процес нечіткого пошуку, після чого формується кінцевий ідентифікатор схожості на основі отриманих оцінок.

### **Аналіз останніх досліджень та публікацій**

У статті [4] автори подають вибірку даних, отриманих у результаті аналізу соціальних мереж серед різних груп користувачів. Використавши парні параметри, вони сформуливали основні результати, отримані в результаті взаємодії користувачів та аналізу повідомлень, які написали користувачі. Розглянуто застосування спеціального фреймворку для соціалізації користувачів у соціальній мережі Twitter, основні особливості формування хештегів та те, як користувачі з різними вхідними параметрами використовують сформовані хештеги. Сформульовано основні питання, на які потрібно дати відповідь, та алгоритми, які потрібно оптимізувати для того, щоб ефективніше аналізувати інформацію про користувачів у соціальних мережах, а також сформовано концепції додаткових алгоритмів, реалізація яких вплине на аналізування користувацької інформації в соціальних мережах.

Важливі програмні питання відображають автори в статті [5], а саме розглянули принципи побудови та підтримки ефективних серверних систем, які дають змогу витримувати велику кількість користувачів одночасно та підтримувати систему в стабільному стані, незалежно від зовнішніх факторів. Ця проблема надзвичайно актуальна, оскільки величезна кількість людей користується соціальними мережами із телефонів і комп'ютерів. Кількість одночасних користувачів популярної соціальної мережі може сягати мільйонів, кожен з них користується різними системами та пристроями, у кожного користувача різна швидкість інтернету, відповідно потрібно забезпечити видавання та опрацювання інформації для всіх користувачів, і це завдання нині дуже актуальне. Охарактеризовано основні принципи побудови вискоелективних серверних систем, які дають змогу використовувати паралельні запити і, відповідно, можуть опрацьовувати запити асинхронно, формуючи відповідь на запит частинками, динамічно зберігаючи проміжну інформацію. Розглянуто способи збереження інформації та визначення правдивості інформації, яку залишив користувач, висвітлено основні способи уникнення втрат і спотворення інформації та розглянуто способи оптимізації роботи системи збереження даних. Розглянуто основні концепції використання нейронних мереж у контексті взаємодії користувачів всередині соціальних мереж та розкрито основні й ефективні форми аналізу користувацького тексту. Висвітлено основну інформацію, отриману в результаті аналізу нейронною мережею користувацьких повідомлень, залишених у соціальній мережі Instagram, проаналізовано довжину залишених повідомлень та фактори, від яких ця довжина залежить.

У статті [6] автор розглянув певні механізми, призначені для взаємодії клієнтських та серверних систем, висвітлив основні особливості максимально ефективного передавання та збері-

гання проміжної інформації, уникаючи втрат даних та формуючи проміжні запити від користувачів. Незважаючи на велику кількість вирішених завдань у контексті роботи ефективних соціальних мереж, визначено чимало завдань та проблем, які потрібно вирішити. Серед них і багатопотокове опрацювання користувацьких даних без втрати ефективності системи, асинхронне опрацювання вхідних запитів, що дає змогу відокремити клієнтську частину програми від серверної і тим самим передати частину роботи клієнтові, щоб оптимізувати роботу системи. Постало також важливе завдання – повноцінно реалізувати крос-доменну модель передавання даних, що дасть змогу оптимізувати формування запиту на сервер та відповідь самого сервера на запит, через особливості формування та передавання поточного запиту.

У статті [7] автори розглядають використання нейронних мереж у соціальних мережах. В сучасному світі величезна кількість людей користуються соціальними мережами, дивляться відео, слухають аудіо, завантажують та переглядають фотографії. Для автоматизації та оптимізації певних дій часто використовують нейронні мережі, які забезпечують широкий спектр застосування, аналіз аудіо, відео, фотографій, формування та аналіз текстової інформації. Усе це надзвичайно актуально в наш час та активно досліджується під час створення програмних систем. Розглянуто реалізацію нейронної мережі, яка може встановлювати основні потреби користувачів, здійснюючи аналіз списку їхніх покупок в онлайн-магазині за визначеними параметрами. Подано реалізацію нейронної мережі, яка може визначати поточний настрій та основні емоції користувачів на основі повідомлень, які користувачі залишають у соціальних мережах.

У статті [8] автор висвітлив форми визначення наявності користувачів на фотографії у соціальній мережі Instagram, запропонував механізми оптимізації аналізу фотографій з використанням нейронних мереж. Розглянуто основні принципи роботи нейронних мереж із визначення різних взаємопов'язаних об'єктів на вибраних фотографіях у соціальній мережі Instagram, на прикладах фотографій з їжею. Визначено основні механізми взаємодії користувачів між собою у соціальних мережах, запропоновано нейронну мережу, яка, навчаючись на повідомленнях користувачів, формує вибірку із повідомлень, використовуючи зазначені параметри. Застосування нейронних мереж надзвичайно широке, але, незважаючи на продемонстровані можливості стосовно аналізу текстової інформації, аудіоданих та фотографій, є ще багато проблем, які необхідно вирішити, і найзручніше це робити за допомогою нейронних мереж. Це важке завдання, оскільки створення нейронної мережі з нуля передбачає розроблення математичної моделі системи з нуля та здійснення правильного та ефективного навчання нейронної мережі, що дасть змогу оптимізувати роботу соціальних мереж та застосувати зручні функції взаємодії користувачів.

Взаємодія та аналіз користувачів, їхніх запитів та внесеної інформації в різних системах соціалізації та в інтернет-магазинах – актуальне питання, яке розглянув автор у статті [9]. В наш час для інтернет-магазинів важливо оптимізувати пропонування користувачам саме того товару, який вони хочуть знайти. Для виконання цього завдання необхідно аналізувати користувачів, їхні пошукові запити, формувати певну вибірку користувачів для здійснення аналізу та застосовувати алгоритми вибору необхідної інформації із певної кількості текстових даних. Висвітлено реалізацію механізмів підбору користувачів за наявністю певних нормалізованих параметрів та подальше форматування отриманих даних для автоматизації взаємодії користувачів. Розглянуто основні принципи опрацювання медіа-файлів, які завантажили користувачі, а саме фотографій та відео, описано основні принципи асинхронного збереження користувацьких даних. Висвітлено основні алгоритми лінійного пошуку інформації користувачів для визначення основних параметрів взаємодії людей між собою та факторів, які впливають на залишення певної інформації в профілях користувачів, проаналізовано одержані дані. Розглянуто реалізацію нейронної мережі, яка може визначати прогнози користувачів та гри футбольних клубів, здійснюючи аналіз сторінок у соціальних мережах футбольних клубів та коментарів користувачів. Опрацьовано механізм ефективної соціалізації користувачів із застосуванням функціоналу виставлення вподобань користувачів та

оптимізації основних процесів взаємодії користувачів у результаті взаємної соціалізації в одній із соціальних мереж. Висвітлено реалізацію основних механізмів визначення та аналізу потреб користувачів стосовно одягу, із використанням згорткових нейронних мереж та алгоритмів аналізу текстових даних та визначенням і порівнянням основних вибраних параметрів для вибірки користувачів.

Автори статті [10] розглянули основні механізми соціалізації користувачів, які використовують ті самі хештеги в соціальній мережі Twitter, визначили особливості взаємодії користувачів. У роботі висвітлено особливості роботи системи, яка аналізує соціальні мережі користувачів і за допомогою алгоритмів нечіткого пошуку серед текстових даних визначає основні потреби користувачів. Незважаючи на низку вирішених завдань, проблема аналізу та опрацювання користувачьких засобів серйозна і досі, постійно оптимізують алгоритми аналізу тексту та створюють нові алгоритми постопрацювання інформації в результаті поділу та отримання даних.

Аналіз та опрацювання текстової інформації користувача в різних програмних системах та подальше визначення емоцій користувача, прогнозування певних дій яскраво висвітлили автори в статті [11]. Розглянуто реалізацію нейронної мережі, яка визначає емоції та настрої користувачів, аналізуючи декілька їхніх фотографій, нормалізуючи інформацію та формуючи вихідні дані. Висвітлено основні механізми соціалізації користувачів у соціальній мережі Facebook та їх реалізацію, оптимізовано алгоритми взаємодії користувачів між собою протягом певного часу. Показано особливості аналізу користувачьких відгуків щодо куплених товарів, реалізацію механізму формування вибірки та визначення певного відсотка нерелевантних відгуків. Обдумано основні алгоритми аналізу соціальних мереж, які дають змогу визначити основні емоції користувача упродовж певного проміжку часу, оптимізовано алгоритм нечіткого пошуку. Розроблено механізм показу таргетованої реклами користувачеві системи відповідно до основної інформації користувача, яка проаналізована із застосуванням алгоритмів нечіткого пошуку, сформовано вибірку параметрів для кожного користувача. Висвітлено алгоритми аналізу серверних даних та їх подальше збереження у базі даних, розроблено алгоритми зберігання даних у базі даних у випадку навантаження на систему.

У статті [12] автори розробили алгоритм аналізу соціальних мереж підлітків та нейронну мережу, яка визначає основні параметри підлітків і допомагає оптимізувати їх спілкування. Здійснено вирішення низки проблем, але проблема аналізу тексту та визначення емоцій користувача досі актуальна, для аналізу інформації використовуються як нейронні мережі, що шукають певні дані в тексті за заданими параметрами, так і алгоритми лінійного пошуку інформації, призначені для опрацювання великої кількості текстової та аналізу всієї інформації, навіть частково пошкодженої.

Проблема аналізу даних та подальшої ефективності опрацювання інформації дуже важлива для створення інформаційної системи і розглянута у статті [13]. В сучасному світі всі програмні системи працюють із різними даними, а результатом опрацювання є сукупність різних даних, які проаналізовані за допомогою або певних алгоритмів, або нейронних мереж. Важливим завданням є опрацювання, зберігання та використання отриманих даних, необхідно комбінувати використання постійних та динамічних даних, що дасть змогу оптимізувати роботу програмного продукту та збільшити ефективність роботи з даними. Розроблено механізм визначення основних робіт користувача із використанням мобільного пристрою за допомогою згорткової нейронної мережі, яка аналізує роботу користувача з мобільним пристроєм та формує основну інформацію про задану вибірку користувачів.

Автор статті [14] розглянув нейронну мережу, яка дає змогу визначити приблизний вік користувача соціальної мережі Instagram, здійснюючи пошук його обличчя на фотографіях та опрацюючи фотографії. Запропоновано механізм аналізу медіа-даних, а саме фільмів та серіалів із використанням нейронної мережі для визначення мінімального віку, з якого дозволено перегляд медіа-контенту, оптимізовано алгоритм для медіа-файлів низької якості. Висвітлено технологію

аналізу голосових повідомлень користувачів соціальних мереж, розроблено нейронну мережу, яка визначає емоції користувача, аналізуючи дані голосових повідомлень. Розглянуто механізм збереження та опрацювання користувацької історії пошуків товарів в інтернет-магазині, розроблено механізм пропонування найрелевантніших товарів користувачу. Розроблено механізм соціалізації користувачів у соціальній мережі Facebook із використанням алгоритму Левенштейна та розроблено механізм асинхронного зберігання даних. Висвітлено оптимізований алгоритм нечіткого пошуку певної інформації, що здійснює пошук даних та аналіз тексту, розроблено оптимізований механізм асинхронного аналізу тексту. Незважаючи на вирішені проблеми, досі дуже важливо оптимізувати роботу системних компонентів та створювати нові алгоритми для опрацювання і подальшого аналізу отриманих даних, що зробить програмні системи гнучкішими та доступнішими.

Використання алгоритмів пошуку інформації у текстових даних розглянули автори в статті [15]. Важливим є використання алгоритмів Левенштейна та нечіткого пошуку, також ці алгоритми є гнучкими, відповідно є можливість створювати окрему реалізацію алгоритму відповідно до поставленого завдання. Розглянуто реалізацію сучасних механізмів пошуку інформації в соціальних мережах, наведено інноваційні алгоритми, які потребують оптимізації для швидшого пошуку та опрацювання даних, викладено алгоритми оптимізованого збереження даних. Висвітлено механізми аналізу користувацьких запитів у інтернет-магазинах, створено алгоритм формування даних, отриманих у результаті аналізу запитів, та використання сформованих даних для збереження інформації про певного користувача. Висвітлено реалізацію алгоритмів асинхронного динамічного розрахунку та опрацювання даних, розроблено методику часткового збереження даних, яка дає змогу уникнути повторного перезавантаження вебсторінки та дає змогу створити динамічнішу інформаційну систему. Розглянуто реалізацію механізмів соціалізації користувачів із застосуванням моделі знерухомлення інформації, що дає змогу оптимізувати аналіз даних та здійснювати аналіз декількома потоками, оптимізовано алгоритм Левенштейна для соціалізації користувачів.

У статті [16] автор подав механізми аналізу користувацьких запитів у інтернет-магазинах, розглянув реалізацію алгоритму нечіткого пошуку серед спотворених даних, що дає змогу відновити початкову втрачену інформацію, вибрати параметри даних та виконати аналіз тексту. Висвітлено реалізацію бінарних алгоритмів для пошуку даних у соціальних мережах, визначено основні аспекти, необхідні для оптимізації пошуку даних та гарантування цілісності переданих даних за навантаження на систему, що дасть змогу захистити дані. В результаті можна сказати, що досі існує проблема оптимізації алгоритмів опрацювання текстових даних, оскільки кожна програмна система унікальна, і для кожної системи необхідно створювати унікальний алгоритм і оптимізувати його відповідно до особливостей системи.

Важливість використання машинного навчання у сучасних інформаційних системах висвітлив автор статті [17]. В наш час всі системи оперують великою кількістю даних, часто необхідно здійснити розподіл даних, що дасть змогу оптимізувати та пришвидшити роботу системних процесів. Цього можна досягти, опрацювавши частину даних із використанням машинного навчання. Висвітлено особливості створення серверних програм, оптимізовано алгоритми взаємодії серверних програм, які належать різним хостам, створено алгоритм оптимізованого пакетного передавання даних у вигляді спеціальних об'єктів та створено алгоритми зберігання отриманих об'єктів. Розглянуто реалізацію алгоритмів аналізу текстових даних у соціальних мережах, створено нейронну мережу, яка порівнює фотографії користувача та текстову інформацію, яку залишив користувач, і на основі проаналізованих даних формує вибірку, визначає тип та основні параметри користувача системи, оптимізовано роботу алгоритмів пошуку.

Автор статті [18] висвітлив основні алгоритми кластеризації даних, збережених на мобільних пристроях, сформовано основні шаблони збереження, опрацювання та формування нової вибірки даних із використанням поточної. Оптимізовано алгоритми пошуку інформації, які дають змогу уникнути нерелевантних даних та сформувати корисну вибірку.

Використання SEO-технологій у програмних системах, що передбачають взаємодію користувачів, показали автори статті [19]. У наш час SEO-технології важливі для оптимізації алгоритмів пошуку та соціалізації користувачів, правильно реалізований алгоритм дає можливість ефективно використовувати всі наявні можливості інформаційної системи для пошуку та взаємодії користувачів між собою, що дає змогу оптимізувати основні процеси системи. Розглянуто реалізацію лінійного алгоритму аналізу текстових даних користувачів, створено алгоритм побудови інформаційних діаграм, який дає змогу використовувати задані параметри стосовно опрацювання даних, оптимізовано алгоритми нечіткого пошуку для точнішого пошуку релевантної інформації в заданому тексті. Висвітлено реалізацію механізмів видобування даних, створено нейронну мережу, яка аналізує на вході всю соціальну інформацію та з використанням вхідних параметрів видаляє всю нерелевантну інформацію, зберігаючи лише необхідну в певний момент вибірку даних. Показано реалізацію механізмів зберігання даних із соціальних мереж у мобільних пристроях, оптимізовано зберігання даних на серверах системи та в локальній базі даних пристрою, оптимізовано алгоритми передавання та ефективного зберігання даних. Автори також запропонували декілька нових алгоритмів для аналізу соціальних мереж користувачів, це і нейронні мережі, які дають змогу аналізувати фотографії та аудіозаписи користувачів, і алгоритми аналізу користувацьких повідомлень, надзвичайно важливі в контексті створення глобальних вебзастосунків, оскільки в такі моменти однією із найважливіших проблем стає автоматизація опрацювання користувацької інформації.

Зробивши певні висновки, можна сказати, що завдання використання SEO-технологій важливе, оскільки існує безліч алгоритмів, які охарактеризували автори в статті [20–27], які необхідно оптимізувати для ефективного використання інформаційних систем соціалізації користувачів. Постає також завдання створення нових ефективних алгоритмів, які дадуть можливість покращити роботу сучасних інформаційних систем та здійснювати ефективне опрацювання даних.

#### **Формулювання мети та постановка задачі**

Основна мета – здійснення системного аналізу бажаної інформаційної системи з використанням функціональних діаграм IDEF0 та IDEF3, побудова дерева цілей та відповідне подальше створення інформаційної системи соціалізації груп користувачів за спільними інтересами. Варто зазначити, що ця мета є глобальною і в межах однієї мети важко описати всю суть дослідження, відповідно потрібно здійснити декомпозицію мети на декілька менших завдань, які необхідно вирішити для реалізації інформаційної системи. Якщо брати алгоритмічну складову, то варто виділити два основні завдання. Це реалізація згорткової нейронної мережі для визначення справжності користувачів та реалізація алгоритмів нечіткого пошуку для опрацювання текстових даних різних обсягів для вирішення основного завдання системи, а саме аналізу інформації про кожного користувача, формування певного рейтингу користувача, порівняння користувача з іншими користувачами для спрощення подальшої соціалізації користувачів, інтереси яких збігаються найбільше.

Завдання створення нейронної мережі, яка справно працює, важливе, оскільки одним із основних аспектів системи є взаємодія користувачів. Взаємодіючи між собою, користувачі повинні бачити один одного та знати, як виглядає співрозмовник, отже, необхідно бути впевненим, що на фотографії реальний користувач, а не шахрай. Щоб це здійснити, необхідно реалізувати згорткову нейронну мережу, яка найкраще справляється з такими завданнями, а саме з пошуком обличчя користувача на фотографії та порівнянням його із уже наявними в базі даних на момент аналізу.

Ще одне важливе завдання – початкове опрацювання тексту, оскільки необхідно здійснити правильний поділ на групи, які надалі проаналізують алгоритми нечіткого пошуку. Важливим завданням є така реалізація зберігання та опрацювання тексту, щоб оптимізувати роботу з базою даних, де весь текст зберігатиметься. Це необхідно, щоб не перевантажувати роботу з базою даних важкими запитами і намагатися тримати використовуваний текст локально в клієнті користувача і звертатися до бази лише тоді, коли це потрібно, і видобувати лише ті дані, які необхідні в певний

момент часу. Будь-яка оптимізація важлива, оскільки це допоможе пришвидшити роботу програми і зробити її максимально зручною для користувача.

Також важливим завданням є робота із фотографіями, оскільки створювана інформаційна система по суті є соціальною мережею, завантажені фотографії є важливою її частиною. Користувачі можуть завантажувати безліч фотографій, деякі великого розміру, це все зберігати в базі даних недоцільно і застаріло, оскільки медіа-файли великих розмірів не оптимальні для зберігання у базі даних і важко оптимізуються для зберігання на сервері, особливо коли кількість таких файлів може постійно збільшуватися. Тому важливим завданням є реалізація зберігання медіа-файлів у хмарному сховищі із використанням публічного інтерфейсу доступу до сховища, що дасть змогу зберігати в базі даних лише посилання та ідентифікатори фотографій, завдяки чому робота з базою буде максимально зручною, легкою та швидкою і не потрібно буде думати та працювати над оптимізацією зберігання файлів великих розмірів. Це дасть змогу витратити більше часу на вирішення важливіших завдань, таких як асинхронна реалізація інформаційної системи, під чим мають на увазі реалізацію більшої частини інформаційної системи, як серверної, так і клієнтської, із використанням асинхронних методів та засобів опрацювання даних.

Для описання інформаційної системи сформульовано основну мету та аспекти її функціонування у вигляді метрик за допомогою дерева цілей. Під головною метою системи в цьому випадку розуміють “Інформаційну систему соціалізації особистостей за спільними інтересами (G)”, оскільки інформаційна система призначена для зберігання та аналізу даних про всіх користувачів системи, опрацювання внесеної ними інформації, порівняння інтересів користувачів та їх подальшої соціалізації. Основна мета надає лише загальне розуміння того, що має виконувати інформаційна система, для детальнішого пояснення роботи інформаційної системи необхідно конкретизувати основні аспекти системи. До першого рівня належать чотири основні аспекти: “Передавання даних (A1)”, “Аналіз даних (A2)”, “Збереження даних (A3)” та “Формування даних (A4)”. У цьому випадку наведено всі аспекти першого рівня деталізації дерева цілей, що дає певне поняття про основний алгоритм роботи систем, але цього все ще недостатньо для повноцінного формулювання особливостей системи. До другого рівня належать: “Структурованість (B1)”, “Функціональність (B2)”, “Захищеність (B3)”, “Цілісність (B4)” та “Масштабованість (B5)”. Відповідно до третього рівня належать критерії, що є дочірніми до другого рівня: “Структурованість” складається із “Шаблонність (B1.1)” та “Ієрархічність (B1.2)”, “Функціональність” складається із “Опрацьованість (B2.1)” та “Представленість (B2.2)”, “Захищеність” – із “Відновлюваність (B3.1)” та “Резервність (B3.2)”, “Цілісність” – із “Синхронізованість (B4.1)” та “Залежність (B4.2)”, “Масштабованість” – із “Адаптивність (B5.1)” та “Мобільність (B5.2)”. У цьому випадку четвертий рівень ієрархії буде останнім, він містить дві альтернативи: інформаційно-аналітична система та інформаційно-довідкова система, позначені на діаграмі як “Альтернатива 1 (C1)” та “Альтернатива 2 (C2)”. На рис. 1 зображено дерево цілей інформаційної системи, що містить чотири рівні ієрархії.

Наступним кроком є визначення пріоритетів ієрархічних елементів, під час оцінювання використано значення від 1 до 5, під 1 розуміють аналогічну важливість, 3 – слабе значення, 5 – сильне значення, 2 або 4 означають проміжні величини між сусідніми елементами.

У табл. 1 подано побудовану матрицю попарних порівнянь.

За допомогою алгоритму Томаса Сааті знайдено ваги критеріїв, значення яких такі: передавання даних – 0,357; аналіз даних – 0,389; збереження даних – 0,192; формування даних – 0,062. Наступний крок – знаходження найбільшого значення:  $\lambda_{max} = 0,357 \times (1 + 0,333 + 0,5 + 0,2) + 0,389 \times (2 + 0,333 + 2 + 1) + 0,192 \times (1 + 0,25 + 3 + 0,5) + 0,062 \times (3 + 5 + 1 + 4) = 2,07 + 0,725 + 0,912 + 0,806 = 4,213$ . Далі знайдено індекс узгодженості  $CI = (4,213 - 4)/3 = 0,071$  та обчислено індекс послідовності  $CR = 0,071/0,90 = 0,078$ . Варто зазначити, що 0,90 – значення, взяте із таблиці випадкових індексів, тому що потрібний нам поточний індекс дорівнює 4. Наступним кроком є побудова матриці попарних порівнянь для аналізу другого рівня ієрархії на основі першого рівня (табл. 2).



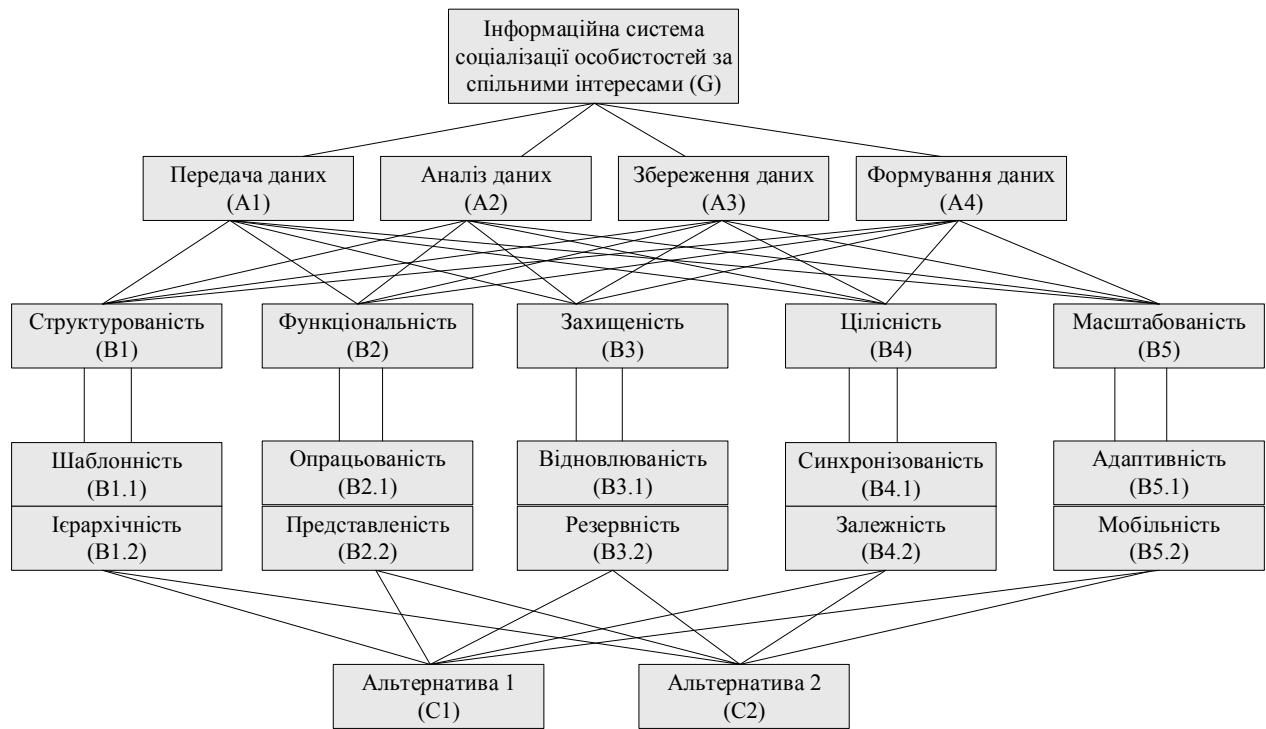


Рис. 1. Дерево цілей інформаційної системи

Таблиця 1

## Матриця попарних порівнянь мети системи

Інформаційна система	Передавання даних	Аналіз даних	Збереження даних	Формування даних
Передавання даних	1	2	3	5
Аналіз даних	0,5	1	0,5	3
Збереження даних	0,333	2	1	4
Формування даних	0,2	0,333	0,25	1

Таблиця 2

## Матриця попарних порівнянь критеріїв

Передавання даних	Структурованість	Функціональність	Захищеність	Цілісність	Масштабованість
1	2	3	4	5	6
Структурованість	1	3	2	0,5	0,333
Функціональність	0,333	1	1	0,333	0,2
Захищеність	0,5	1	1	0,5	0,2
Цілісність	2	3	2	1	1
Масштабованість	3	5	5	1	1
<b>Аналіз даних</b>	Структурованість	Функціональність	Захищеність	Цілісність	Масштабованість
Структурованість	1	4	2	3	1
Функціональність	0,25	1	1	0,5	0,2
Захищеність	0,5	1	1	0,333	0,2
Цілісність	0,333	2	3	1	0,333
Масштабованість	1	5	5	3	1
<b>Збереження даних</b>	Структурованість	Функціональність	Захищеність	Цілісність	Масштабованість
Структурованість	1	3	5	3	0,333
Функціональність	0,333	1	2	1	0,2
Захищеність	0,2	0,5	1	3	0,25
Цілісність	0,333	1	0,333	1	0,5
Масштабованість	3	5	4	2	1

Продовження табл. 2

1	2	3	4	5	6
<b>Формування даних</b>	Структурованість	Функціональність	Захищеність	Цілісність	Масштабованість
Структурованість	1	2	3	3	2
Функціональність	0,5	1	3	2	1
Захищеність	0,333	0,333	1	0,5	0,333
Цілісність	0,333	0,5	2	1	3
Масштабованість	0,5	1	3	0,333	1

Виконавши розрахунки взятих із матриць попарних порівнянь критеріїв оцінювання інформаційних систем, аналогічні до здійснених на першому рівні ієрархії, ми отримали значення пріоритетів для другого рівня ієрархії. Для передавання даних: структурованість – 0,164; функціональність – 0,182; захищеність – 0,255; цілісність – 0,089; масштабованість – 0,310. Для аналізу даних: структурованість – 0,164; функціональність – 0,182; захищеність – 0,255; цілісність – 0,089; масштабованість – 0,310. Для збереження даних: структурованість – 0,164; функціональність – 0,182; захищеність – 0,255; цілісність – 0,089; масштабованість – 0,310. Для формування даних: структурованість – 0,164; функціональність – 0,182; захищеність – 0,255; цілісність – 0,089; масштабованість – 0,310.

Далі здійснено підрахунок глобальних пріоритетів відповідно до критеріїв другого рівня ієрархії та їх підкритеріїв, отриманих із третього рівня. Вибрано основні критерії: структурованість, функціональність, захищеність, цілісність та масштабованість. Матрицю попарних порівнянь для третього рівня ієрархії відносно другого рівня щодо критеріїв подано в табл. 3.

Таблиця 3

**Матриця попарних порівнянь критерію “Структурованість”**

Структурованість	Шаблонність	Ієрархічність
Шаблонність	1	3
Ієрархічність	0,333	1
<b>Функціональність</b>	Опрацьованість	Представленість
Опрацьованість	1	2
Представленість	0,5	1
<b>Захищеність</b>	Відновлюваність	Резервність
Відновлюваність	1	0,2
Резервність	5	1
<b>Цілісність</b>	Синхронізованість	Залежність
Синхронізованість	1	5
Залежність	0,2	1
<b>Масштабованість</b>	Адаптивність	Мобільність
Адаптивність	1	3
Мобільність	0,333	1

Виконавши всі необхідні розрахунки, ми одержали значення глобальних пріоритетів (табл. 4). Потім здійснено перевірку узгодженості, знайдено значення усіх  $\lambda_{max}$ , CI, CR, як результат доведено, що всі значення CR не перевищують обмеження – 10 % (табл. 4).

Останнім кроком була побудова матриць відповідно до розміру наявних альтернатив, де здійснюється порівняння на основі критеріїв третього рівня. Отримані матриці наведено в табл. 5.

Проаналізувавши матриці попарних порівнянь критеріїв останнього рівня ієрархії щодо двох поточних наявних критеріїв інформаційної системи, ми отримали значення ваг поточних альтернатив до кожного критерію, відтак виконали підрахунок глобальних ваг як суми добутків даних глобальних значень другого рівня ієрархії, а саме: структурованість, функціональність, захищеність

ність, цілісність та масштабованість, на значення поточних критеріїв, отриманих з третього рівня ієрархії, а саме: шаблонність, ієрархічність, опрацьованість, представленість, відновлюваність, резервність, синхронізованість, залежність, адаптивність, мобільність.

Таблиця 4

**Отримані глобальні пріоритети та індекси**

Критерій	Порівняння	Значення	$\lambda_{\max}$	CI	CR
Структурованість (0,188)	Шаблонність	0,090	2,034	0,055	0,071
	Ієрархічність	0,098			
Функціональність (0,092)	Опрацьованість	0,029	2,065	0,045	0,065
	Представленість	0,063			
Захищеність (0,322)	Відновлюваність	0,277	2,080	0,049	0,063
	Резервність	0,045			
Цілісність (0,121)	Синхронізованість	0,076	2,043	0,052	0,072
	Залежність	0,045			
Масштабованість (0,277)	Адаптивність	0,192	2,011	0,058	0,075
	Мобільність	0,085			

Таблиця 5

**Матриця попарних порівнянь критерію “Шаблонність”**

Критерій		A1	A2	Вага альтернативи
Шаблонність	A1	1	3	0,75
	A2	0,333	1	0,25
Ієрархічність	A1	1	0,25	0,2
	A2	4	1	0,8
Опрацьованість	A1	1	1	0,5
	A2	1	1	0,5
Представленість	A1	1	1	0,5
	A2	1	1	0,5
Відновлюваність	A1	1	0,25	0,2
	A2	4	1	0,8
Резервність	A1	1	5	0,83
	A2	0,2	1	0,17
Синхронізованість	A1	1	1	0,5
	A2	1	1	0,5
Залежність	A1	1	2	0,75
	A2	0,333	1	0,25
Адаптивність	A1	1	0,25	0,2
	A2	4	1	0,8
Мобільність	A1	1	3	0,75
	A2	0,333	1	0,25

Значення альтернатив такі: A1 (інформаційно-аналітична система) = 0,602, A2 (інформаційно-довідкова система) = 0,398. Як висновок можна сказати, що “Інформаційно-аналітична система” є найкращою альтернативою відповідно до всіх критеріїв.

**Виклад основного матеріалу**

Побудувавши дерево цілей, ми здійснили подальшу конкретизацію функцій системи, вибравши для деталізації структури системи функціональні діаграми IDEF0. Для узагальненого відображення головних функцій розробленої інформаційної системи насамперед створено контекстну діаграму, яка складається із основних вхідних та вихідних даних, а також основних даних для

зображення механізмів та контролю. До вхідних даних належать: “Логін користувача”, “Пароль користувача”, “База даних”, “Фотографії користувача” та “Запит користувача”. До вихідних аспектів – “Соціалізація користувачів” та “Оновлена база даних”. Серед основних аспектів контролю треба виділити “Рівні доступу” та “Вимоги”, що впливають на процеси інформаційної системи. За виконання основних функцій інформаційної системи відповідають механізми системи, а саме: “Алгоритми взаємодії”, “Згорткова неймережа”, “Алгоритми нечіткого пошуку” та “Алгоритми збереження даних”. На рис. 2 зображено контекстну діаграму поточної інформаційної системи.

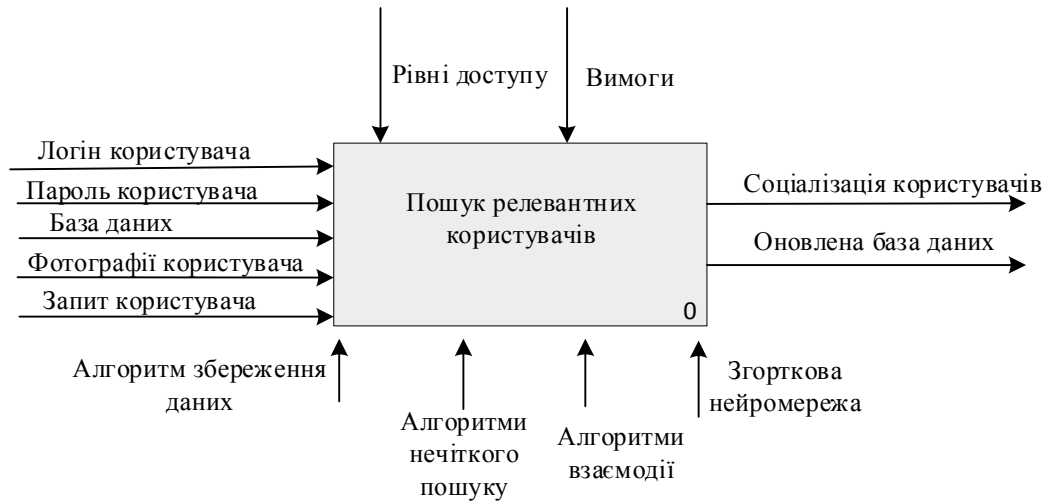


Рис. 2. Контекстна діаграма інформаційної системи

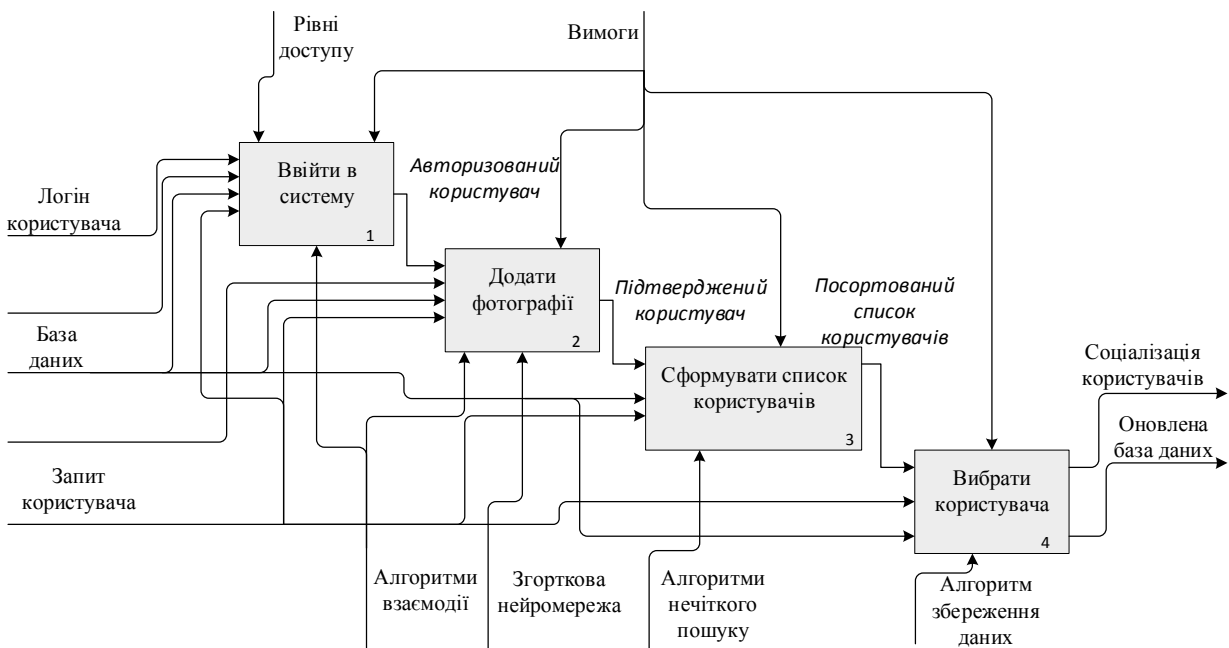


Рис. 3. Декомпозиція інформаційної системи

На рис. 3–8 наведено декомпозицію інформаційної системи та її блоків, на рис. 9 – декомпозицію блока “Авторизуватися”.

На рис. 10–12 подано декомпозицію блоків “Додати фотографію”, “Перевірити наявність обличчя на фотографії” та “Перевірити наявність обличчя у базі”, на рис. 13 – декомпозицію блока “Відформатувати параметри користувачів”.

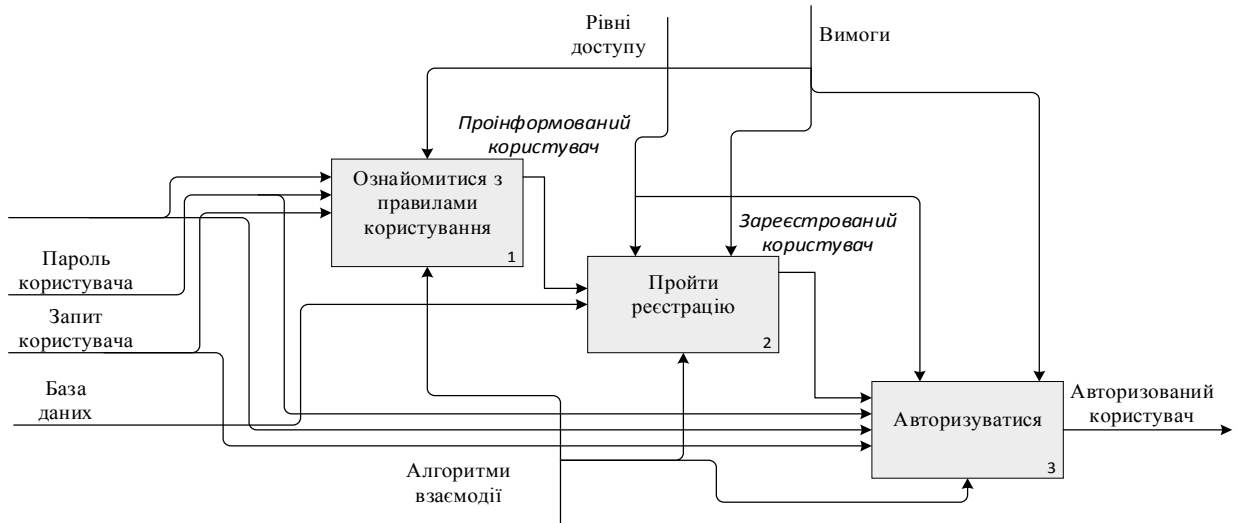


Рис. 4. Декомпозиція блока "Увійти в систему"



Рис. 5. Декомпозиція блока "Додати фотографії"

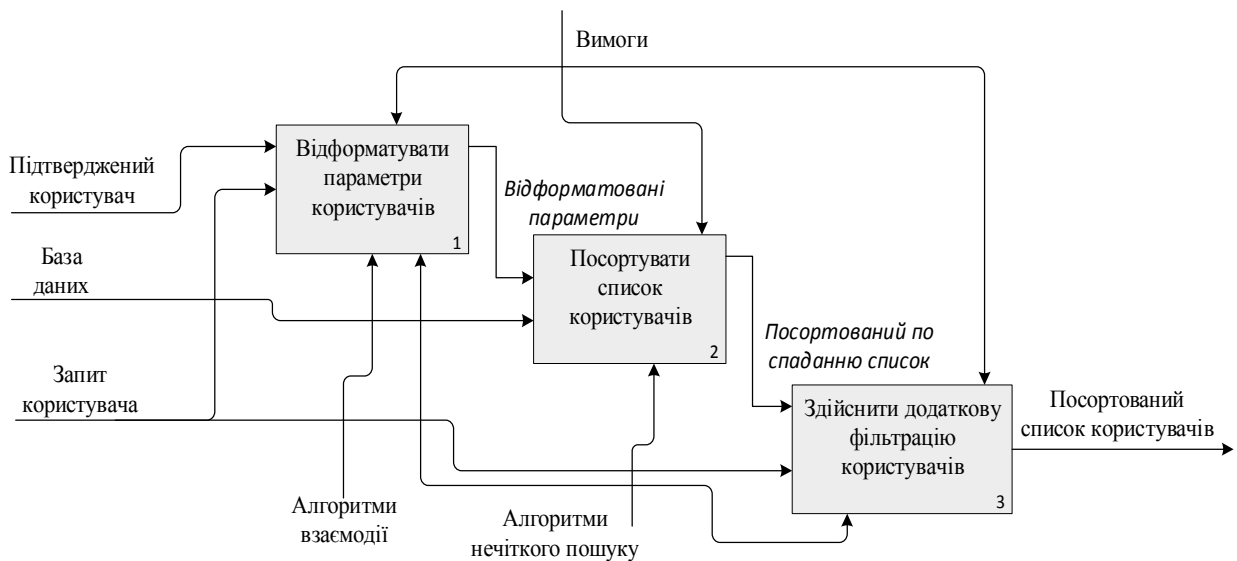


Рис. 6. Декомпозиція блока "Сформувати список користувачів"

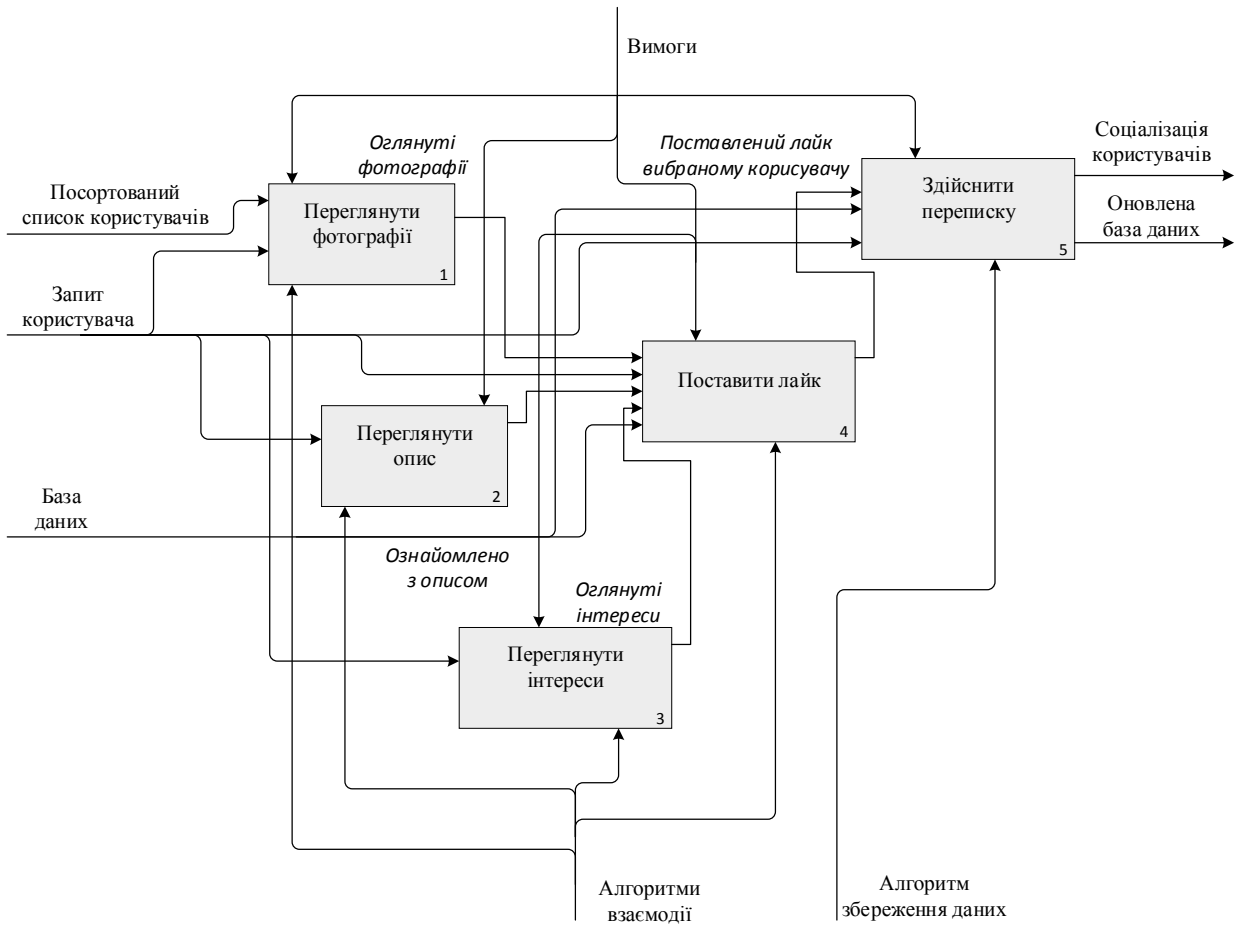


Рис. 7. Декомпозиція блока "Вибрати користувача"

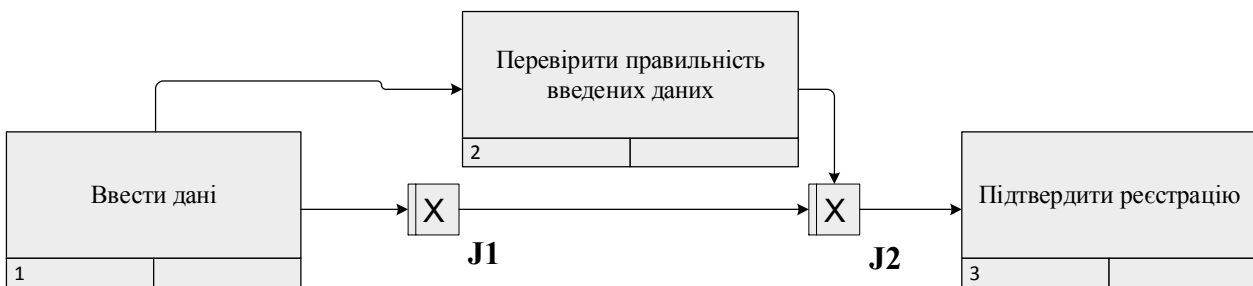


Рис. 8. Декомпозиція блока "Пройти реєстрацію"

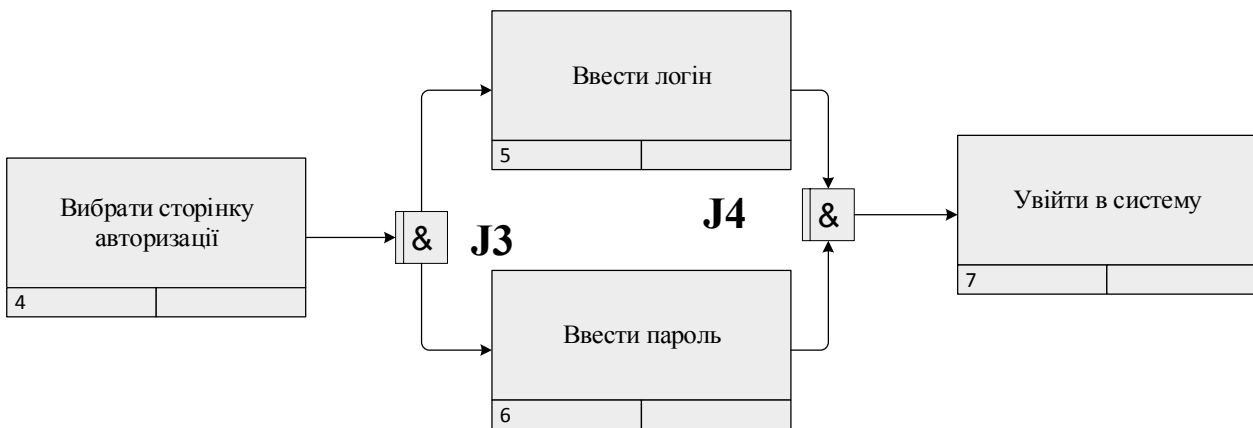


Рис. 9. Декомпозиція блока "Авторизуватися"

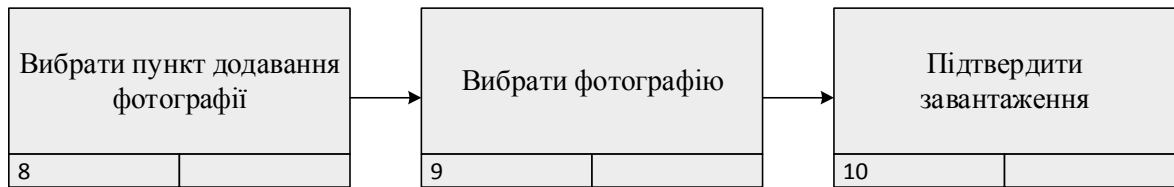


Рис. 10. Декомпозиція блока "Додати фотографію"

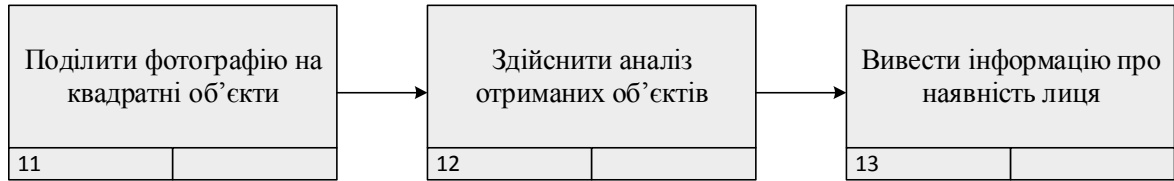


Рис. 11. Декомпозиція блока "Перевірити наявність лиця на фотографії"

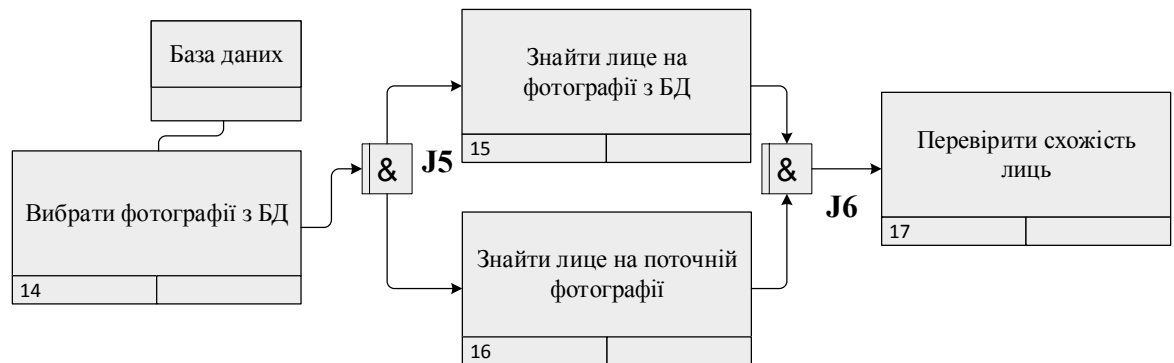


Рис. 12. Декомпозиція блока "Перевірити наявність лиця в базі"

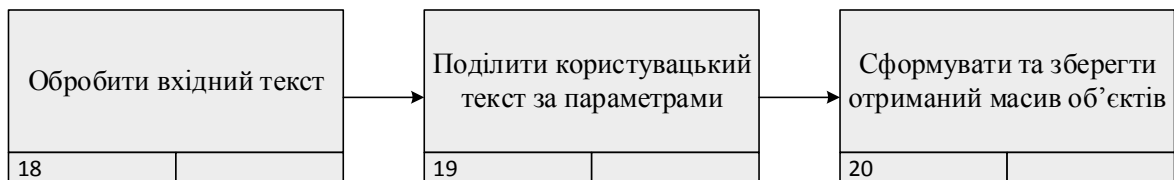


Рис. 13. Декомпозиція блока "Відформатувати параметри користувачів"

На рис. 14–15 подано декомпозицію блока "Посортувати список користувачів" та "Здійснити додаткову фільтрацію користувачів", на рис. 16 – декомпозицію блока "Поставити лайк".

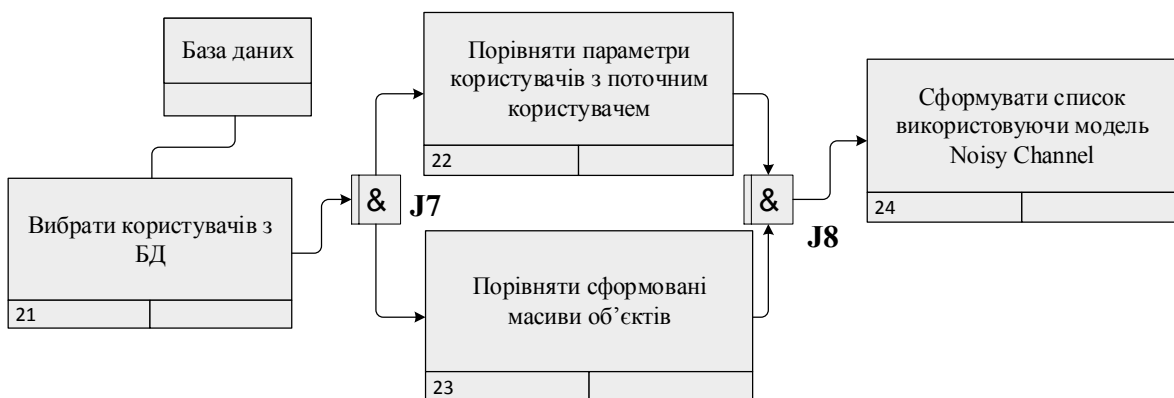


Рис. 14. Декомпозиція блока "Посортувати список користувачів"



Рис. 15. Декомпозиція блока “Здійснити додаткову фільтрацію користувачів”

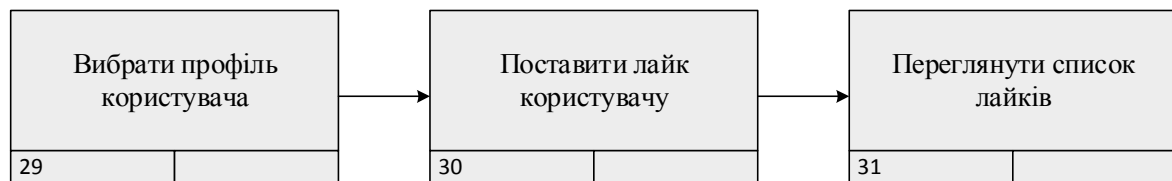


Рис. 16. Декомпозиція блока “Поставити лайк”

На рис. 17 зображено декомпозицію блока “Здійснити листування”.

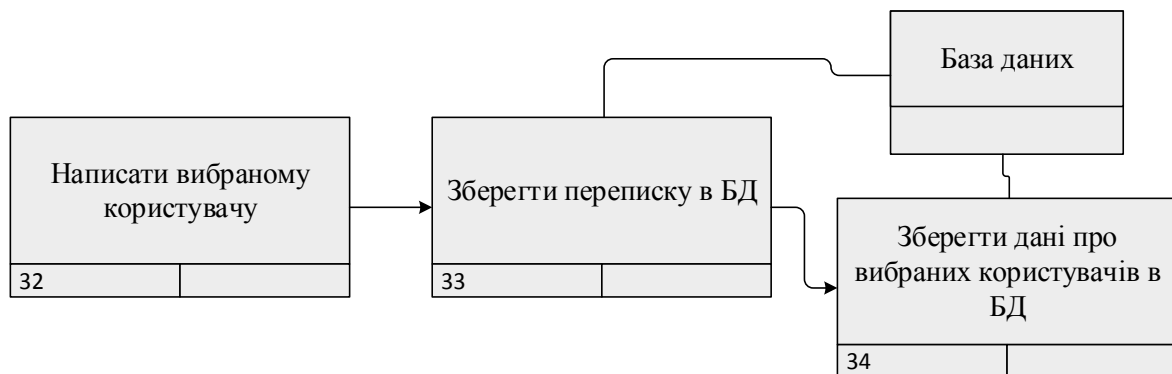


Рис. 17. Декомпозиція блока “Здійснити листування”

Для відображення чіткої структури інформаційної системи та ієрархії процесів використано деревоподібну структуру відображення, оскільки її найлегше подати і вона зрозуміла для відображення чіткої ієрархії елементів системи. Основним кореневим процесом інформаційної системи є “Пошук релевантних користувачів”, оскільки це найузагальненіше пояснення усіх процесів, що відбуваються упродовж роботи системи. Кореневий елемент поділено на чотири конкретні елементи, а саме “Ввійти в систему”, “Додати фотографії”, “Сформувати список користувачів” та “Вибрати користувача”. Кожен із цих чотирьох процесів виконує свої унікальні функції, приймаючи лише певні дані на вхід та віддаючи сформовані вибірки даних на вихід. Якщо продовжити ієрархію процесів, то кожен із цих процесів складається із певної кількості елементів.

Процес “Ввійти в систему” складається із основних елементів, які відповідають за початкову взаємодію користувача із системою: “Ознайомитися з правилами користування”, “Пройти реєстрацію”, “Авторизуватися”. Процес “Додати фотографії” містить підпроцеси, що відповідають за додавання та опрацювання фотографій користувача: “Додати фотографію”, “Перевірити наявність обличчя на фотографії”, “Перевірити наявність обличчя в базі”. Наступний важливий процес “Сформувати список користувачів” містить підпроцеси, що відповідають за підбір користувачів інформаційної системи: “Відформатувати параметри користувачів”, “Посортувати список користувачів” та “Здійснити додаткову фільтрацію користувачів”. У процес “Вибрати користувача” входять



підпроцеси, що відповідають за соціалізацію користувачів: “Переглянути фотографії”, “Переглянути опис”, “Переглянути інтереси”, “Поставити лайк” та “Здійснити листування”.

Ієрархію процесів наведено на рис. 18, на рис. 19 – головне вікно програми.

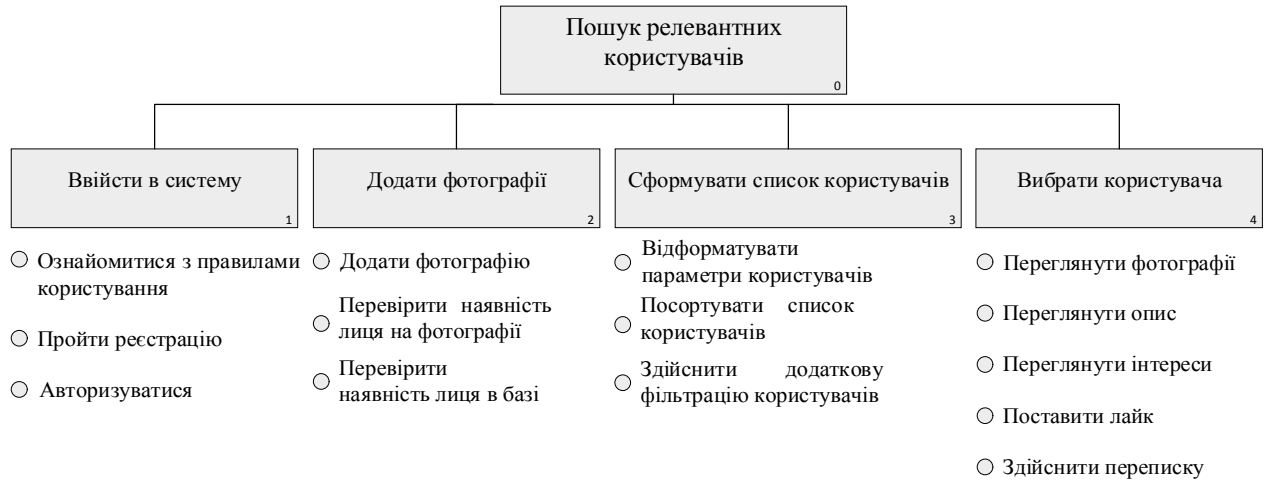


Рис. 18. Ієрархія процесів інформаційної системи

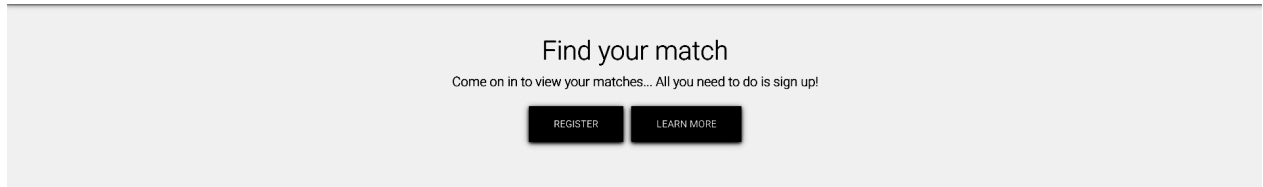


Рис. 19. Головне вікно програми

Якщо описувати інформаційну систему соціалізації особистостей за спільними інтересами, то варто почати із реалізованих механізмів реєстрації та авторизації користувачів системи. На рис. 20, а зображено кнопки головного вікна програми, на рис. 20, б – форму реєстрації користувача.

а

б

Рис. 20. Кнопки головного вікна програми (а) та форма для реєстрації користувача (б)

Реєстрацію та авторизацію користувачів реалізовано із використанням механізму Identity, який надає API для створення зручної інформаційної системи роботи із профілями різних користувачів. Для реєстрації користувачів використано метод Register, який відповідно виконує функції створення користувацького профілю. На рис. 21, а подано авторизацію користувача, введення логіну та паролю, на рис. 21, б – повідомлення про успішну авторизацію.



Рис. 21. Авторизація користувача (а) та повідомлення про успішну авторизацію (б)

На вхід метод приймає два необхідні параметри – змінну типу User, яка містить в собі всі властивості, які вніс користувач для збереження у базі даних, і змінну password – пароль, який ввів користувач. Для збереження пароля користувача використовується шифрування, для реалізації шифрування – дві змінні, які мають тип масиву байтів: passwordHash та passwordSalt. Серед усіх шифрувань вибрано HMACSHA512, це сучасний криптографічний механізм перевірки цілісності інформації, який використовує хеш-функції та гарантує цілісність інформації. В ході виконання алгоритму створюється новий об'єкт, ключ із якого присвоюється змінній passwordSalt, а підрахований хеш переданого пароля у passwordHash. Також варто зазначити, що passwordSalt – це спеціальний рядок, який містить дані, що передаються одночасно із вхідним масивом даних і є образом для розрахунку хешу. Повернувши із методу створені passwordSalt та passwordHash, алгоритм записує їх у переданий раніше в метод об'єкт типу User та асинхронно записує нового користувача в базу даних, асинхронно здійснює оновлення бази даних і збереження змін і в кінці повертає створеного користувача. Також перед створенням користувача перевіряють вже наявність такого користувача методом UserExists. Авторизація користувача відбувається доволі легко на серверному рівні та має певні особливості в клієнтській частині інформаційної системи. На серверній частині в метод Login передаються логін та пароль користувача, здійснюється перевірка, чи існує користувач, який має такий логін і пароль у базі даних; якщо такого користувача не існує, то повертається null. Якщо користувач існує, тоді здійснюється порівняння хешів паролів, за допомогою методу VerifyPasswordHash, в який передається поточний пароль користувача, та passwordHash і passwordSalt, створені в методі користувача, взятого з бази даних. Якщо хеші збігаються, то з методу повертається об'єкт користувача. Після цього здійснюється авторизація за допомогою JWT – Json Web Token. Це порівняно нова технологія, яка дає змогу зручно здійснювати авторизацію користувача, зберігаючи частину токена в базі даних браузера. Спочатку для користувача визначають необхідні дозволи відповідно до рівня доступу за допомогою властивостей ClaimTypes.NameIdentifier та ClaimTypes.Name. Далі із захищеного файлу дістають токен, який є ключем шифрування, і шифрують симетричний ключ доступу. Відтак видають основні властивості користувача за допомогою створення окремої сигнатури. Наступним кроком є створення спеціального дескриптора токена – це об'єкт, який містить в собі всі дозволи для поточного користувача, час дії та валідності поточного токена для користувача та властивості користувача із сигнатурою. Далі здійснюється створення об'єкта типу JwtSecurityTokenHandler, який відповідає за генерацію вже самого токена.

Відтак генеруємо токен, використавши дані, які містяться у створеному дескрипторі. Останнім кроком буде повернення з методу анонімного об'єкта, який містить у собі створений токен та об'єкт із усіма властивостями поточного користувача. В клієнті приймається запит з сервера та за допомогою директиви AuthGuard здійснюється перевірка валідності частини токена, призначеного для зберігання всередині локальної бази даних браузера, та відбувається асинхронне збереження токена. Всі дані зберігаються в таблицях у реляційній базі даних, всі наявні дані взаємопов'язані за допомогою зв'язків. Схему поточної бази даних подано на рис. 22.

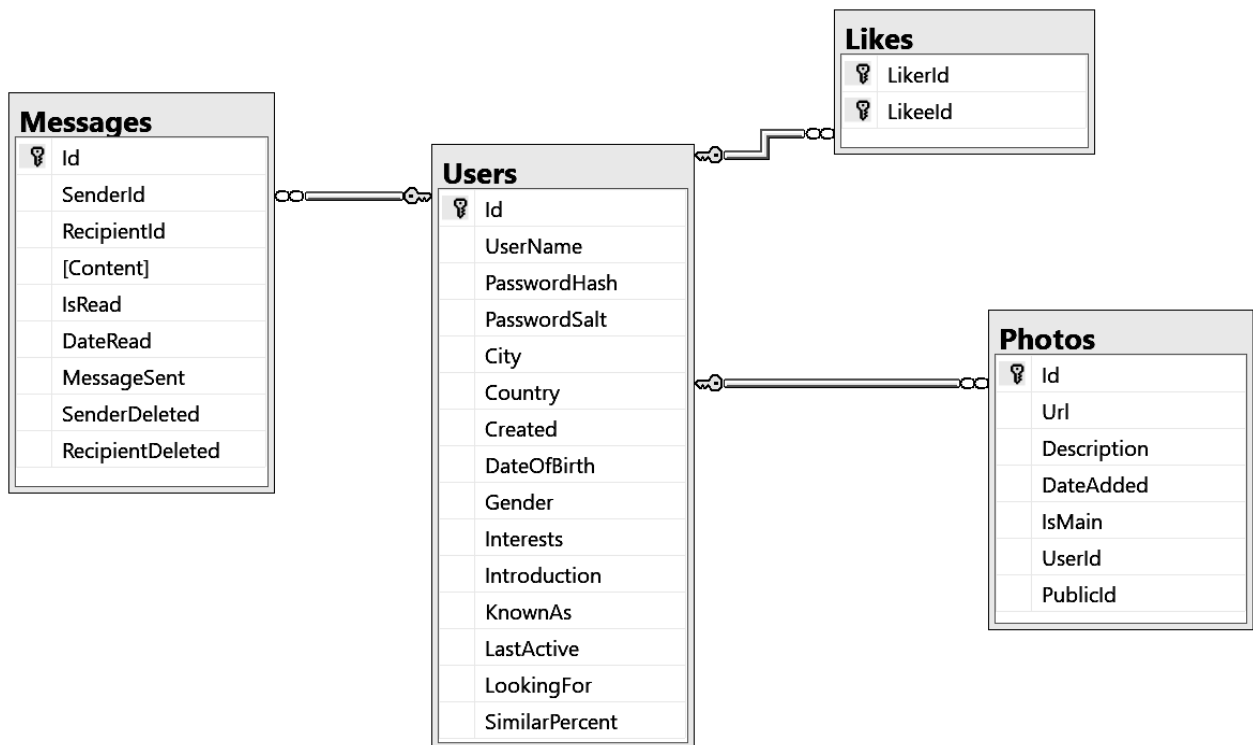


Рис. 22. Схема бази даних інформаційної системи

Насамперед варто описати ключову таблицю бази даних – Users, яка представляє поточного користувача системи. Вона має такі основні поля: Id – унікальний ідентифікатор користувача, UserName – логін користувача; PasswordHash – хеш пароля користувача; PasswordSalt – образ хешу користувача; Gender – стать користувача; DateOfBirth – дата народження; KnownAs – нікнейм користувача всередині системи; Created – дата реєстрації; LastActive – дата останніх відвідин користувача; Introduction – вступна інформація про користувача; LookingFor – основна інформація про те, що саме користувачу потрібно і що він шукає; Interests – описання інтересів користувача системи; City – місто проживання; Country – країна проживання; SimilarPercent – тимчасовий параметр схожості користувача. Також, як видно зі схеми бази даних, таблиця Users пов’язана з іншими таблицями: Photos, Likes та Messages за допомогою зв’язку “один до багатьох”, відповідно Users є центральною таблицею. Далі йде таблиця Photos, в якій є всі основні властивості фотографій користувачів інформаційної системи, а саме Id – унікальний ідентифікатор; Url – посилання на картинку; Description – опис; DateAdded – дата додавання; IsMain – булеве поле, яке визначає аватарку користувача; PublicId – ідентифікатор знаходження картинки в хмарному сервісі зберігання даних та UserId – зовнішній ключ, який означає, що картинка належить одному із користувачів. Далі таблиця Likes, в якій містяться ідентифікатори користувачів, які виставляли лайки, а саме LikerId – унікальний ідентифікатор користувача, який поставив лайк, та LikeeId – унікальний ідентифікатор користувача, який одержав лайк у системі. Потім – таблиця Messages, в якій вказано всі основні властивості повідомлень, що надіслали користувачі інформаційної системи, а саме Id – унікальний ідентифікатор повідомлення; SenderId – унікальний ідентифікатор відправника; RecipientId – унікальний ідентифікатор отримувача; Content – вміст повідомлення; IsRead – позначка чи прочитане повідомлення; DateRead – дата прочитання; MessageSent – дата відправлення; SenderDeleted – чи видалив відправник; RecipientDeleted – чи видалив отримувач.

*Інструкція користувача.* Щоб зрозуміти, як користуватися ІС соціалізації особистостей за спільними інтересами, створено спеціальну інструкцію користувача в його профілі. На рис. 23, а подано параметри профілю користувача, на рис. 23, б – заповнений профіль користувача.

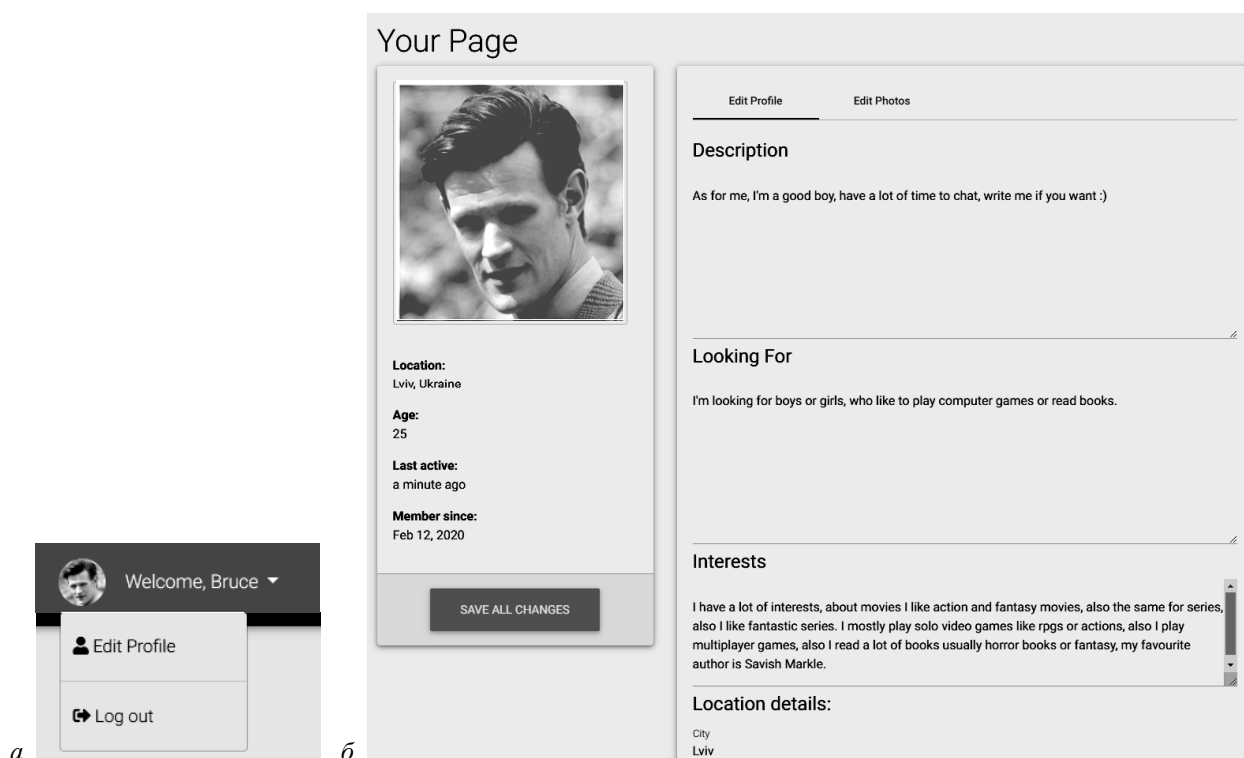


Рис. 23. Параметри користувацького профілю (а) та заповнений профіль користувача (б)

**Призначення документа.** Інформаційна система соціалізації особистостей призначена для знайомства та взаємодії користувачів системи, для цього використовуються нейронна мережа для ідентифікації користувачів та алгоритми аналізу тексту для підбирання користувачів.

**Короткий виклад основної частини документа.** ІС призначена для соціалізації користувачів, умовою застосування є реєстрація користувачів у системі, внесення даних і завантаження власної фотографії. На рис. 24 наведено процес завантаження фотографій у систему. Можна завантажувати одночасно одну і більше фотографій, перетягнувши їх вручну або за допомогою провідника. Призначення – соціалізація користувачів системи. Вхідними даними є інформація про користувача та користувацькі фотографії, вихідними даними, сформованими системою, – підбірка користувачів, посортована за зниженням відсоткового співвідношення схожості користувачів. Також програма дає можливість ставити лайк вибраному користувачу та здійснювати приватне листування всередині системи, оновлювати всі дані про себе та переглядати прочитані й відправлені повідомлення. На рис. 25 подано завантажені фотографії користувача. Можна видалити всі фотографії, крім поточної головної фотографії, та нейронні мережі, що опрацювали всі фотографії, і ті, на яких не знайдено обличчя, недоступні для виставлення основних фотографій користувача.

**Позначення і назви програми.** Повна назва: Інформаційна система соціалізації особистостей за спільними інтересами на основі SEO-технологій та методів машинного навчання. Скорочена назва: Інформаційна система соціалізації особистостей за спільними інтересами.

**Мови програмування, на яких написана програма.** C#, JavaScript, TypeScript, ASP .NET Core MVC, Angular.

**Призначення програми.** Соціалізація користувачів за допомогою формування відсоткового співвідношення схожості користувачів на основі вказаної інформації.

**Можливості програми.** Інформаційна система ідентифікує наявність обличчя на фотографії користувача, тим самим дає змогу зробити певну фотографію основною, здійснює соціалізацію користувачів. Під час формування списку користувачів надається можливість вибрати користувача, поставити лайк та вести приватне листування всередині інформаційної системи, а надалі переглянути всі повідомлення.

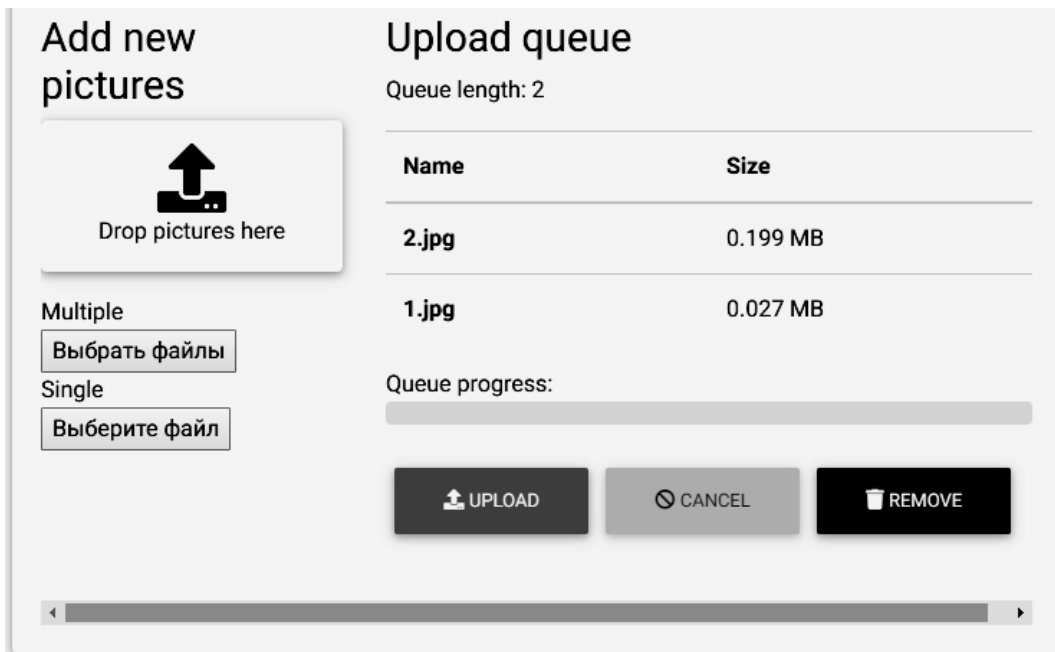


Рис. 24. Завантаження фотографій

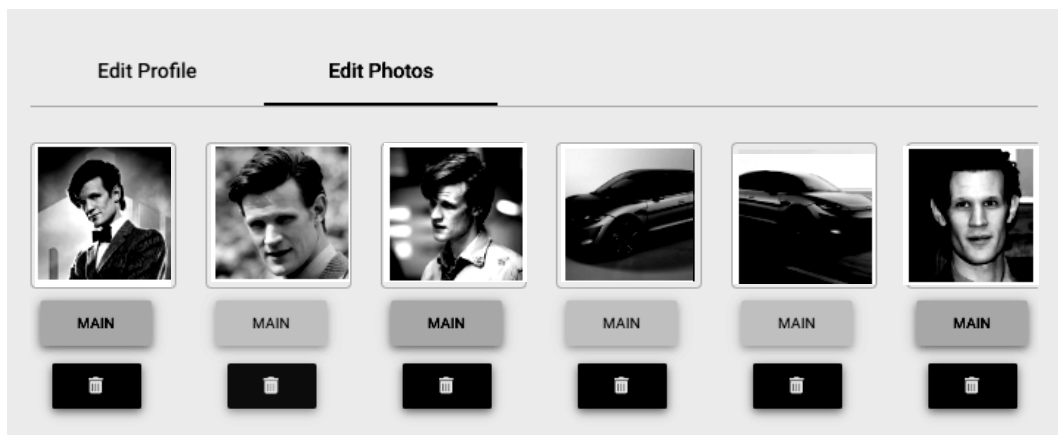


Рис. 25. Завантажені фотографії користувача

*Описання завдань:* завдання реєстрації та авторизації користувачів всередині системи; збереження інформації користувача; збереження фотографій користувача; аналіз фотографій користувача та пошук обличчя на фотографіях; зміна основної фотографії; формування списку користувачів і визначення відсоткового співвідношення. На рис. 26 подано список користувачів, сформований за допомогою алгоритмів опрацювання тексту та посортований за зниженням відсоткового співвідношення схожості користувачів. Завдання: збереження даних про вибраних користувачів; збереження приватних повідомлень та діалогів користувачів; збереження фотографій у хмарному сервісі; динамічне завантаження даних у профіль користувача.

*Методи вирішення завдань.* Для реєстрації та авторизації використано засоби Identity та JWT токени, збереження інформації відбувається всередині бази даних із використанням ORM функціоналу Entity Framework Core. Завдання: збереження фотографій за допомогою хмарного сервісу Cloudinary; аналіз фотографій за допомогою згорткової та сіамської нейронних мереж; зміна головної фотографії за допомогою внутрішніх функцій системи; формування списку користувачів і формування відсоткового співвідношення – виконуються за допомогою таких алгоритмів, як алгоритм Левенштейна, алгоритм розширення вибірки, алгоритм N-грам та моделі Noisy Channel. Для взаємодії користувачів використано методи утримання сесії, методи збереження приватних

повідомлень та методи збереження інформації про кожне повідомлення окремо, завдання динамічного завантаження даних виконуються за допомогою асинхронних методів та методів “реактивного” програмування.

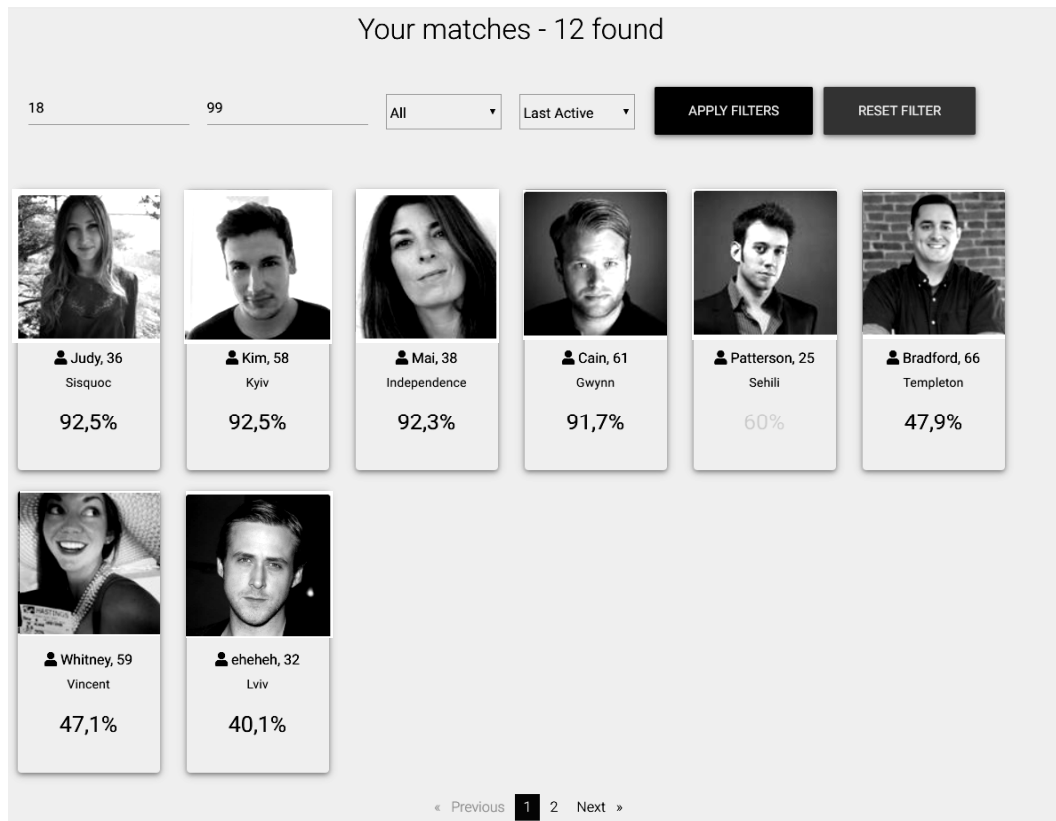


Рис. 26. Сформований список користувачів

*Функції, що виконує програма.* Register – основна функція для реєстрації користувачів; Login – функція для авторизації користувачів; GetUsers – функція для формування списку користувачів та виведення сформованого списку; GetUser – функція для виведення повної інформації про вибраного користувача; UpdateUser – функція для оновлення інформації користувача; LikeUser – функція для виставлення позначки вподобання вибраному користувачу; GetPhoto – функція для вибору окремої фотографії; AddPhotoForUser – функція пошуку обличчя та додавання фотографії; SetMainPhoto – функція встановлення головної фотографії користувача; DeletePhoto – функція видалення фотографії користувача; GetMessage – функція для вибору зі списку окремого повідомлення користувача; GetMessagesForUser – функція виведення надісланих, отриманих або непрочитаних повідомлень; GetMessageThread – функція виведення усіх повідомлень із діалогу між користувачами; CreateMessage – функція створення та збереження створеного повідомлення; DeleteMessage – функція видалення повідомлення; MarkMessageAsRead – функція позначення повідомлення як прочитаного. На рис. 27, а зображено використання фільтрів для пошуку у вже сформованому списку, на рис. 27, б – вибір профілю користувача, продемонстровано можливість переглянути профіль користувача, поставити лайк і написати приватне повідомлення.

*Часові характеристики.* Програма оптимізує всі виконувані завдання через те, що відсутні перезавантаження сторінки і всі запити асинхронні, аналіз фотографій нейронними мережами відбувається асинхронно, із використанням “реактивного” програмування, алгоритми аналізу тексту оптимізовано для опрацювання великих об’ємів даних. Сама система оптимізує соціалізацію користувачів, оскільки визначає відсоток схожості усіх користувачів і дає можливість вибрати потрібного користувача. На рис. 28, а, б зображено перегляд вкладки із інформацією про вподобання користувачів, які вибрали нас і яких вибрали ми.

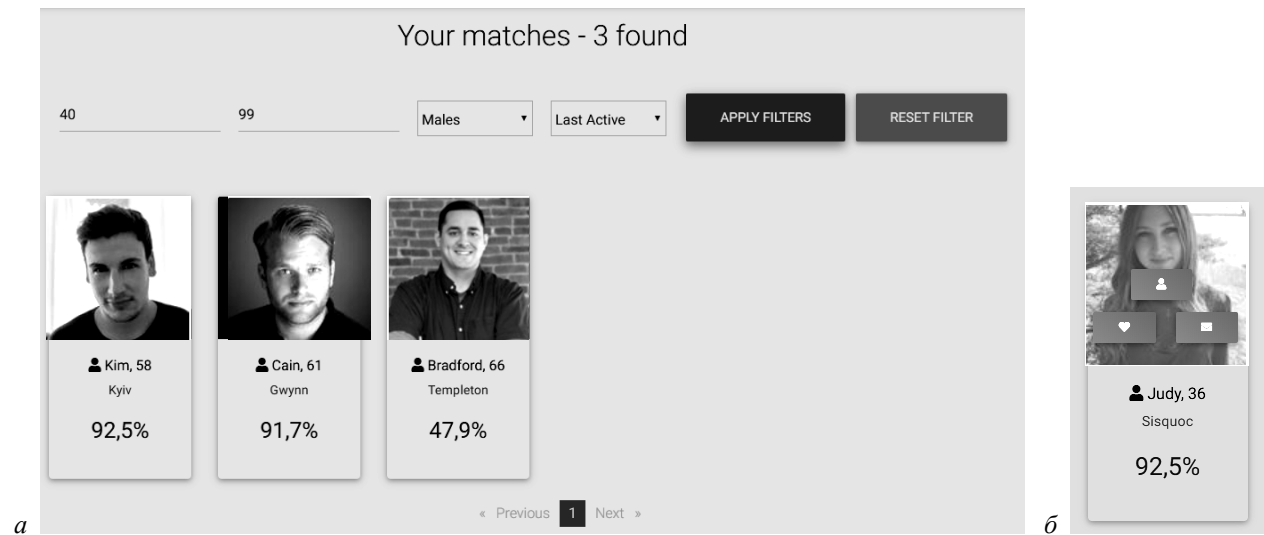


Рис. 27. Фільтрація списку (а) та вибір користувача (б)

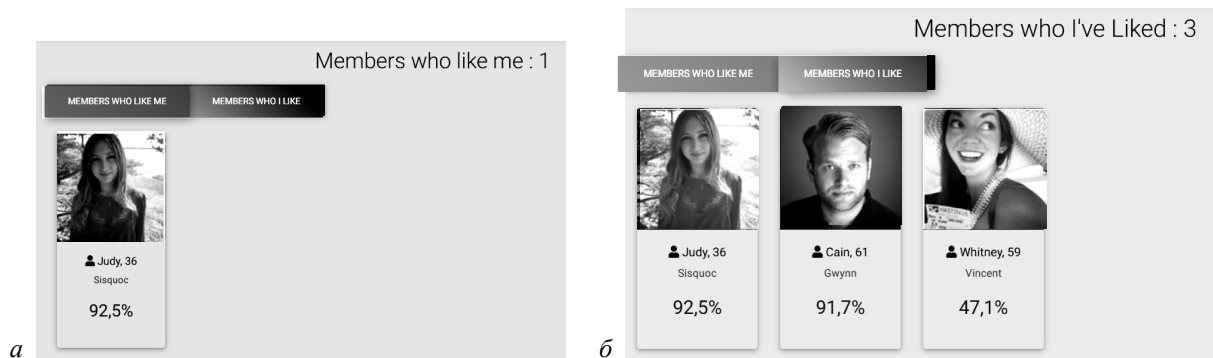


Рис. 28. Користувачі, які вибрали нас (а), та користувачі, яких вибрали ми (б)

**Режим роботи.** Систему створено у вигляді вебзастосунку. Відповідно, вона працює цілодобово і робота з нею починається у момент першого користувацького запиту всередині системи. Під час першого запиту та в момент реєстрації чи авторизації користувача за допомогою JWT токена створюється сесія на 24 год, всі проміжні дані, які завантажив або вибрав користувач, зберігаються, після закінчення сесії всі дані очищуються.

**Засоби контролю правильності виконання і самовідновлення програми.** В системі є два види помилок: помилки завантаження даних та помилки завантаження сторінок. Практично усі функції системи у разі помилки повертають або повідомлення про помилку, або повідомлення про виняткову ситуацію без збоїв у роботі всієї системи. У випадку помилки завантаження сторінки відбувається перенаправлення користувача на спеціальну вебсторінку із повідомленням про помилку.

**Обмеження сфери застосування програми.** Основні обмеження: програма не може гарантувати реальність користувача системи за фотографією; не може абсолютно точно визначити відсоткове співвідношення користувачів; програму не може використовувати користувач, який не авторизований у системі; система не може опрацьовувати дані в офлайн-режимі.

**Умови, необхідні для виконання програми.** Основні обмеження полягають у тому, що програмою може користуватися лише зареєстрований та авторизований у системі користувач, для формування вибірки користувачів. Необхідно, щоб користувач інформаційної системи мав свою фотографію, виставлену основною, заповнив усі дані, в користувача має бути доступ до інтернету. На рис. 29–30 подано основну інформацію профілю вибраного користувача, вкладку з інтересами користувача та вкладку з усіма фотографіями користувача.

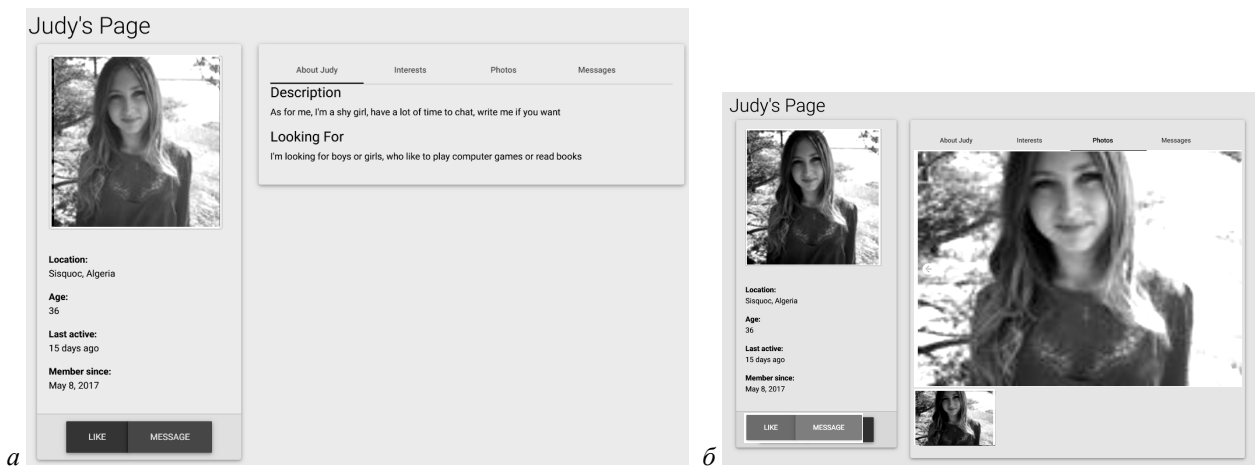


Рис. 29. Основна інформація про користувача (а) та фотографії користувача (б)

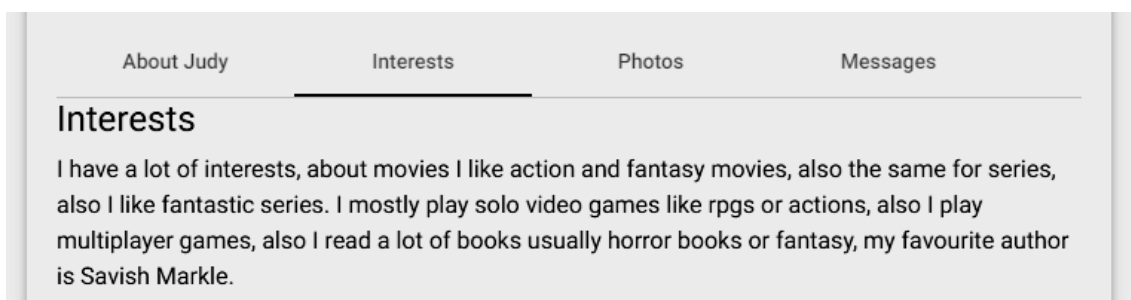


Рис. 30. Інтереси користувача

На рис. 31 зображено вкладку із приватним листуванням із користувачем. У листуванні наведено ніки користувачів, фотографії, час надсилання і час прочитання повідомлень.

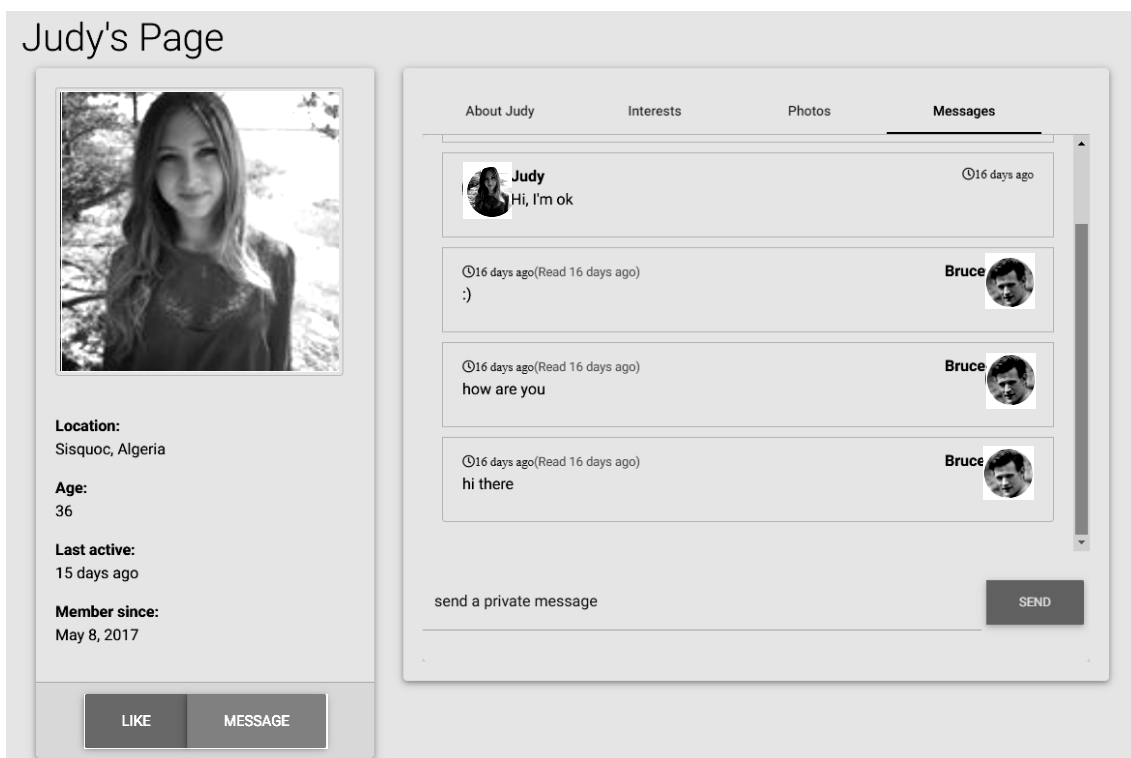


Рис. 31. Приватне листування із користувачем



На рис. 32 та 33 наведено сторінку із інформацією про всі повідомлення, непрочитані, отримані та надіслані. Можна здійснювати керування повідомленнями, а саме перегляд вибраного повідомлення за допомогою переходу в діалог з користувачем, або видалення вибраного свого повідомлення для всіх, або чужого повідомлення лише для себе.

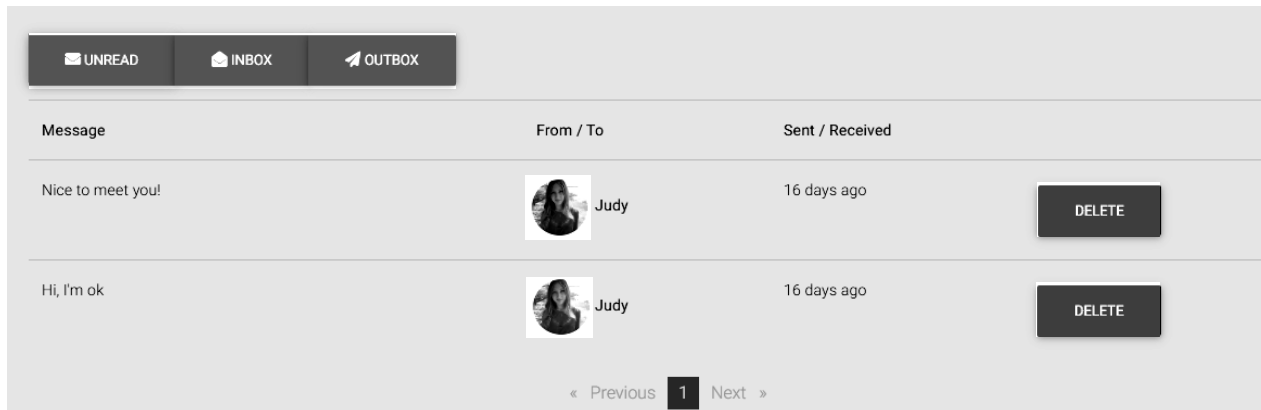


Рис. 32. Отримані повідомлення

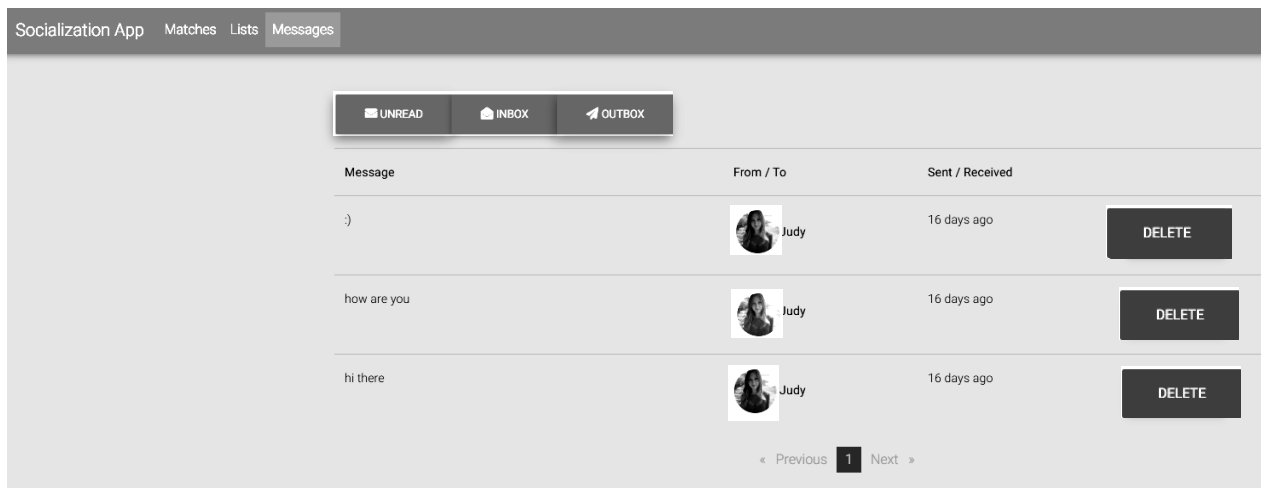


Рис. 33. Надіслані повідомлення

На рис. 34 зображено вхід із профілю іншого користувача, якого вибрано першим користувачем системи, та перегляд списку користувачів, які нас вибрали, що дає змогу почати приватне листування між двома користувачами, що вибрали один одного.

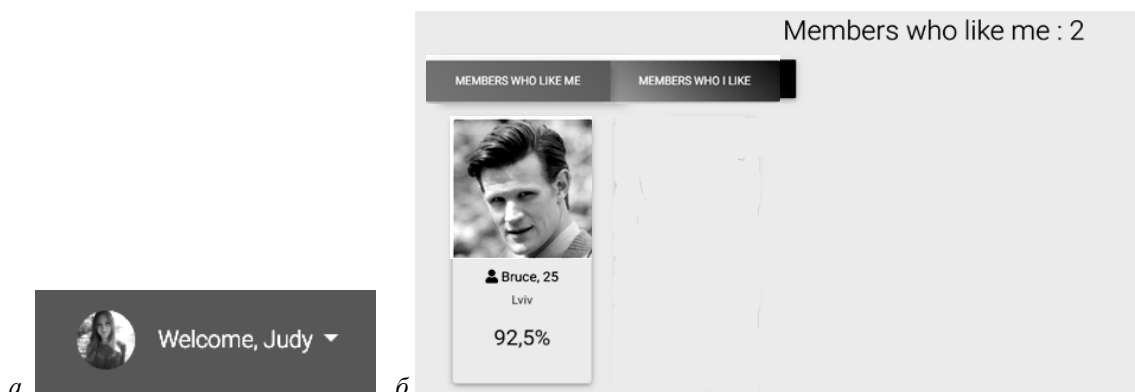


Рис. 34. Вхід іншого користувача в систему (а) та користувачі, що вибрали поточного користувача (б)

*Відомості про технічні та програмні засоби, що забезпечують виконання програми.* Інформаційну систему створено у вигляді вебзастосунку, відповідно в користувача системи має бути комп'ютер або телефон із доступом до інтернету та браузер.

*Вимоги до складу і параметрів периферійних пристроїв.* Серед периферійних пристроїв варто виділити клавіатуру та мишу для введення даних та монітор для відображення даних.

*Програмне забезпечення, необхідне для функціонування програми.* Серед програмних засобів потрібен лише браузер, бажано останньої версії із доступом до інтернету, та можливість використання соокіе файлів.

*Вимоги до програмного забезпечення.* Повинна бути будь-яка робоча операційна система, встановлені драйвери для використання клавіатури, миші та монітора та одна з останніх версій будь-якого браузера.

*Вимоги та організаційні, технічні й технологічні умови.* Для повноцінної роботи системи повинен бути постійний доступ до інтернету, увімкнені соокіе файли в браузері та мають працювати пристрої введення та виведення даних користувачеві. На рис. 35 подано приватне листування початкового користувача з вибраним користувачем інформаційної системи.

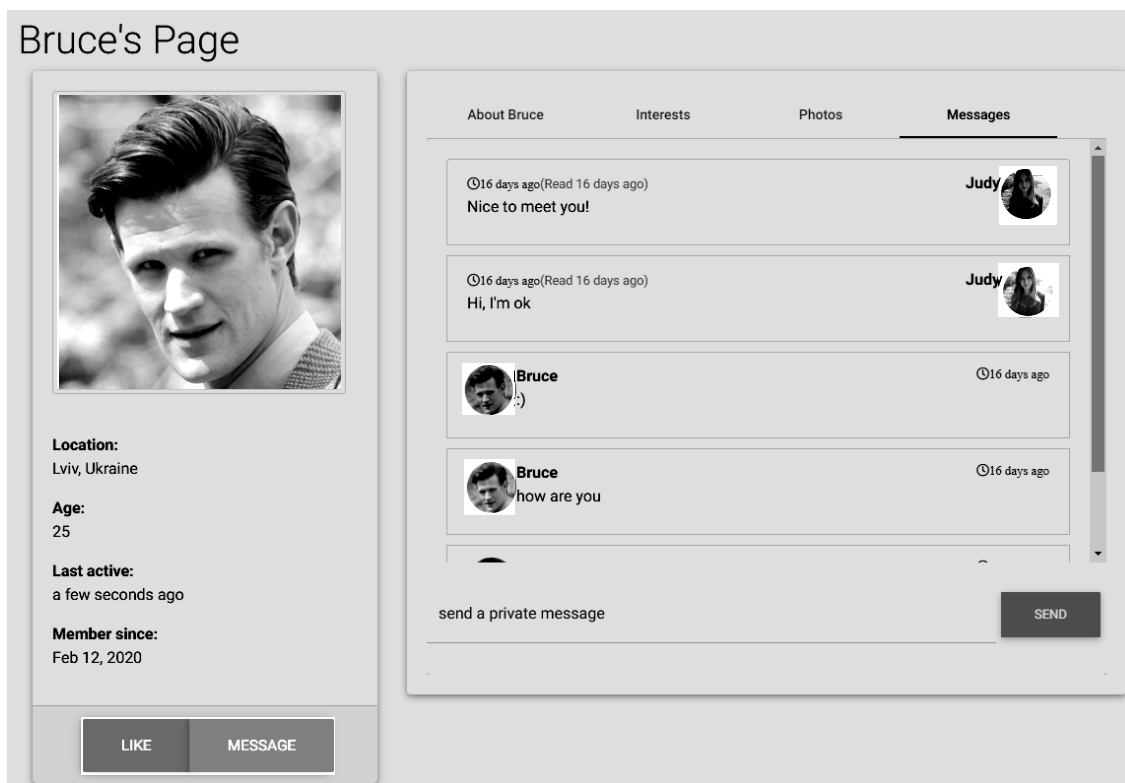


Рис. 35. Приватне листування користувачів

Реалізувавши інформаційну систему, ми виконали статистичний аналіз двох параметрів системи, а саме порівняння швидкості формування вибірки користувачів та точності отриманого відсоткового співвідношення. Порівняно реалізовану в інформаційній системі комбінацію алгоритмів Левенштейна, N-грам, розширення вибірки, моделі Noisy Channel та звичайного алгоритму Левенштейна, який найчастіше використовують у схожих системах соціалізації особистостей. Спочатку проаналізовано ефективність формування вибірки користувачів, вибрано 12 користувачів системи і для кожного сформовано вибірку із використанням комбінації алгоритмів та звичайного алгоритму Левенштейна. Отриману діаграму наведено на рис. 43. З неї можна зробити висновок, що реалізована в системі комбінація алгоритмів ефективніша та точніша приблизно на 25–30 % порівняно зі звичайним алгоритмом Левенштейна. Далі проаналізовано швидкість формування вибірки користувачів, знову вибрано 12 користувачів системи і для кожного сформовано вибірку із використанням комбінації алгоритмів та з використанням звичайного алгоритму Левенштейна. Із

отриманої діаграми (рис. 37) можна зробити висновок, що реалізована в системі комбінація алгоритмів здійснює вибірку приблизно в десять разів швидше, ніж звичайний алгоритм Левенштейна.



Рис. 36. Відсоткове співвідношення між користувачами



Рис. 37. Час формування вибірки даних

### Висновки

У наш час соціалізація особистостей за спільними інтересами надзвичайно важлива, оскільки більшість людей намагаються спростити та автоматизувати всі основні життєві процеси, які зазвичай займають багато вільного часу. Те саме стосується і процесу соціалізації. Створена інформаційна система в цьому плані відіграє важливу роль. Під час виконання роботи спершу

проаналізовано основні наукові роботи, які є дотичними до цієї теми і відіграють дуже важливу роль у досягненні розуміння основних процесів під час реалізації роботи інформаційної системи соціалізації особистостей за спільними інтересами. Досліджено основні алгоритми, важливі в соціальних мережах та у різних системах взаємодії користувачів, визначено алгоритми, які найкраще підходять для поточної інформаційної системи. Досліджено використання сучасних нейронних мереж та їх важливість у роботі інформаційної системи в питаннях опрацювання даних. У ході системного аналізу для загального окреслення суті інформаційної системи побудовано дерево рішень та визначено основну мету функціонування системи. Здійснено також конкретизацію структури інформаційної системи за допомогою IDEF0 та IDEF3 діаграм, за допомогою деревоподібної структури побудовано ієрархію процесів системи. Вибір правильних програмних рішень та необхідного програмного забезпечення для написання інформаційної системи є дуже важливим кроком, тому що від цього залежить більшість особливостей створеної системи. Для створення інформаційної системи вибрано мови програмування C# і JavaScript та їх основні фреймворки, а саме ASP .NET Core та Angular і відповідні середовища розроблення програмного коду Visual Studio та Visual Studio Code. Це дало змогу створити багаторівневу систему, яка поділена на серверну і клієнтську частини, що динамічно та повністю асинхронно взаємодіють між собою у ході роботи системи. Останнім та найважливішим кроком була практична реалізація інформаційної системи соціалізації особистостей за спільними інтересами. Передусім було написано простий та надійний функціонал реєстрації та здійснено подальшу авторизацію користувача системи із використанням методів Identity та JWT токенів, що дало змогу надійно зберігати паролі користувачів у базі даних, оптимізувати процедури створення сесії та надавати увесь необхідний функціонал під час роботи користувача в системі. Відтак забезпечено реалізацію роботи двох нейронних мереж: згорткової та сіамської, що дало змогу здійснити пошук людського обличчя на фотографіях, що завантажує користувач, і порівняти знайдене обличчя з уже наявними в базі даних. Це дає можливість ефективно ідентифікувати справжність користувача та гарантувати, що цього користувача на певний момент немає в базі даних, відповідно він є реальним. За допомогою алгоритмів нечіткого пошуку, алгоритму Левенштейна та моделі Noisy Channel створено алгоритм аналізу та порівняння користувачької інформації, який для поточного користувача формує список наявних користувачів системи, посортований за зменшенням відсоткового співвідношення схожості користувачів, та вказує, наскільки інтереси інших користувачів збігаються з інтересами поточного користувача.

У ході виконання роботи здебільшого досягнуто основні цілі відповідно до поставленої мети, однак ще багато аспектів можна удосконалити всередині створеної інформаційної системи. Проте функціонал, доступний всередині інформаційної системи на цей момент, відповідає поставленій меті створення інформаційної системи і забезпечено можливістю пройти реєстрацію усім користувачам системи, вказати всі необхідні дані про себе, завантажити фотографії зі своїм обличчям і після цього здійснювати ефективну соціалізацію, а саме переглядати користувачів зі сформованого списку, ставити позначку про вподобання користувача та здійснювати приватне листування з вибраним користувачем всередині інформаційної системи.

### Список літератури

1. Chu, S. C. Using a consumer socialization framework to understand electronic word-of-mouth (eWOM) group membership among brand followers on Twitter. *Electronic Commerce Research and Applications*. 2016. No. 14 (4). С. 251–260.
2. Batiuk, T., Vysotska, V., Lytvyn, V. Intelligent System for Socialization by Personal Interests on the Basis of SEO-Technologies and Methods of Machine Learning. *Computational Linguistics and Intelligent Systems (COLINS 2020): 4th International Conference, Lviv, 23–24 April 2020* : CEUR workshop proceedings. 2020. No. 2604. С. 1237–1250.
3. Vysotska, V. Information Technology for Internet Resources Promotion in Search Systems Based on Content Analysis of Web-Page Keywords. *Radio Electronics, Computer Science, Control*. 2021. No. 3. С. 133–151.

4. De-Gregorio, F., Sung, Y. Understanding attitudes toward and behaviors in response to product placement. *Journal of Advertising*. 2010. No. 39 (1). C. 83–96. DOI: <http://doi.org/10.2753/JOA0091-3367390106>.
5. Elaheebocus S. M., Weal M., Morrison L. Peer-based social media features in behavior change interventions: Systematic review. *Journal of Medical Internet Research*. 2018. No. 20 (2). C. 1–20. DOI: <http://doi.org/10.2196/jmir.8342>.
6. Erkan I. The influence of e-WOM in social media on consumers' purchase intentions: An extended approach to information adoption. *Computers in Human Behavior*. 2016. No. 4. C. 47–55.
7. Ferrara E., Interdonato R., Tagarelli A. Online popularity and topical interests through the lens of Instagram. *Hypertext and Social Media*. 2014. No. 2. C. 24–23. DOI: <http://doi.org/10.1145/2631775.2631808>.
8. Gao L. Online consumer behavior and its relationship to website atmospheric induced flow: Insights into online travel agencies in China. *Journal of Retailing and Consumer Services*. 2014. No. 21 (4). C. 653–655.
9. Geurin-Eagleman A. N. Communicating via photographs: A gendered analysis of Olympic athletes' visual self -presentation on Instagram. *Sport Management Review*. 2015. No. 19 (2). C. 133–145. DOI: <http://doi.org/10.1016/j.smr.2015.03.002>.
10. Guidry J. D., Messner M., Jin Y. From McDonalds fail to Dominos sucks: An analysis of Instagram images about the 10 largest fast food companies. *Corporate Communications: An International Journal*. 2015. No. 20 (3). C. 344–359.
11. Hanna R., Rohm A., Crittenden V. L. We're all connected: The power of the social media ecosystem. *Business Horizons*. 2011. No. 54 (3). C. 265–273. DOI: <http://doi.org/10.1016/j.bushor.2011.01.007>.
12. Salganik M. Social Research in the Digital Age. *Journal of Interactive Marketing*. 2019. No. 2 (9). C. 345–358.
13. Hudson S., Roth M., Madden T. J. The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees. *Tourism Management*. 2015. No. 2 (8). C. 68–76. DOI: <http://doi.org/10.1016/j.tourman.2014.09.001>.
14. Jeff M., Jennifer R., Catherine J., Elke P. Managing brand presence through social media: The case of UK football clubs. *Internet Research*. 2014. No. 24 (2). P. 181–204.
15. Kim E. Brand followers' retweeting behaviour on Twitter: How brand relationship influence brand electronic word-of-mouth. *Computers in Human Behavior*. 2014. No. 38 (8). C. 18–25.
16. Kudeshia C., Sikdar P., Mittal A. Spreading love through fan page liking: A perspective on small scale entrepreneurs. *Computers in Human Behavior*. 2016. No. 8 (19). C. 257–270. DOI: <http://doi.org/10.1016/j.chb.2015.08.003>
17. Lueg J. E. Interpersonal communication in the consumer socialization process: Scale development and validation. *Journal of Marketing Theory and Practice*. 2007. No. 15 (1). C. 25–39. DOI: <http://doi.org/10.2753/MTP1069-6679150102>.
18. Mousavijad M. The effect of socialization factors on decision making of teenagers consumers in schools. *Journal of School Administration*. 2017. No. 5 (1). C. 217–234.
19. Schnell R. Enhancing Surveys with Objective Measurements and Observer Ratings. *Journal of Interactive Marketing*. 2015. No. 44 (1). C. 288–302.
20. Parry M. E., Kawakami T., Kishiya K. The effect of personal and virtual word-of-mouth on technology acceptance. *Journal of Product Innovation Management*. 2012. No. 29 (6). C. 952–966. DOI: <http://doi.org/10.1111/j.1540-5885.2012.00972.x>.
21. Quan-Haase A., Sloan L. Introduction to the Handbook of Social Media Research Methods: Goals, Challenges and Innovations. *The Sage Handbook of Social Media Research Methods*. 2017. No. 10 (5). C. 606–859. DOI: <http://doi.org/10.4135/9781473983847.n1>.
22. Murphy S. T. Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*. 2001. No. 8 (3). C. 723–739. DOI: <http://doi.org/10.1037/0022-3514.64.5.723>.
23. Ntale P. D. Word of mouth communication: A mediator of relationship marketing and customer loyalty. *Cogent Business and Management*. 2019. No. 6 (18). C. 3–36.
24. Schmah M., Wilke T., Rossmann A. Electronic word of mouth: A systematic literature analysis. *Digital Enterprise Computing*. 2017. C. 147–158.
25. Ozdemir A., Tozlu B., Şen E., Ateşoğlu A. Analyses of word-of-mouth communication and its effect on students' university preferences. *Procedia - Social and Behavioral Sciences*. 2016. No. 8 (5). C. 22–35.
26. Park J., Ciampaglia G. L., Ferrara F. Style in the age of Instagram: Predicting success within the fashion industry using social media. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 2016. No. 22 (8). C. 64–72. DOI: <http://doi.org/10.1145/2818048.2820065>.

27. Ranjbaran B., Jamshidian M., Dehghan Z. A survey of identification of major factors influencing customers attitude toward machine made carpet brands. *Journal of Business Strategies*. 2007. No. 5 (23). С. 109–118.

### References

1. Chu, S. C. (2016). Using a consumer socialization framework to understand electronic word-of-mouth (eWOM) group membership among brand followers on Twitter. *Electronic Commerce Research and Applications*, No. 14 (4), 251–260.
2. Batiuk, T., Vysotska, V., Lytvyn, V. (2020). Intelligent System for Socialization by Personal Interests on the Basis of SEO-Technologies and Methods of Machine Learning. *Computational Linguistics and Intelligent Systems (COLINS 2020): 4th International Conference*, Lviv, 23–24 April 2020 : CEUR workshop proceedings, No. 2604, 1237–1250.
3. Vysotska, V. (2021). Information Technology for Internet Resources Promotion in Search Systems Based on Content Analysis of Web-Page Keywords. *Radio Electronics, Computer Science, Control*, No. 3, 133–151.
4. De-Gregorio, F., Sung, Y. (2010). Understanding attitudes toward and behaviors in response to product placement. *Journal of Advertising*, No. 39 (1), 83–96. DOI: <http://doi.org/10.2753/JOA0091-3367390106>.
5. Elaheebocus, S. M., Weal, M., Morrison, L. (2018). Peer-based social media features in behavior change interventions: Systematic review. *Journal of Medical Internet Research*, No. 20 (2), 1–20. DOI: <http://doi.org/10.2196/jmir.8342>.
6. Erkan, I. (2016). The influence of e-WOM in social media on consumers' purchase intentions: An extended approach to information adoption. *Computers in Human Behavior*, No. 4, 47–55.
7. Ferrara, E., Interdonato, R., Tagarelli, A. (2014). Online popularity and topical interests through the lens of Instagram. *Hypertext and Social Media*, No. 2, 24–23. DOI: <http://doi.org/10.1145/2631775.2631808>.
8. Gao, L. (2014). Online consumer behavior and its relationship to website atmospheric induced flow: Insights into online travel agencies in China. *Journal of Retailing and Consumer Services*, No. 21 (4), 653–655.
9. Geurin-Eagleman, A. N. (2015). Communicating via photographs: A gendered analysis of Olympic athletes' visual self-presentation on Instagram. *Sport Management Review*, No. 19 (2), 133–145. DOI: <http://doi.org/10.1016/j.smr.2015.03.002>.
10. Guidry, J. D., Messner, M., Jin, Y. (2015). From McDonalds fail to Dominos sucks: An analysis of Instagram images about the 10 largest fast food companies. *Corporate Communications: An International Journal*, No. 20 (3), 344–359.
11. Hanna, R., Rohm, A., Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. *Business Horizons*, No. 54 (3), 265–273. DOI: <http://doi.org/10.1016/j.bushor.2011.01.007>
12. Salganik, M. (2019). Social Research in the Digital Age. *Journal of Interactive Marketing*, No. 2 (9), 345–358.
13. Hudson, S., Roth, M., Madden, T. J. (2015). The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees. *Tourism Management*, No. 2 (8), 68–76. DOI: <http://doi.org/10.1016/j.tourman.2014.09.001>.
14. Jeff, M., Jennifer, R., Catherine, J., Elke, P. (2014). Managing brand presence through social media: The case of UK football clubs. *Internet Research*, No. 24 (2), 181–204.
15. Kim, E. (2014). Brand followers' retweeting behaviour on Twitter: How brand relationship influence brand electronic word-of-mouth. *Computers in Human Behavior*, No. 38 (8), 18–25.
16. Kudeshia, C., Sikdar, P., Mittal, A. (2016). Spreading love through fan page liking: A perspective on small scale entrepreneurs. *Computers in Human Behavior*, No. 8 (19), 257–270. DOI: <http://doi.org/10.1016/j.chb.2015.08.003>.
17. Lueg, J. E. (2007). Interpersonal communication in the consumer socialization process: Scale development and validation. *Journal of Marketing Theory and Practice*, No. 15 (1), 25–39. DOI: <http://doi.org/10.2753/MTP1069-6679150102>.
18. Mousavijad, M. (2017). The effect of socialization factors on decision making of teenagers consumers in schools. *Journal of School Administration*, No. 5 (1), 217–234.
19. Schnell, R. (2015). Enhancing Surveys with Objective Measurements and Observer Ratings. *Journal of Interactive Marketing*, No. 44 (1), 288–302.
20. Parry, M. E., Kawakami, T., Kishiya, K. (2012). The effect of personal and virtual word-of-mouth on technology acceptance. *Journal of Product Innovation Management*, No. 29 (6), 952–966. DOI: <http://doi.org/10.1111/j.1540-5885.2012.00972.x>.

21. Quan-Haase, A., Sloan, L. (2017). Introduction to the Handbook of Social Media Research Methods: Goals, Challenges and Innovations. *The Sage Handbook of Social Media Research Methods*, No. 10 (5), 606–859. DOI: <http://doi.org/10.4135/9781473983847.n1>.
22. Murphy, S. T. (2001). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, No. 8 (3), 723–739. DOI: <http://doi.org/10.1037/0022-3514.64.5.723>.
23. Ntale, P. D. (2019). Word of mouth communication: A mediator of relationship marketing and customer loyalty. *Cogent Business and Management*, No. 6 (18), 3–36.
24. Schmah, M., Wilke, T., Rossmann, A. (2017). Electronic word of mouth: A systematic literature analysis. *Digital Enterprise Computing*, 147–158.
25. Ozdemir, A., Tozlu, B., Şen, E., Ateşoğlu, A. (2016). Analyses of word-of-mouth communication and its effect on students' university preferences. *Procedia – Social and Behavioral Sciences*, No. 8 (5), 22–35.
26. Park, J., Ciampaglia, G. L., Ferrara, F. (2016). Style in the age of Instagram: Predicting success within the fashion industry using social media. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, No. 22 (8), 64–72. DOI: <http://doi.org/10.1145/2818048.2820065>.
27. Ranjbaran, B., Jamshidian, M., Dehghan, Z. (2007). A survey of identification of major factors influencing customers attitude toward machine made carpet brands. *Journal of Business Strategies*, No. 5 (23), 109–118.

## INFORMATION SUPPORT FOR PERSONALITIES SOCIALIZATION PROCESSES BASED ON COMMON INTERESTS

Taras Batiuk<sup>1</sup>, Victoria Vysotska<sup>2</sup>

Lviv Polytechnic National University, Information Systems and Networks Department

<sup>1</sup> e-mail: [taras.batiuk.mnsa.2020@lpnu.ua](mailto:taras.batiuk.mnsa.2020@lpnu.ua), ORCID: 0000-0001-5797-594X,

<sup>2</sup> e-mail: [Victoria.A.Vysotska@lpnu.ua](mailto:Victoria.A.Vysotska@lpnu.ua), ORCID: 0000-0001-6417-3689

© Batiuk T., Vysotska V., 2022

The main objective of this article is to create an information system project for socialization by personal interests on the basis of SEO-technologies and methods of machine learning. The main purpose of this information system is to identify the user within the system using neural networks and to select similar users by analysing the user's current information. An information system was created that, through Identity and JWT tokens, provides optimized and secure authorization, logging, and support functions for the current system user session. Finding a face in a user's photo and checking the presence of a similar user in the database are implemented using convolutional and Siamese neural networks. The analysis and formation of similar user beeps were implemented using fuzzy search algorithms, the Levenshtein algorithm and the Noisy Channel model, which made it possible to maximize the automation of the user selection process and to optimize the time spent in this process. Tools have also been created to view other users' profiles, preferences and private correspondence. All private correspondence and information about it are stored in the current database. Each user of the system can view all information about sent and received messages. The created information system implements the process of user identification, analysis, selection and further socialization of system users.

**Key words:** project; fuzzy search; convolutional neural network; Siamese neural network; Levenshtein distance; Noisy Channel.