# ВСТУП

Теперішній розвиток інформаційних технологій має великий вплив на всі сфери діяльності людини. Саме це стає рушійним фактором для їх впровадження у більшість галузей. Виключенням не стала і генеалогія та генетика, адже за останні роки рівень використання інформаційних систем у цих напрямках значно підвищився.

Майже у кожної особи нашого суспільства виникало прагнення до сімейної історії. Із приходом інформаційних технологій це прагнення стає значно доступнішим, адже з’являються можливості легкого збору, зберігання та аналізування історичних даних за допомогою існуючих програмних продуктів.

Питання використання інформаційних систем для створення генеалогічного дерева родини є цілком актуальним. Проведення всіх етапів збору та аналізу інформації без використання сучасної обчислювальної техніки вимагає кропіткої праці, а також великої кількості часу у відповідного спеціаліста. На практиці часто виникають помилки у неправильному знаходженні та зберіганні інформації про родовід. Тому на сьогоднішній день програмні продукти, що пов’язані з побудовою генеалогічних дерев, є досить важливими в практичному застосуванні, але вони через велику вартість є не зовсім доступними.

Для того щоб інформаційна система мала широке використання, в основу програмного продукту закладають велику базу даних та базу знань. За допомогою даної технології можна отримати змогу збирати та впорядковувати величезну кількість даних, а також проводити різноманітні операції з ними.

Темою дипломної роботи є розробка інформаційної системи генеалогічного дерева родини.

**Актуальність теми** полягає у тому, що дана інформаційна система дозволить покращити зберігання історичної інформації про сім’ю. Є можливість визначення ймовірності захворювання генетичними хворобами від пращурів до нащадків, побудова графіків за масивом інформаційних даних від усіх родичів. Також актуальним є реалізація онлайн мережі у якій є можливість додавання користувачів до друзів з функцією перегляду, та копіювання інформації про рідню. Створення бази знань генетичних захворювань, що можуть проявлятися у людей. Все це надасть можливість зберігати більш різноманітнішу інформацію про свій родовід.

**Об’єкт дослідження** – методи та засоби створення інформаційної системи генеалогічного дерева родини.

**Предмет дослідження** – це програмні засоби для автоматизації зберігання та аналізу інформації про родовід, а також можливість знаходження ймовірності передачі генетичних захворювань від пращурів до нащадків .

**Область застосування –** інформаційна система є корисною для користувачів, які поставили собі за мету сформувати своє генеалогічне дерево, що міститиме всю відому інформацію про предків та дає можливість передбачити генетичні захворювання у нащадків.

**Мета даної роботи** полягає у створенні інформаційної системи генеалогічного дерева родини, що полегшує формування загальної картини роду. Дана система розробляється для роботи з великим потоком даних, а також з метою скоротити витрати часових ресурсів користувачів у цілому.

**Практичне значення** одержаних результатів полягає в формуванні генеалогічного дерева родини, визначення ймовірності прояву генетичних захворювань у нащадків. Автоматизації розрахункових задач даної предметної області.

Дипломна робота складається із вступу, \_\_\_розділів, висновків, списку використаної літератури (\_\_\_ найменувань). Робота містить \_\_\_ таблиці, \_\_\_ рисунка. Роботу викладено на \_\_\_ сторінках друкованого тексту.

# 1 ОГЛЯД ТА порівняльний АНАЛІЗ аналогів ІНФОРМАЦІЙНОЇ СИСТЕМИ «генеалогічного дерева родини»

## 1.1 Призначення та область застосування інформаційної системи генеалогічного дерева родини

Генеалогічне (родовідне) дерево – схематичне представлення родинних зв'язків у вигляді умовно-символічного дерева, в коренях якого знаходиться родозасновник, а на кроні та вітках – представники різних ліній родоводу[Савелов Л. М. Лекции по русской генеалогии. — М.: Археографический центр. 1994. — Репринт издания 1909 года. ISBN 5-86169-015-4 (Т. 3); ISBN 5-86169-015-4 (рос.)]. Таке схематичне зображення надає користувачеві більш зручний вигляд представлення родоводу.

Користувач має можливість пошуку та зберігання різноманітної інформації про своїх предків, але робота із великою кількістю даних одночасно, не є досить зручною та займає велику кількість часу. Тому доцільно використовувати інформаційну систему генеалогічного дерева родини.

Інформаційна система – це сукупність організаційних і технічних засобів для збереження та обробки інформації з метою забезпечення інформаційних потреб користувачів[Грицунов О. В. Інформаційні системи та технології. Навчальний посібник. — Х.: ХНАМГ, 2010. — 222 с.].

Інформаційна система генеалогічного дерева має можливість визначення ймовірності прояву генетичних захворювань, що передались від предків до нащадків, це стає можливим завдяки генетичним алгоритмам, зокрема за допомогою законів Менделя, що становлять основу класичної генетики спадковості. Генетичні захворювання це захворювання, обумовлені порушеннями в процесах збереження, передачі та реалізації генетичної інформації. В основі спадкових захворювань лежать мутації: генні, хромосомні та геномні. Відповідно до цього всі спадкові хвороби людини можна об’єднати в 4 великі групи:

1. генні (зміни на рівні окремих нуклеотидів),
2. геномні (зміни кількості цілих хромосом),
3. хромосомні (внутрішньо - і міжхромосомні перебудови),
4. мультифакторіальні (на розвиток хвороби впливають і гени, і фактори навколишнього середовища).

Спадкові захворювання і захворювання із спадковою схильністю є основною причиною смертності, особливо в економічно розвинутих країнах. На підставі великого статистичного матеріалу встановлено, що через генетичні порушення 1 із 130 зародків гине вже в перші дні, 25 % припиняють своє існування на пізніших строках вагітності. 40 % дитячої смертності зумовлено спадковими дефектами, 5-12 % новонароджених мають природжені генетичні дефекти, які зумовлені дефектом хромосом або мутаціями окремих генів. Кожна людина має 5-10 потенціально шкідливих генів, які передаються нащадкам разом з нормальними генами.[http://www.pharmencyclopedia.com.ua/article/591/spadkovi-xvorobi].

Основним призначенням продукту проектування є побудова генеалогічного дерева родини, яке представлене у класичній схемі, формування графіків на основі відомих даних та знаходження ймовірності прояву спадкових генетичних захворювань у майбутніх нащадків.

Основною областю застосування даної інформаційної системи генеалогічного дерева родини є діяльність користувачів які прагнуть дізнатися більше про власний родовід та про його подальший розвиток.

## 1.2 Огляд способів і засобів вирішення поставлених задач

### 1.2.1 Поняття «колективного розуму»

Вислів «колективний розум» поширюється вже кілька десятиліть, але став важливим і популярним з приходом нових комунікаційних технологій. Під цим поняттям розуміють витяг нового знання з об'єднаних переваг, поведінки і уявлень деякої групи людей.

Колективний розум був можливий і до появи Інтернету. Для того щоб збирати дані від розрізнених груп людей, об'єднувати їх і аналізувати, Всесвітня павутина не потрібна. До найважливіших форм подібних досліджень входять соціологічні опитування та переписи. Одержання відповідей від великого числа людей дозволяє робити про групу такі статистичні висновки, які на основі одиничних даних зробити неможливо. Породження нових знань, виходячи з даних, отриманих від незалежних респондентів, – це і є суть колективного розуму.

Прикладом можуть бути фінансові ринки, де ціна встановлюється не за бажанням індивідуума чи шляхом скоординованих зусиль, а в результаті поведінки на торгах безлічі незалежних людей, які діють у власних інтересах. Вважається, що ф'ючерсні ринки, де численні учасники укладають контракти виходячи зі своїх уявлень про майбутні ціни, здатні передбачати ціни більш точно, ніж незалежні експерти. Пояснюється це тим, що такі ринки акумулюють знання, досвід та інтуїцію тисяч людей, а експерт-одинак може покладатися лише на себе. Хоча методи колективного розуму існували і до появи Інтернету, можливість отримувати інформацію від тисяч і навіть мільйонів людей у Всесвітній павутині відкрила найширший спектр нових можливостей. У будь-який момент часу хтось користується Інтернетом для здійснення покупок, в дослідницьких цілях, в пошуках розваг або заради створення власного сайту. Ці дії можна відстежувати і вилучати з них інформацію, навіть не ставлячи користувачеві питань. Існує безліч способів обробити та інтерпретувати цю інформацію. Розглянемо кілька приклади, які ілюструють принципово різні підходи.

Вікіпедія – це онлайн енциклопедія, створювана виключно самими користувачами. Будь-яка людина може створити нову статтю або відредагувати вже існуючу, а повторювані спроби недозволеного використання відсікають не багаточисленні адміністратори. У Вікіпедії більше статей, ніж у будь-якій іншій енциклопедії, і, незважаючи на маніпуляцію деяких зловмисних користувачів, вважається, що за більшістю тем інформація є точною. Це приклад колективного розуму, оскільки кожна стаття підтримується великою групою людей, а в результаті виходить творіння, що набагато перевершує все, що могла б створити одна організована група. Програмне забезпечення Вікіпедії не піддає зібрану інформацію якийсь особливо інтелектуальній обробці; воно просто відстежує зміни і відображає останню версію.

Пошукова система Google – найпопулярніша в світі пошукова машина. У ній вперше для ранжирування веб-сторінок був застосований підхід, заснований на кількості посилань на дану сторінку. Для цього необхідно зібрати інформацію про те, що думають про дану сторінку тисячі людей, і скористатися нею для сортування результатів пошуку. Це зовсім інший приклад колективного розуму. Якщо Вікіпедія запрошує користувачів зробити свій внесок у створення сайту, то Google витягує інформацію з того контенту, який користувачі вже розмістили на своїх власних сайтах, і застосовує її для генерування оцінок від імені користувачів.

Термін «колективний розум» використовується у Всесвітній павутині, означаючи мережеві служби, що обробляють дані, зібрані з різних джерел і від різних учасників, та знаходять серед них найкориснішу інформацію для користувачів. Приклади реалізації: вікі-проекти за спільним накопиченням, вдосконалення та публікації знань (Вікіпедія); сервіси запитань і відповідей Google Answers; Habrahabr; групова розробка відкритого програмного забезпечення та ін[Говард Блум «Глобальный мозг: эволюция коллективного разума от Большого Взрыва до 21-го века» (Howard Bloom, Global Brain: The Evolution of Mass Mind from the Big Bang to the 21st Century, 2000)].

### 1.2.2 Особливості виконання генеалогічного дослідження

Генеалогічне дослідження передбачає пошук інформації у архівах щодо окремої родини на протязі кількох поколінь. Для виконання дослідження необхідна точна біографічна інформація: імена, прізвища та по батькові, місце та роки народження, місце проживання, здобутки відомих членів родини.

Повнота і чіткість відомостей є гарантією правильного пошуку необхідних даних.

Основними джерелами генеалогічних досліджень та історії родини є метричні книги, реєстрації актів громадянського стану. Іншими, також важливими матеріалами є документи посімейно-майнового обліку (для підросійської України це наприклад подвірні, посімейні (погосподарчі) списки, ревізькі «сказки», списки платників податку на прибуток, платників обов'язкового страхування, виборців та громадян, що були позбавлені виборчого права; справи про опікунство, межові книги, документація переписів 1917, 1921, 1926, 1939 рр.), джерела біографічного характеру (послужні (формулярні) списки та особові справи, переліки працівників підприємств, установ, та організацій, учасників релігійних спільнот, учнів, документація про стосунок до військової служби, фільтраційні і архівно-слідчі справи). Опис цих джерел подамо далі.

Виконання генеалогічного дослідження потрібно починати з наявної інформації, це такі дані як рік, дата народження одного з членів сім’ї. Спершу шукаємо актовий запис про народження особи в метричній книзі за місцем її народження.

У разі віднайдення цієї інформації можна дізнатися імена та по батькові батьків. Для того, щоби віднайти інших членів родини (сестер та братів) потрібний уважний перегляд рік за роком актових матеріалів про народження та смерті. Актові записи про шлюб потрібно переглянути для визначення років народження та місця приписки (можливого місця народження) батьків або дівочого прізвища матері.

Потрібно звертати увагу на всіх людей з таким самим прізвищем, оскільки вони можуть виявитися родичами).

Значно прискорить та полегшить виконання дослідження перегляд посімейних списків та ревізьких «сказок». За наведеними у цих документах роками смерті та народження можливо, уникаючи повного перегляду усіх наявних метричних книг, знайти потрібні актові записи. У разі, якщо деякі дані. потрібні для проведення дослідження відсутні, то можна звернутися до інших документів. Характеристику цих документів подамо у наступних статтях.

Для полегшення побудови родоводу рекомендується складати генеалогічні картки на кожну знайдену людину

[http://rodovody.com.ua/index.php/ua/genealogy].

## 1.3 Аналіз переваг та недоліків існуючих аналогів

Програмних продуктів які дозволяють будувати генеалогічні дерева досить багато. Коротко розглянемо переваги та недоліки деяких з них.

### 1.3.1 Програмний продукт «Family Tree Builder»

«Family Tree Builder» - проста зручна у використанні программа, російськомовна. Дозволяє вибудувати особисте генеалогічне дерево та має можливість збереження його в зручних графіках або звітах. Можна додавати до кожного учасника сімейного дерева дату народження або смерті, його фотографії, контактні дані, стать та інші важливі дані. До кожної персони можна додати чоловіка/дружину, дітей, батьків, братів або сестер. Створене дерево можна зберегти у вигляді графіка або звіту з подальшою можливістю друку для зручного перегляду. Інтерфейс програми показана на рисунку 1 та 2.

Можливості:

* експорт свого дерева в Excel;
* створення графіків і звітів;
* збереження графіків і звітів в зручних форматах і можливість друку;
* синхронізація з сайтом розробника для публікації родоводу;
* пошук родичів з понад 65 мільйонів зареєстрованих користувачів;
* додавання точної інформації до кожної персони (ПІБ, дата народження або смерті, інформація про подружжя або дітей, фотографії та ін.).

Плюси:

* потужна програма, що дозволяє створювати сімейне дерево;
* простий і зручний інтерфейс;
* можливість скачати безкоштовно Family Tree Builder;
* наявність російськомовної локалізації.

Мінуси:

* одномовний інтерфейс;
* безкоштовна версія має обмежений функціонал.

Інтерфейс програмного продукту зображено на рисунку 1.1 та рисунку 1.2.

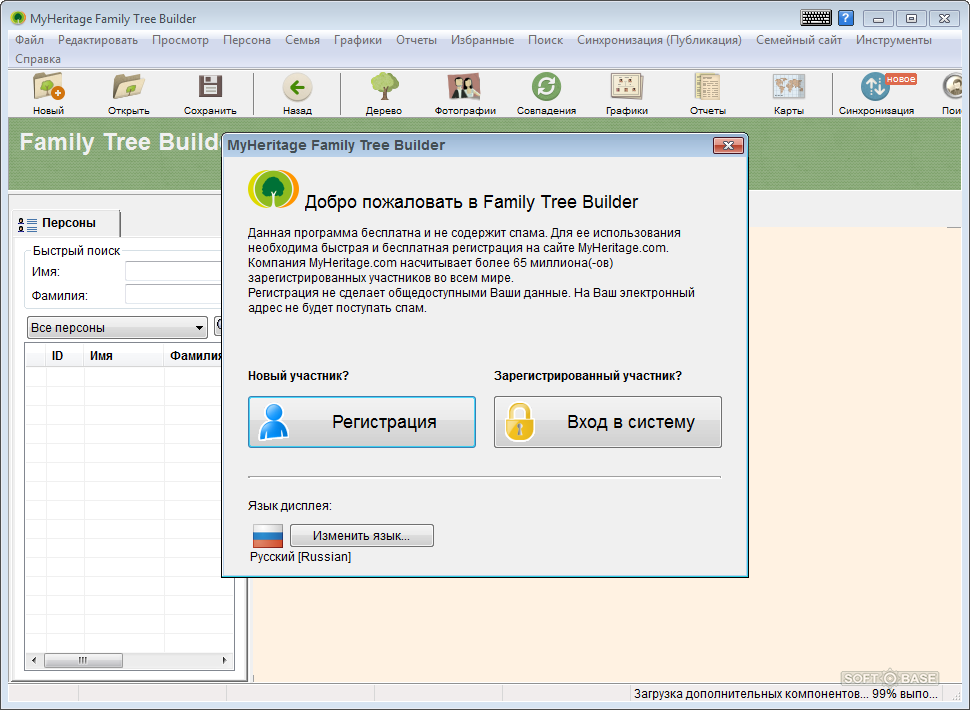


Рисунок 1.1 – Інтерфейс програми «Family Tree Builder»

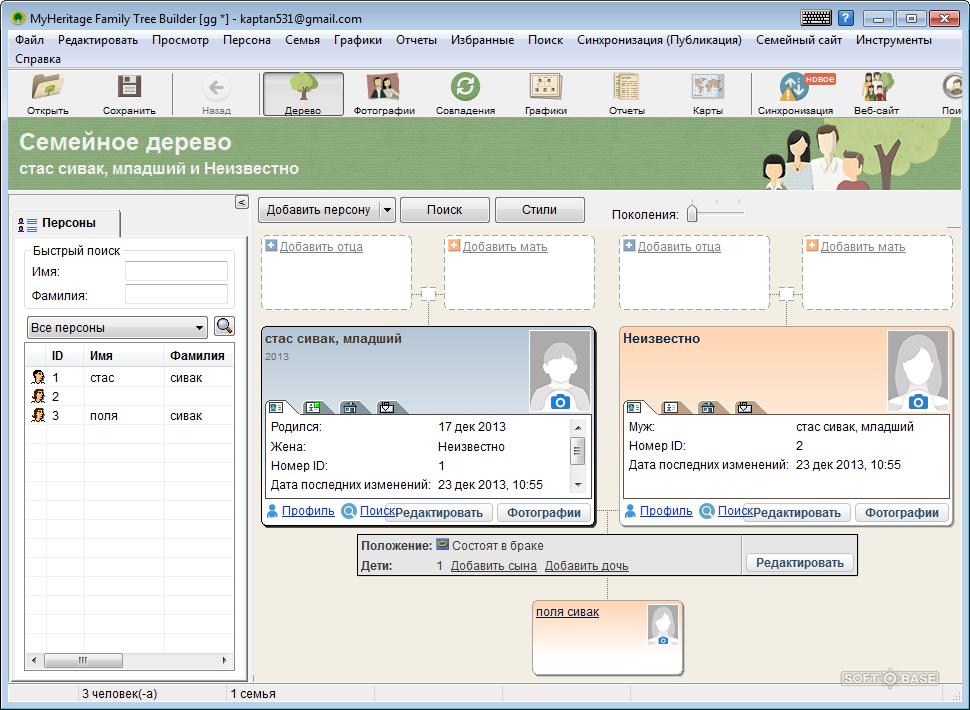


Рисунок 1.2 – Інтерфейс програми «Family Tree Builder»

### 1.3.2 Програмний продукт «SimTree»

Додаток допоможе візуально уявити генеалогічне дерево і внести коротку інформацію про кожну персони, включаючи фотографії. Інтерфейс простий і зрозумілий користувачам. Саме тому SimTree можна скачати і використовувати спільно з дітьми в навчальних цілях. Інтерфейс програми показана на рисунку 3 та 4.

Можливості:

* побудова генеалогічного дерева;
* табличне відображення інформації про кожну особу;
* додавати особисті дані до карток;
* автоматичне визначення родинних зв'язків при додаванні нових членів сім'ї.

Плюси:

* простий інтерфейс;
* зберігання практично необмеженого кількість даних;
* можливість прикріплювати файли.

Мінуси:

* обмежений функціонал;
* єдина схема візуального відображення спорідненості.

Інтерфейс програмного продукту представлено на рисунку 1.3.

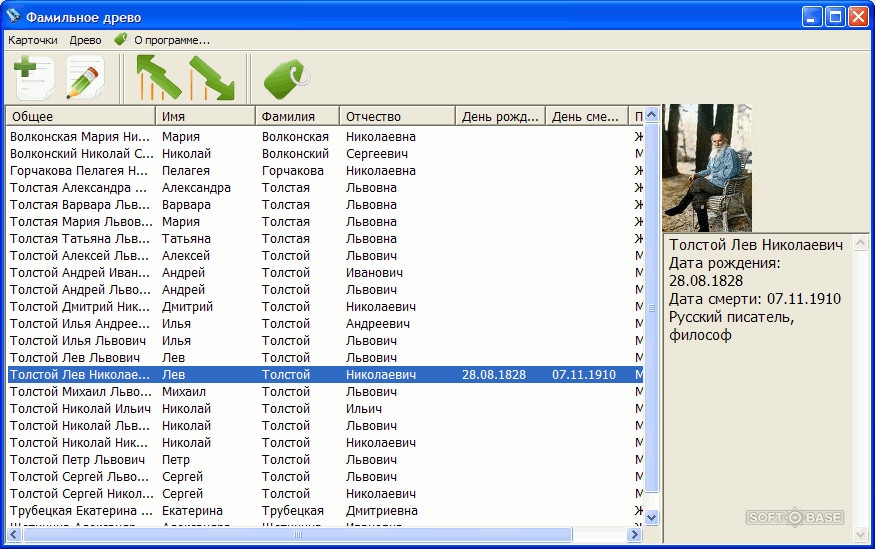


Рисунок 1.3 – Інтерфейс програми «SimTree»

### 1.3.3 Програмний продукт «Genopro»

Програма дозволяє швидко створити сімейне дерево навіть не дуже досвідченим користувачам. Завдяки наявності «гарячих клавіш» ви можете легко додати потрібних вам людей і інформацію про них. Програма умовно-безкоштовна: безкоштовна версія має обмежений термін дії (15 днів) і обмежена кількість персон для додавання (до 25 осіб). Щоб підвищити функціонал програми, необхідно буде купити ліцензійний ключ. Інтерфейс програми показана на рисунку 5 та 6.

Можливості

* наявність вбудованого генератора звітів;
* всі дані доступні для перегляду і редагування в табличному форматі;
* створення призначених для користувача вкладок даних для введення додаткової інформації;
* додавання інформації про відносини між людьми (наприклад, закоханість або конфлікт у відносинах);
* збереження резервної копії онлайн;
* відправка створеної родоводу іншим родичам по e-mail;
* друк створеної родоводу.

Плюси:

* підтримка Unicode;
* програма має російськомовну локалізацію;
* наявність «гарячих клавіш»;
* можливість безкоштовного завантаження Genopro;
* експорт в різні формати, наприклад XML, Gedcom або TXT;
* імпорт сімейного дерева в форматі Gedcom, створеного в аналогічних програмах.

Мінуси:

* обмежений термін дії безкоштовної версії (15 днів);
* обмежена кількість створених персон в пробної версії (25 осіб).

Інтерфейс програмного продукту зображено на рисунках

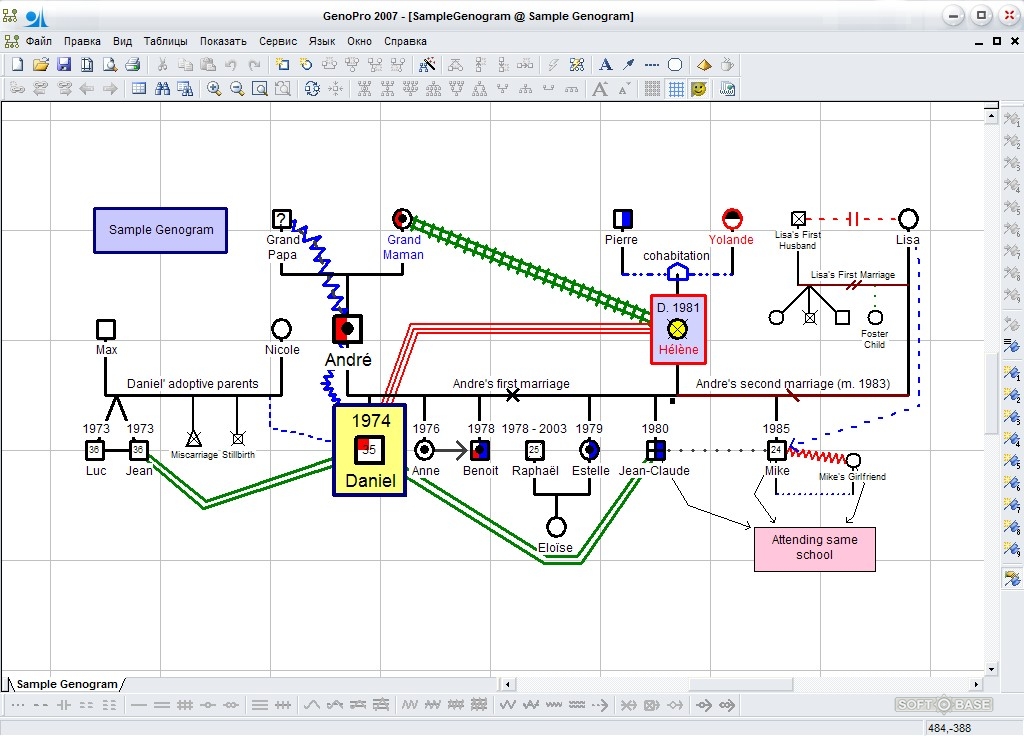


Рисунок 1.4 – Інтерфейс програми «Genopro»

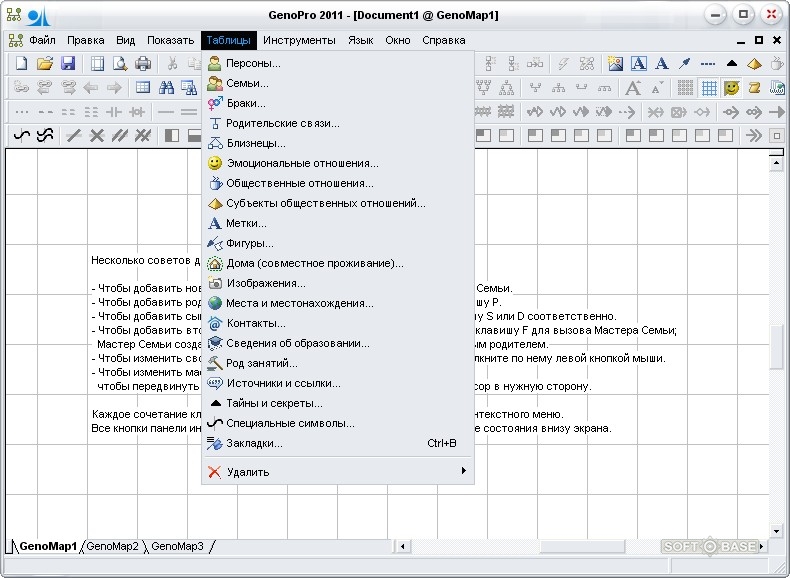


Рисунок 1.5 – Інтерфейс програми «Genopro»

### 1.3.4 Програмний продукт «RootsMagic Essentials»

Програма «RootsMagic Essentials» одна з найстаріших, її творець Bruce Buzbee написав першу версію ще в далекому 1987 році. З тих пір вона звичайно неодноразово змінювалася і допрацьовувалася і на даний момент представляє дуже симпатичну і функціональну платформу, що володіє масою достоїнств. Давайте поглянемо на цю програму ближче. Інтерфейс програми показана на рисунку 7 та 8.

Можливості:

* складання генеалогічного дерева за допомогою зручної візуальної схеми;
* дослідження предків за допомогою спеціальних алгоритмів;
* прив’язку людей до місцевості;
* пошук інформації про осіб у відкритих джерелах.

Плюси:

* простий, красивий і доступний інтерфейс;
* широкий функціонал;
* можливість публікувати своє дерево онлайн і пов'язувати його з іншими.
* Мінуси:
* програма не русифікована;
* багато важливих функцій доступні тільки в платній версії.

Інтерфейс програмного продукту представлено на рисунку 1.6 та рисунку 1.7.

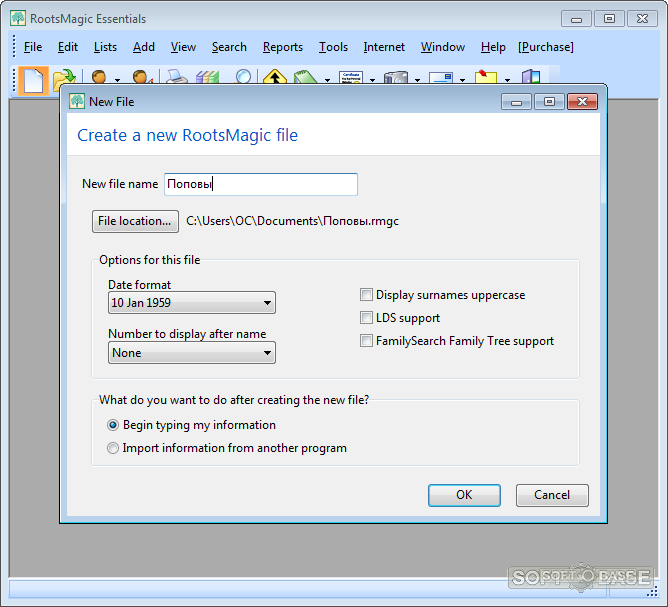


Рисунок 1.6 – Інтерфейс програми «RootsMagic Essentials»

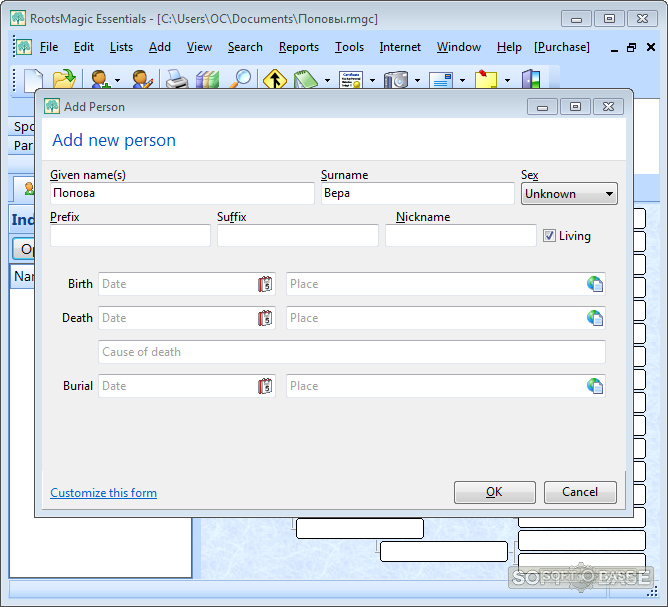


Рисунок 1.7 – Інтерфейс програми «RootsMagic Essentials»

## 1.4 Постановка задачі на розробку

Інформаційна система генеалогічного дерева родини повинна відповідати вимогам та виконувати такі функції, як:

* реєстрація нового користувача;
* вхід зареєстрованого користувача;
* створення генеалогічного дерева для користувача;
* додавання родичів до дерева користувача;
* можливість доручення до друзів;
* можливість дозволу перегляду дерева іншим користувачам;
* автоматичне формування генеалогічного дерева;
* ведення статистичного аналізу даних по родичам;
* можливість копіювання інформації із інших дерев;
* додавання нових генетичних захворювань;
* можливість знаходження ймовірності захворювання спадковою хворобою нащадка;

Також вся система повинна працювати завдяки Web-API, адже це дозволить поширити її на всі існуючі платформи. Всі дані які повертатимуться із сервера повинні бути у форматі JSON. Для забезпечення захисту особистих даних користувача, потрібно реалізувати шифрування його паролю для входу.

Для правильного визначення ймовірності прояву спадкових захворювань необхідно вказати достовірну інформацію про хворобу таку, як можливість передачі її жінкам або чоловікам, а також її форму прояву у пращура домінантна чи рецесивна.

## 1.5 Висновки до першого розділу

В даному розділі було проаналізовано методи та засоби для розробки якісного програмного продукту на тему «Інформаційна система «Генеалогічне дерево родини»». Було вивчено конкретну область, для врахування особливостей, якими буде володіти дана інформаційна система.

Проаналізувавши опрацьовану літературу, можна сказати, що генеалогія та генетика посідає досить значне місце у розвитку сімейної історії. Використання даного програмного продукту дозволить скласти умови для більш ефективного проведення збору інформацій про власну родину.

У першому розділі розглянуто аналоги для розроблюваної інформаційної системи. Існує велика кількість програмного забезпечення для побудови генеалогічного дерева родини, але немає такої системи , яка передбачає наявність у осіб певного спадкового захворювання. Тому розробка даної інформаційної системи є досить актуальна.

Також під час написання першого розділу було проаналізовано фактичний стан задачі, що розв’язується, на основі матеріалів, які характеризують об’єкт проектування; містить теоретичний виклад важливих аспектів проблеми, що досліджується, та критичний огляд наукової літератури і періодичних джерел з визначеної тематики.

Сформовано технічне завдання, в якому передбачено всі переваги та недоліки існуючих аналогів. Та включені ті функції, які не було знайдено у аналогів, що і робить даний програмний продукт унікальним.