

МЕТОДИ ТА ЗАСОБИ СТРУКТУРУВАННЯ І ЗБЕРІГАННЯ ДАНИХ В ЕЛЕКТРОННИХ МЕДИЧНИХ КАРТКАХ

© Завалій Т.І., Нікольський Ю.В., 2010

Описано інформаційну систему “Імуноскрин 2.0”, призначену для централізованого збирання і зберігання медичних даних у галузі імунології. Подано основні принципи проектування і розроблення системи. У системі використано спеціальну OWL-онтологію для структурування даних, автоматизації введення інформації та створення звітів.

Ключові слова: накопичення даних, онтологія, електронна медична картка.

The authors describe “Imunoskrin 2.0” web-based information system for centralized collecting and storing medical data in the immunology domain. Design principles and implementation are briefly described. The system uses OWL-ontology for data structuring, automated data acquisition and reports creation.

Keywords: data acquisition, ontology, EMR.

Вступ

Актуальною є потреба розробляти масштабні багатовимірні моделі та приймати рішення на основі ретроспективних даних, накопичених у предметній області. Забезпечення якості, відкритості та повторного використання даних потребує багатьох зусиль. В умовах постійних змін предметної області (ПО), розроблення технологій збирання та інтегрування великих обсягів потенційно суперечливих і неповних даних є актуальним завданням. Розв'язання цієї задачі необхідне для побудови сховищ даних та консолідованих інформаційних ресурсів у будь-якій галузі знань. Це дасть змогу агрегувати великі обсяги даних з метою їх статистичного та інтелектуального аналізу. Як показує досвід, результати такого аналізу значною мірою залежать від якості й кількості первинних даних. Покращення якості первинних даних дасть змогу поліпшити якість рішень, прийнятих на їх основі.

У статті розглянуто задачу централізованого накопичення структурованих даних з метою їх подальшого аналізу. У пропонованому підході онтологія предметної області задає структуру даних, які зберігаються у електронній картці пацієнта. Розроблено програмні засоби збирання та зберігання даних, які використовують в імунологічних відділеннях для обліку і лікування пацієнтів, формування звітності.

Аналіз останніх досліджень

У галузі аналізу даних можна виділити проблематику, якій раніше не приділялось достатньої уваги – накопичення великих обсягів даних в умовах постійних змін у предметній області. Виникають задачі отримання, структурування та інтегрування даних з локальних чи віддалених джерел, розроблення анкет для отримання даних, розроблення засобів автоматизованого введення даних у базу даних, способів перетворення та інтегрування даних, формування таблиць даних для застосування методів аналізу. Ці задачі виникають на перших етапах процесу аналізу даних – збирання, відбирання та перетворення даних. Результатом є сформована таблиця з даними про об'єкти предметної області.

Первинне накопичення медичних даних відбувається у медичній картці пацієнта, журналах консультацій, скеруваннях на обстеження тощо. Електронну медичну картку (ЕМК) опишемо як сукупність цифрових медичних даних про окремого пацієнта. Така картка може містити різнотипні дані – персональну інформацію, історію хвороби, призначені ліки, відмітки про вакцинування,

результати лабораторних тестів, рентгенівські знімки тощо. До світових галузевих стандартів належать ISO 18308, CEN 13606, HL7, DICOM, openEHR й такі спеціалізовані онтології, як MeSH, SNOMED CT, UMLS, ICD-10. В Україні єдиного затвердженого стандарту електронної медичної картки немає. Тому для зберігання медичної інформації необхідно використовувати максимально гнучкі структури даних. Перелічимо основні проблеми впровадження ЕМК, які зумовлюють актуальність досліджень:

- дані надходять з різних джерел;
- дані збирають за різними методиками;
- виникають помилки під час введення даних;
- набір потрібних ознак не фіксований і може змінюватись з часом.

Аналіз публікацій показує як актуальні проблеми в інформатизації медичної галузі, так і наявність успішних прикладів впровадження комп'ютерних технологій. Концепцію ЕМК реалізовано у різних комерційних та науково-дослідних розробках. У медичній інформаційній системі Doctor Eleks [1] впроваджено світові стандарти для подання медичних даних. У ній використано єдину медичну термінологію та деревоподібну схему введення даних на основі словесних шаблонів. Всі дані медичних обстежень пацієнта об'єднані в єдину картку і збережені в базі даних. У [2] описано досвід впровадження та переваги електронних карток у країнах, що розвиваються. Зокрема, наведено приклад веб-системи РІН-EMR для обліку хворих за програмою лікування стійких форм туберкульозу. Завдяки впровадженню системи зменшилась кількість помилок під час введення даних (з 17 % до 3 %), спрощено формування звітності, лікарям надано прості засоби аналізу і перевірки даних, введено уніфікований реєстр лікарських препаратів. Загальна кількість пацієнтів на обліку становила 4300. Варто згадати і про такі масштабні веб-сервіси, як Google Health, Microsoft HealthVault і Dossia, які забезпечують користувачів засобами введення, редагування, імпорту та експорту електронних медичних карток. Огляд та класифікацію медичних інформаційних систем, аналіз їх архітектури та вимог до підсистем наведено у [3, 4]. У [5] описано експертну систему діагностування імунодефіцитних станів. У системі використано статистичні методи для аналізу даних про температуру ділянок тіла.

Упровадження ЕМК вимагає розроблення спеціальних методів подання та опрацювання структурованих даних. У працях [6, 7] описано підходи до збирання та подання ієрархічно організованої інформації на прикладі накопичення даних про музейні експонати та про пацієнтів імунологічних відділень. Авторами здійснено структурування предметних областей, реалізовано фреймову модель подання даних з використанням наслідування класами-нащадками властивостей базового класу, наведено схему реструктурування інформації. Об'єкти предметної області та ієрархічні зв'язки між ними описують засобами середовища ZOPE і зберігають у файлах спеціального формату. У роботі [8] запропоновано подання структурованих даних за допомогою вкладених ненормалізованих реляційних відношень. Показано, що такі відношення утворюють ядро сховища даних, а опрацювання даних можливе стандартними операціями над гіперкубом даних. Як приклад, розглянуто опрацювання даних з медичного паспорту пацієнта. Для подання даних авторами використано 13 ненормалізованих відношень, описано атрибути та їхні домени; для статистичного аналізу даних розроблено запити на побудову перехресних таблиць. У медичній галузі часто застосовують іншу модель зберігання даних – EAV (Entity attribute value) [9]. Її використовують під час побудови баз даних, призначених для зберігання інформації про об'єкти з великою кількістю властивостей, які можуть змінюватись. У цій моделі інформацію про об'єкти, їхні атрибути та значення атрибутів зберігають в окремих таблицях. Для зв'язку типу “багато до багатьох” потрібна четверта таблиця, яка містить поля з номерами сутностей, номерами атрибутів та номерами їх значень. Такий підхід дає змогу створювати гнучку та нормалізовану базу даних. Недоліками є істотне ускладнення SQL-запитів на вибірку та неможливість зберігати структуру атрибутів і зв'язки між ними.

Перспективним є підхід, у якому онтологія предметної області задає структуру даних, є семантичним прошарком для доступу до баз даних і використовується у процесах аналізу даних

[10, 11, 12]. У роботі [13] досліджено різні формати подання метаданих та створено онтологію для опису медичних інформаційних ресурсів. Онтологію створено на базі Дублінського ядра (Dublin Core) та класифікації хвороб ICD-9-CM, додано можливість фіксації географічної та персональної інформації про автора. У праці [14] досліджено можливість застосування медичних онтологій для автоматизації основних клінічних задач, наприклад, відбору пацієнтів для клінічних досліджень, моніторингу інфекційних захворювань, підтримки прийняття рішень. У роботі використано онтологію SNOMED CT і річні дані про 240 тис. пацієнтів. У проекті NCI Thesaurus [15] Американського інституту раку побудовано OWL-онтологію обсягом десятки тисяч класів і десятки відношень з метою структурувати онкологічні інформаційні ресурси. У праці [16] описано впровадження інформаційної системи в імунології для централізованого накопичення інформації про пацієнтів з рідкісним червоним вовчаком. Засобами системи проаналізовано дані 100 пацієнтів з п'яти болгарських лікарень.

Існує декілька підходів до накопичення даних з використанням онтології. За першим дані зберігаються у базі знань або у вигляді RDF- чи OWL-файлів (Resource Description Framework та Ontology Web Language, стандарти W3C) і запити на вибірку здійснюються засобами спеціальної мови (наприклад, SPARQL). За другим підходом онтологія ПО відображається у структурі реляційної або об'єктно-орієнтованої бази даних і її опрацьовують засобами СУБД. RDF-модель напряму може відображатись у реляційній моделі [17]:

1. Рядок таблиці відповідає об'єкту в RDF-моделі.
2. Назва стовпця таблиці є властивістю об'єкта (propertyType).
3. Значення поля є значенням властивості.

Таке відображення не є ефективним за наявності великої кількості властивостей та якщо потрібно враховувати структурні залежності між ознаками об'єктів. Зміна множини властивостей неминує веде до зміни схеми бази даних. Тому розроблення нових способів зберігання даних зі змінною структурою є актуальним завданням досліджень. Вирішивши його, можна будувати відкриті та гнучкі інформаційні системи, зокрема і для опрацювання електронних медичних карток.

Постановка задачі

Авторами вирішено проблему структурування й накопичення ретроспективних даних у предметній області завдяки впровадженню єдиної інформаційної технології. Розроблено інформаційну систему для створення й обліку електронних карток консультацій пацієнтів імунологічного відділення. Особливість задачі полягає в організації такого процесу збирання та опрацювання даних, який би забезпечував легку масштабованість, можливість модифікації структури ознак, простоту використання, стійкість до помилок, анонімність пацієнтів. Для досягнення такої мети розв'язано задачі структурування предметної області та формування онтології ознак; розроблення веб-орієнтованої архітектури інформаційної системи; проектування та розроблення онтологічної компоненти та форм уведення даних; розроблення, тестування і впровадження інформаційної системи у лікарську практику.

Основний матеріал

Першим етапом усього процесу аналізу є збирання даних, під час якого виникають проблеми неповноти та неточності даних. У предметній області з часом можливі зміни, тому структура даних не є сталою. Виникає потреба у гнучкій структурі даних, яка дасть можливість організувати тривале накопичення і подальше перетворення даних. У процесі накопичення збирають дані про певні групи ознак об'єктів з предметної області. Ці ознаки можна подати у вигляді деревоподібної структури зі встановленими між ними семантичними зв'язками.

Якщо припустити, що в кожному джерелі даних використовують свою онтологію предметної області, то задача інтегрування цих даних зводиться до задачі інтегрування гетерогенних онтологій [18]. Синтаксично гетерогенні онтології описані різними мовами та потребують додаткової трансляції при об'єднанні. Термінологічно гетерогенні онтології містять різні назви тих самих понять. Це пов'язано з використанням різних мов і діалектів, технічних термінологій, синонімів тощо. Семантично гетерогенні онтології побудовані на основі різних підходів до моделювання

понять і використовують для цього різні аксіоми. Причинами семантичної гетерогенності можуть бути:

- різна область покриття, коли дві онтології описують різні частини предметної області, які можуть перетинатись, але не збігаються;
- рівень деталізації (гранулювання) при описі понять;
- різні погляди на поняття ПО за однакового покриття та деталізації. Наприклад, політична і геологічна карти світу.

Існує і семіотична гетерогенність, яка виникає через різну інтерпретацію онтології і контексту окремих понять людиною. Надалі розглядатимемо лише термінологічну та семантичну гетерогенність, зокрема проблеми різних областей опису та рівнів деталізації понять. Вирішення цих проблем забезпечить успішність процесу структурування і накопичення даних, уможливить створення відкритих інформаційних систем, у яких онтологія є семантичним прошарком для доступу до даних.

Структурування даних

Для уникнення проблем гетерогенності на етапах збирання та попереднього опрацювання даних пропонуємо використовувати онтологію предметної області, у якій відображено структуру ознак, дані яких буде внесено у таблицю даних T (1). Найпростішим випадком такого підходу є використання словників даних. Кожна ознака об'єкта може набувати лише наперед узгоджених та фіксованих значень зі словника. Так усувають проблему механічних помилок під час введення даних в електронному вигляді. У складнішому випадку в предметній області будуємо повноцінну онтологію, що дасть змогу враховувати семантичні зв'язки між ознаками об'єктів. Наявність прихованих причинно-наслідкових зв'язків між ознаками може спричинити надлишковість даних у таблиці. Дані у цю таблицю заносять поступово, протягом тривалого часу, або за допомогою експерта відбирають з різних джерел. Зі зміною експерта чи його поглядів змінюється набір ознак, і, відповідно, структура таблиці даних. Таблицю даних визначено як

$$T = (U, A \cup \{d\}), \quad (1)$$

де $U = \{u_1, u_2, \dots, u_n\}$ – непорожня скінченна множина об'єктів, кожен з яких поданий кортежем таблиці; $A = \{a_1, a_2, \dots, a_m\}$ – непорожня скінченна множина умовних атрибутів, які є описом ознак об'єктів; d – класифікаційний атрибут, за значенням якого об'єкт зараховано до певного класу [19].

Визначимо онтологію O як трійку

$$O = \langle X, R, I \rangle, \quad (2)$$

де X – множина понять предметної області, R – множина відношень на елементах множини X , I – множина функцій інтерпретації (аксіом), заданих на множинах X і R .

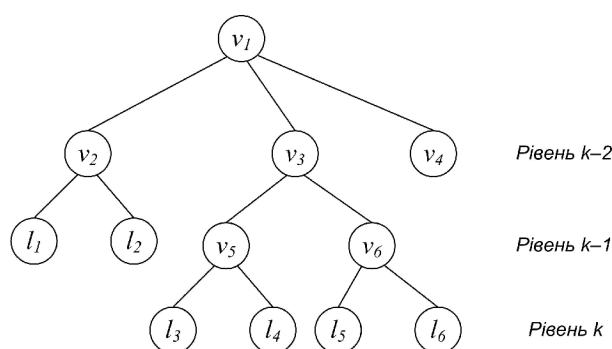


Рис. 1. Древоподібне подання структури понять в онтології

Структура понять предметної області в загальному випадку має ієрархічний вигляд, тому структуру онтології можна подати у вигляді дерева

$$S = \langle V, L, E \rangle, \quad (3)$$

у якому множина вершин V відповідає класам онтології, множина листків L – екземплярам класів, множина ребер E – відношенням наслідування (наприклад, клас-підклас, або частина-ціле). Введемо рівень деталізації $k \in N$ як міри віддаленості вершини від кореня дерева (рис. 1).

Визначимо множину дій над деревом S , у результаті яких утворюється нове дерево S' . Ці дії виконуються над структурою дерева з метою його трансформації або для узгодження з іншим деревом [20]. Такими операціями є:

- 1) додавання нової вершини дерева;
- 2) видалення вершини дерева;
- 3) заміна батьківської вершини (переміщення гілки);
- 4) вилучення гілки дерева;
- 5) додавання гілки дерева;
- 6) перейменування вершини дерева;
- 7) об'єднання вершин.

Ці операції реалізовано засобами процедурної мови програмування для розробленої схеми бази даних (рис. 6). Зазначимо, що операція об'єднання двох вершин зводиться до їх перейменування і переміщення гілок з підпорядкованими вершинами. Підмножина вершин і листків дерева S відповідає множині атрибутів таблиці даних: $V \cup L = A$. Якщо структура та набір ознак змінюється в часі, то можна говорити про введення нової версії онтології. Онтології, які використовують в різних джерелах, теж можуть відрізнятися. Тому в загальному випадку вестимо мову про множину онтологій

$$U_O = (O_1, O_2, \dots, O_i), i=|U_O|, \quad (2)$$

які задають структуру даних у кожному джерелі даних і відображають зміни цієї структури в часі. Кожній онтології O_i поставимо у відповідність таблицю даних T_{O_i} , яка формується відповідно до структури онтології (табл. 1). Атрибут a_m таблиці даних T_{O_i} має домен D_m . У загальному випадку такий атрибут є двійковим, тобто домен $D_m = \{0, 1\}$. Значення атрибута відображає наявність (1) чи відсутність (0) в об'єкта ознаки a_m . За великої кількості ознак таблиця з двійковими значеннями стає розрідженою. Тому пропонуємо формувати множину атрибутів з ознак $k-1$ рівня ієрархії, а домен атрибута – з множини підпорядкованих вершин рівня k разом зі значенням "Null":

$$A = \{a_m=l \mid l \in L_k\}, D_m = \{l \mid l \in L_{k-1}\} \cup \text{Null}. \quad (3)$$

Значення "Null" позначає відсутність у об'єкта певної ознаки. За такого підходу дані у таблицю можна відбирати з різних рівнів ієрархії, формуючи навчальні вибірки з різною множиною ознак. За потреби, дані з різною структурою ознак можна інтегрувати у єдину таблицю даних.

Таблиця 1

Таблиця з ієрархічною структурою атрибутів (див. рис. 1)

	v_1			
	v_2	v_3		v_4
		v_5	v_6	
u_1	$\{l_1, l_2, \text{Null}\}$	$\{l_3, l_4, \text{Null}\}$	$\{l_5, l_6, \text{Null}\}$	[0, 1]
...				
u_n				

Архітектура інформаційної системи

Задача збирання, тривалого зберігання та опрацювання даних з карток пацієнтів імунологічних відділень має такі обмеження:

- кількість ознак, інформацію про які фіксуємо, є дуже великою (>1000);
- структура ознак ієрархічна і може змінюватись в часі;
- потрібно фіксувати зміну значень ознак у часі;
- ознаки набувають значення символічного типу.

Схему бази даних та архітектуру веб-орієнтованої інформаційної системи “Імуноскрин 2.0” спроектовано з використанням онтології ознак як її складової частини. Опишемо процес проектування, концептуальну та логічну моделі ІС, модулі системи.

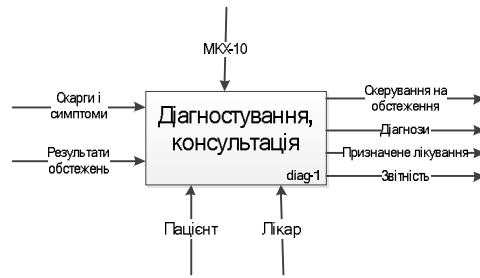


Рис. 2. Контекстна діаграма предметної області

Концептуальна модель предметної області “Імунологія” – це онтологія верхнього рівня, схема головних понять та зв’язків між ними (рис. 3). До неї входять такі поняття, як Лікар, Пацієнт, Установа, Діагноз, Препарат, Обстеження, Місце. Дослідження процесів опрацювання даних у відділенні медичної установи дало змогу виділити такі групи ознак, інформацію про які потрібно збирати: загальні дані про пацієнта, дані про скерування, дані діагностичних обстежень, виявлені відповідно до МКХ-10 хвороби, призначене лікування та його ефект. У результаті тривалої роботи лікарів-імунологів сформовано ієрархію ознак, що складається з чотирьох частин: “Скерування та обстеження”, “Діагноз імунодефіциту”, “МКХ-10” та “Лікування”.

Побудову OWL-онтології здійснено у редакторі Protege [21]. Цей редактор є модульним і підтримує різні мови моделювання – Clips, UML, XML, RDF та OWL. Процес створення онтології мав п’ять етапів:

1. Визначено основні поняття галузі та надано їм імена. Основні поняття – це *Особа*, *Установа*, *Діагноз*, *Препарат*. Другорядні поняття – *Обстеження*, *Місце* (адреса пацієнта чи розташування медичного закладу).

2. Кожне поняття утворює клас, який або поділений на підкласи, або містить набір екземплярів класу. На цьому етапі задано ієрархію класів предметної області (рис. 3). Наприклад, *Особа* → *Лікар* → *Імунолог*, або *Діагноз* → *МКХ-10* → *Імунодефіцити*. Для визначення підкласів класу *Діагноз* використано міжнародну класифікацію хвороб МКХ-10 [22].



Рис. 3. Фрагмент онтології. Ієрархія класів зі зв’язками “is_a”

3. Визначено властивості класів. Властивість в OWL – це бінарне відношення. Наприклад, Особа “проживає у” Місце, Пацієнт “приймає” Препарат. Також задано такі характеристики відношень, як рефлексивність, транзитивність, симетричність тощо.

4. Задано обмеження на деякі типи даних. Зокрема, визначено кардинальність властивостей, діапазони значень.

5. Введено екземпляри класів. Наприклад, назви окремих препаратів (“Біовен”, “Веноіmun”, “Аміксил”) чи точні діагнози (“поствірусний синдром втоми”, “синдром стійкої гіпертермії”, “саркоїдоз”).

На цьому етапі онтологія є спрощеною, оскільки в основній її частині є залежності тільки типу клас-підклас, не містить складних класів та заперечень понять. Всі можливості OWL-DL не реалізовано. Наприклад, на цьому етапі можна відобразити лише факт захворювання пацієнта, і не можна – встановлений факт відсутності хвороби. Загальна кількість ознак в онтології – понад 1000. Описана структура онтології повністю відповідає структурі, закладеній в обліковій картці імунологічного хворого – стандартній формі медичної документації, затвердженій МОЗ України.

Розглянемо діаграму потоків даних імунологічного відділення (рис. 4). Лікар, що оглядає пацієнта, ставить діагнози, виписує скерування на обстеження, призначає лікування, заносить основну інформацію у журнал консультацій. Наступний крок – це заповнення для кожної консультації облікової картки спеціальної форми, розробленої на підставі стандарту МКХ-10. Сукупність облікових карток одного пацієнта утворює його історію хвороби і використовується для обчислення кількісних показників діяльності лікаря та установи загалом. До таких показників належать кількість обстежень сільських і міських жителів, кількість скерувань, окремих діагнозів та призначень окремих препаратів, кількість вакцинацій тощо. Загальна кількість показників у кварталному і річному звітах – 178.



Рис. 4. Діаграма потоків даних предметної області

Після впровадження інформаційної системи процес обліку пацієнта спрощено (рис. 5). Лікар здійснює огляд і діагностування пацієнта, виписує скерування, призначає лікування, заносить основну інформацію у журнал консультацій. Інформацію з журналу консультацій вводять у базу даних через веб-інтерфейс інформаційної системи. Засоби інформаційної системи дають змогу автоматично формувати облікову картку пацієнта, історію хвороби та звіт за потрібний період.



Рис. 5. Діаграма потоків даних після упровадження інформаційної системи

Збереження онтології у реляційній базі даних

Для зберігання ієрархічної структури онтології у базі даних використано традиційне подання графів списком суміжних вершин (adjacency list model). У цьому випадку схема відношення, що містить ієрархію, є такою $R=(id, label, parent)$. Ця схема, вперше запропонована Е. Коддом, з деякими модифікаціями широко застосовується у реляційних базах даних [23]. Подання графа методом вкладених множин (nested set model) забезпечує швидкий пошук підпорядкованих вершин, але й має недоліки. Вершини графа нумеруються двічі методом обходу в глибину з поверненням, тому зміна структури графа, додавання та видалення вершин вимагає повторного нумерування всіх вершин.

У пропонованому підході відображення OWL-онтології предметної області у реляційну модель бази даних відбувається у такий спосіб:

1. Кожне поняття з онтології позначене міткою l , яка збережена у таблиці понять (R_1) та має унікальний ідентифікатор.
2. Інтерпретація відношень “is_a” (ієрархія понять) зберігається у таблиці таксономії (R_2).
3. Аналогічно, інтерпретація інших бінарних відношень на множині понять зберігаються в окремій таблиці.

Мітки, що відповідають поняттям онтології, зберігаємо в таблиці $R_1=(term_id, label)$. Ієрархію міток зберігаємо у таблиці $R_2=(term_id, parent_id)$ (рис. 6). У цьому разі обидві таблиці є нормалізованими.

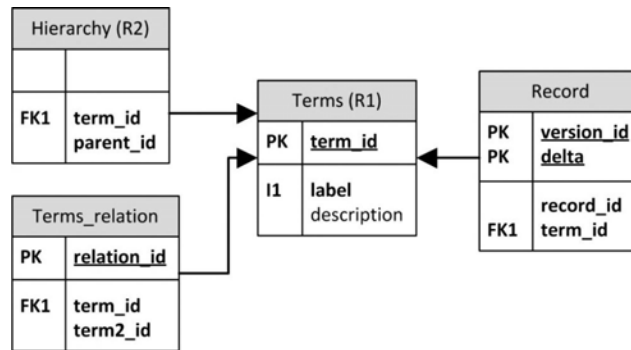


Рис. 6. Схема бази даних для зберігання електронних карток пацієнтів

Наведемо приклад запиту на вибірку множини L листків дерева (екземплярів класів):

```
SELECT label FROM  
(R2 AS t1 LEFT JOIN R2 AS t2 ON t1.term_id = t2.parent_id)  
INNER JOIN R1 ON t1.term_id = R1.term_id  
WHERE t2.term_id IS NULL;
```

Додавання і видалення вершин засобами мови SQL є тривіальною задачею. Складніші операції, такі як видалення термінів з різних рівнів ієрархії, модифікація структури дерева тощо, виконують засобами процедурної мови програмування. В базі даних всі мітки мають символічний тип. Можливе подальше розширення системи для роботи з іншими типами даних. Наприклад, додаткова мітка може містити адресу зображення чи файла, збереженого у файльовій системі.

Структурна схема ІС

Інформаційна система складається з бази даних та чотирьох модулів – введення інформації, формування звітів, редагування онтології та адміністрування (рис. 7). Користувач системи через веб-інтерфейс взаємодіє з модулями ведення даних та формування звітності. Адміністратор системи через веб-інтерфейс працює з редактором онтології. Цей модуль призначений для створення й редагування онтології ознак (структури даних) і таксономії звіту (структури полів у підсумковому звіті). Підсистема адміністрування потрібна для надання прав доступу, керування користувачами та введеними картками, архівування даних. База даних містить службову інформацію, дані з карток пацієнтів, онтологію ознак та таксономію полів підсумкового звіту.

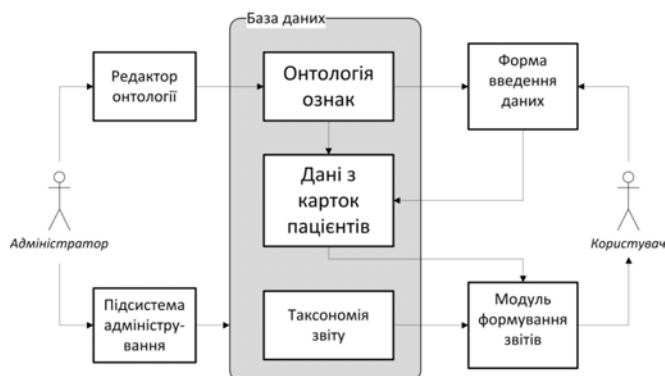


Рис. 7. Структурна схема інформаційної системи

Інформаційну систему реалізовано на основі відкритої системи керування контентом (CMS) із застосуванням мов PHP, JavaScript, HTML та СУБД MySQL. Використання API-функцій та програмних модулів CMS спростило побудову веб-інтерфейсу системи, реалізацію підсистеми адміністрування та модуля формування звітів. Функції системи:

1. Введення та редагування медичних карток пацієнтів.
2. Прослідковування історії хвороби в часі.
3. Пошук у базі даних за ключовими словами.
4. Формування звітних даних за вибраний період.
5. Розподіл прав доступу та забезпечення анонімності даних.

6. Адміністрація розробленої інформаційної системи має можливості: 1) створення нових користувачів та медичних установ; 2) редагування онтології ознак; 3) налаштування інтерфейсу для введення даних окремо для кожної установи; 4) редагування прав доступу до інформації; 5) відстеження активності різних груп користувачів.

7. Користувачі системи мають змогу: 1) створювати та редагувати нові записи про огляд пацієнтів; 2) вносити зміни в окремі частини онтології ознак; 3) формувати звіти потрібної форми за будь-який період часу; 4) простеження зміни стану пацієнта; 5) здійснювати пошук за ключовими словами.

Введення інформації в електронну картку пацієнта виконують на єдиній веб-сторінці з використанням різних методів введення (рис. 8).

Результати впровадження

Опишемо результати впровадження системи у лікарську практику на основі дворічного досвіду експлуатації на кафедрі клінічної імунології та алергології Львівського медичного університету ім. Д. Галицького. Було створено п'ять користувачів з правами лікаря та п'ять користувачів з правами медсестри. Медсестри здійснювали введення інформації про пацієнтів відповідного лікаря. Під час експлуатації системи виникла необхідність незначної модифікації онтології ознак. Зокрема, розширено розділ "Лікування" (додано нові препарати). Змінено розташування в ієрархії терміна "Риб'ячий жир". Зміни в онтології ознак автоматично відобразились у раніше введених картках пацієнтів.

Вибором веб-орієнтованої архітектури вдалося мінімізувати проблеми з упровадженням системи в установі, де відсутній допоміжний технічний персонал, немає внутрішньої корпоративної комп'ютерної мережі. Для доступу до інформаційної системи необхідне лише комп'ютеризоване робоче місце з під'єднанням до Інтернету і сучасний веб-браузер з підтримкою JavaScript. Обслуговує систему один адміністратор у віддаленому режимі. До його обов'язків належить підтримка користувачів і розподіл прав доступу, періодичне резервне копіювання даних, періодичне оновлення системи.

Проблеми впровадження пов'язані з навчанням персоналу, усуненням помилок та забезпеченням даних. Зокрема, існує проблема використання користувачами паролів недостатньої складності та безпечного зберігання особистих даних пацієнтів. Основні помилки, що виникають

під час введення даних, пов'язані з місцем проживання пацієнта. Часто користувачі, не користуючись динамічними підказками з вибору термінів, неправильно або не повністю вводили місто, район та область. Наприклад, “м Львів” замість “м. Львів”, “Ів.Франківська обл” замість “Івано-Франківська обл.” тощо. Помилки, пов'язані з неправильним введенням дат консультацій, неправильним вибором термінів з дерева, помилками в прізвищах виникають рідко і виправляються користувачами системи самостійно. Наведемо кількісні результати експлуатації системи:

- середня швидкість введення картки – 1.5 хв;
- разом введено карток пацієнтів – 556;
- разом консультацій – 1089;
- об'єм бази даних карток – 2 мб;
- об'єм службової бази даних – 80 мб.

Рис. 8. Комбінований інтерфейс для введення інформації у картку

Головним результатом впровадження системи Імуноскрин 2.0 стала інформатизація обліку пацієнтів і повна автоматизація підрахунку кількісних показників діяльності установи. Як наслідок, значно спрощено процес формування кварталних і річних звітів за вимогами Міністерства охорони здоров'я. Тестова експлуатація системи дала змогу усунути виявлені недоліки та підготувати систему до швидкого впровадження у будь-якому імунологічному відділенні України.

Висновки

У статті описано проектування та створення інформаційної системи для накопичення даних у галузі імунології. Система призначена для централізованого введення даних з карток пацієнтів, тривалого обліку пацієнтів, відстеження історії хвороби, формування звітності. Діаграми потоків

даних до і після впровадження інформаційної системи демонструють спрощення процесу обліку пацієнтів. Онтологію предметної області використано як компонент системи. Розроблено спосіб подання онтології ознак у реляційній базі даних. Автори планують подальші дослідження зібраних медичних даних методами інтелектуального аналізу даних.

1. Качмар В.О. Електронна медична карта пацієнта. Взаємосумісність та стандартизація / В.О. Качмар, А.І. Хвищун // *Український журнал телемедицини та медичної телематики*. – Т.6, №1. 2008. – С. 76–79.
2. Fraser H. Implementing electronic medical record systems in developing countries. / H. Fraser, P. Biondich, D. Moodley, S. Choi, B. Mamlin, P. Szolovits. // *Informatics in Primary Care. British Computer Society*. #13. 2005. р. 83–95.
3. Мельникова Н.І. Аналітичний огляд засобів програмного забезпечення в медичній галузі. / Мельникова Н.І., Шаховська Н.Б. // *Вісник Нац. ун-ту “Львівська політехніка”: Інформаційні системи та мережі*. – № 673. – Львів, 2010. – С. 146–154.
4. Цмоць І.Б. Принципи побудови та базові компоненти медичної інформаційної технології. / І.Б. Цмоць, А.Є. Батюк, С.І. Пилипчук, С.А. Батюк // *Матеріали конф. “Комп’ютерні науки та інформаційні технології”*. – (Львів, 2009 р.) / – Вежа і Ко. – 2009. – С. 346–348.
5. Нечипуренко О. Використання комп’ютерних технологій для діагностики і профілактики імунодефіцитних станів людини // *Матеріали конф. “Комп’ютерні науки та інформаційні технології”*, – (Львів, 2009 р.) / – Вежа і Ко. – 2009. – С. 322–325.
6. Войчишин В.К. Інформаційний портал Львівського державного природознавчого музею / В.К. Войчишин, О.С. Климишин, Є.Я. Лециньський, Ю.В. Нікольський, Ю.М. Чорнобай. // *Вісник Нац. ун-ту “Львівська політехніка”: Інформаційні системи та мережі*. – № 549. – Львів, 2005. – С.44–54.
7. Лециньський Є.Я. Моделювання процесу збирання та подання ієрархічно організованої інформації на прикладі спостережень за імунологічними пацієнтами / Є.Я. Лециньський // *Вісник Нац. ун-ту “Львівська політехніка”: Інформаційні системи та мережі*. – № 519. – Львів, 2004. – С. 214–225.
8. Григорович А.Г. Система опрацювання медичних даних “Дільничий терапевт” / А.Г. Григорович, В.Г. Григорович, В.В. Пасічник // *Вісник Нац. ун-ту “Львівська політехніка”: Інформаційні системи та мережі*. – № 653. 2009. – С. 71–102.
9. Carl Welch. EAV modeling. / Carl Welch // *PHP Architect. Volume 7, Issue 6*. 2008. р.25–33.
10. Data mining with ontologies: Implementations, Findings, and Frameworks. / Eds. H. Nigro, S. Cisaró, D. Xodo. – IGI Global. – 2008. – 311 p.
11. D. Calvanese. Ontology-based database access / D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, A. Poggi, and R. Rosati // *Proc. of SEBD–2007*. 2007. p. 324–331.
12. Calvanese D. Data integration through DL-LiteA ontologies / D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini et al. // K.-D. Schewe and B. Thalheim, editors. *Semantics in Data and Knowledge Bases: Postproc. of the Int. Workshop SDKB–2008*. 2008. – p. 26–48.
13. Kamel Boulos M.N. Towards a semantic medical web: HelthCyberMap's Dublin Core ontology in Protege-2000. / Kamel Boulos M.N., Roudsari A.V, Carson E.R. Kamel Boulos M.N., Roudsari A.V, Carson E.R. // *Medical Science Monitor*, July 8(7). 2002. p.124–136.
14. Patel Ch. Matching Patient Records to Clinical Trials. / Ch. Patel, J. Cimino, J. Dolby et al. *Proceedings of the 6th International Semantic Web Conference ISWC/ASWC2007, Vol. 4825 – Heidelberg: Springer-Verlag, 2007, p. 809–822*.
15. Golbeck J. The national cancer institute's thesaurus and ontology / J. Golbeck, G. Fragosó, F. Hartel, J. Hendler, B. Parsia, J. Oberthaler // *Journal of Web Semantics*, №1(1). – Elsevier, 2003. p. 75–80.
16. I. Penjurov. MIS Immunolog – Practical Results and Conclusions / I. Penjurov // *Український журнал телемедицини та медичної телематики*. – 2009. – Т.7, № 1. – С. 71–72. Режим доступу: <http://telemed.org.ua/journ/archive.html>.
17. Berners-Lee T. What the Semantic Web can represent (The Semantic Web and Relational Databases) / Berners-Lee, T. // [Електронний ресурс] Режим доступу: <http://www.w3.org/Designissues/rdfnot.html> 17.09.2010.
18. J. Euzenat, P. Shvaiko. *Ontology matching*. – Berlin: Springer-Verlag, 2007. – 333 p.
19. Komorowski J., Pawlak P., Polkowski L., Skowron A. *Rough Sets: A Tutorial*. [авт. книгу] Eds. S.K.Pal and A. Skowron. *Rough Fuzzy Hybridization: A New Trend in Decision-Making*. Singapore : Spriner-Verlag, 1998, сс. 3-98.
20. Нікольський Ю.В. Моделі даних у структурах, що трансформуються / Ю.В. Нікольський // *Вісник Нац.ун-ту “Львівська політехніка”: Інформаційні системи та мережі*. – № 519. – 2004. – С. 233–243.
21. Protégé [Електронний ресурс]. Режим доступу: <http://protege.stanford.edu>. 29.09.20010. – Заголовок з екрана.
22. *International Classification of Diseases (ICD)* [Електронний ресурс]. Режим доступу: <http://www.who.int/classifications/icd>. 22.10.20010. – Заголовок з екрана.
23. Joe Celko. *Trees and Hierarchies in SQL for Smarties*. – Morgan Kaufmann Publishers. – 2004. – 240 p.