

4. ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ



Науковий вісник НЛТУ України
Scientific Bulletin of UNFU

<http://nv.nltu.edu.ua>

<https://doi.org/10.15421/40270525>

Article received 22.06.2017 р.

Article accepted 29.06.2017 р.

УДК 681.518:004.93.1

ISSN 1994-7836 (print)

ISSN 2519-2477 (online)

✉ Correspondence author

O. M. Berezsky

ob@tneu.edu.ua

О. М. Березький¹, О. Й. Піцун¹, С. О. Вербовий¹, Т. В. Дацко²

¹ Тернопільський національний економічний університет, м. Тернопіль, Україна

² Тернопільський національний медичний університет ім. І. Я. Горбачевського, м. Тернопіль, Україна

РОЗРОБЛЕННЯ РЕЛЯЦІЙНОЇ БАЗИ ДАНИХ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ АВТОМАТИЗОВАНОЇ МІКРОСКОПІЇ

Проаналізовано сучасні технології проектування розподілених баз даних, що дало змогу виділити їх переваги та недоліки. Розроблено даталогічну модель бази даних, що дозволило здійснювати зберігання та обмін інформацією між користувачами системи автоматизованої мікроскопії для аналізу гістологічних і цитологічних зображень. Базовими функціями розробленої системи є такі: облік пацієнтів, користувачів, автоматичне оброблення зображень, підрахунок та зберігання кількісних та якісних характеристик мікрооб'єктів, механізми комунікації між користувачами, класифікація зображень тощо. Більшість наявних систем для опрацювання будь-якого типу зображень володіють тільки засобами для аналізу і не мають у своєму складі механізмів для зберігання, обміну інформації. Перевагою розробленої інтелектуальної системи автоматизованої мікроскопії є наявність модулів для збереження кількісних та якісних характеристик мікрооб'єктів. Застосування "master-slave" реплікації дає змогу розподілити навантаження між наявними серверами та здійснити резервне копіювання даних, що є незамінним у медичних системах. Спроектвана даталогічна модель забезпечує мінімальне дублювання інформації та відповідає вимогам нормалізації. У майбутньому буде спроектовано нереляційну базу даних для телемедичної системи, що дасть змогу підвищити швидкість опрацювання даних.

Ключові слова: реплікація; FTP-сервер; даталогічна модель; гістологічні та цитологічні зображення; master-slave.

Вступ. Інформаційні системи різного масштабу активно впроваджують у медичних та дослідницьких закладах. Із збільшенням кількості, різновидів медичних досліджень та відповідно збільшенням обсягу даних виникає потреба у їх структуризації.

Система автоматизованої мікроскопії (САМ) – це програмно-апаратна система, що дає змогу автоматизувати аналіз біомедичних (гістологічних та цитологічних) зображень в умовах клінічної практики. До складу САМ входить мікроскоп, цифрова камера, комп'ютер та спеціалізоване програмне забезпечення (Berezkyi et al., 2016).

Автоматизація аналізу гістологічних та цитологічних структур пришвидшує діагностику захворювання, дає змогу розширити межі наукових пошуків у медицині. Автоматичне вимірювання параметрів гістологічних мікрооб'єктів дає змогу уточнити лікування і управління терапевтичними процесами. Більшість САМ складаються з апаратної (мікроскоп, відеокамера) та програмної частин. Основне завдання програмної частини полягає в опрацюванні вхідного зображення та визначенні мікрооб'єктів та ознак для подальшої постановки діаг-

нозу лікарем-фахівцем.

База даних (БД) є складовою частиною більшості інформаційних систем незалежно від галузі їх застосування. У медицині системи на основі БД використовують здебільшого для зберігання, обліку та подальшого оброблення інформації про пацієнтів. На основі статистичних даних про пацієнтів та їх діагнози застосовують методи для прогнозування захворювань.

Під час проектування структури БД для системи, що працює з великими масивами даних, та в якій одночасно може брати участь різна кількість користувачів, важливим постає завдання нормалізації самої БД.

Отже, *актуальність цієї роботи* полягає у проектуванні структури бази даних для забезпечення механізму спільного доступу до системи автоматизованої мікроскопії різних типів користувачів (лікарів).

Аналіз останніх досліджень та публікацій. БД є ключовим елементом більшості інформаційних систем, зокрема і спеціалізованих. Особливу увагу науковці приділяють особливостям проектування та роботі з БД у медицині. Автори M. Goldacre та L. Kurina у своїй статті (Goldacre et al., 2000) використовували статис-

Інформація про авторів:

Березький Олег Миколайович, д-р техн. наук, професор. Email: ob@tneu.edu.ua

Піцун Олег Йосипович, аспірант. Email: o.pitsun@tneu.edu.ua

Вербовий Сергій Олегович, аспірант. Email: vso@tneu.edu.ua

Дацко Тамара Вікторівна, канд. мед. наук, доцент. Email: datskotv@gmail.com

Цитування за ДСТУ: Березький О. М., Піцун О. Й., Вербовий С. О., Дацко Т. В. Розроблення реляційної бази даних інтелектуальної системи автоматизованої мікроскопії. Науковий вісник НЛТУ України. 2017. Вип. 27(5). С. 125–129.

Citation APA: Berezkyi, O. M., Pitsun, O. Yo., Verbovyi, S. O., & Datsko, T. V. (2017). Relational database of intelligent automated microscopy system. Scientific Bulletin of UNFU, 27(5), 125–129. <https://doi.org/10.15421/40270525>

тичні дані з медичних БД Oxford Record Linkage Study (ORLS) для пошуку закономірностей між захворюваннями пацієнтів та прогнозуванням захворювань. Miguel Delgado та Daniel SaÂnchez у роботі (Delgado et al., 2001) вводять свій підхід для побудови бази правил серед кількісних значень у реляційних БД. Основною метою цієї роботи є ефективне використання даних про пацієнта. Метою досліджень Brian S. Alper є аналіз електронних БД для забезпечення можливості відповіді на запитання сімейних лікарів (Alper et al., 2001), тобто створення електронного довідника для лікарів на основі наявних медичних БД. Із збільшенням обсягів даних постає потреба у дослідженні та впровадженні розподілених БД в інформаційних системах. Наприклад, Wiragama Wedashwara та Shingo Mabu у своїй роботі (Wedashwara et al., 2006) пропонують алгоритм підтримки прийняття рішень для проведення кластеризації в БД. Останнім часом дедалі більшої популярності набувають системи з можливістю інтелектуального аналізу даних, класифікацією чи кластеризацією даних. Тому пошук нових методів для зменшення надмірності інформації є актуальним завданням. Метод для обчислення нечітких функціональних залежностей між атрибутами в нечітких реляційних системах управління базами даних пропонують Miljan Vucetica та Miroslav Hudcib у роботі (Vucetic et al., 2013). Цей метод базується на використанні нечітких правил.

Структурна схема інтелектуальної системи автоматизованої мікроскопії. Для забезпечення доступу різних користувачів до системи автоматизованої мікроскопії передбачено наявність віддаленої бази даних. Хост, паролі та порт для підключення до БД користувачі отримують під час реєстрації у системі. Структурну схему розробленої інтелектуальної системи автоматизованої мікроскопії наведено на рис. 1.

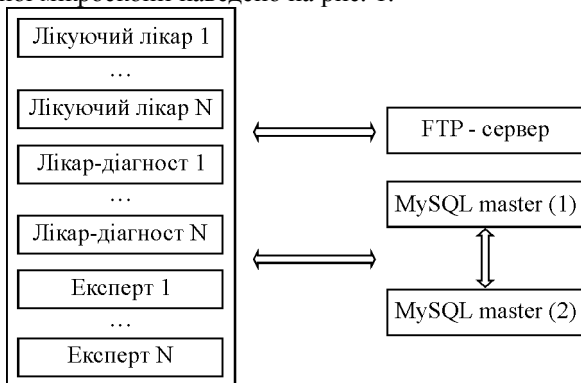


Рис. 1. Структурна схема інтелектуальної САМ

Одночасно з розробленою системою можуть працювати кілька користувачів. Усі вони під'єднуються до віддаленої бази даних та FTP-сервера. На FTP-сервері фізично зберігаються зображення, що були отримані внаслідок проведених досліджень. Система автоматично завантажує та скачує зображення на локальний комп'ютер користувача для подальшого перегляду та опрацювання (Berezsky et al., 2017).

Важливим елементом роботи будь-якого програмного додатку, а особливо медичного призначення, є висока відмовостійкість. Використання віддаленої бази даних, з одного боку, полегшує процес комунікації між користувачами, підвищуючи швидкість оброблення, а з іншого – поява будь-яких технічних проблем із сервером БД чи відсутність з'єднання з сервером призводить до непридатності системи (Thomson et al., 2012).

Реплікація – одна з технік масштабування баз даних (Wiesmann et al., 2000). Суть цього підходу полягає у тому, що дані постійно копіюються з одного сервера на інший. Структуру master-slave реплікації наведено на рис. 2.



Рис. 2. Master-slave реплікація

У Master-slave передбачено окремі сервери для зчитування даних та окремих серверів для операцій додавання, редагування, видалення записів. Master – Master реплікація дає змогу розподілити навантаження на БД між серверами. На відміну від Master-slave реплікації усі сервери в системі призначені для зчитування та додавання, редагування, видалення інформації (Pacitti et al., 2003; Chang et al., 2008). Структуру Master – Master реплікації наведено на рис. 3.



Рис. 3. Master – Master реплікація

Даталогічну модель розробленої бази даних наведено на рис. 4.

Модулі інтелектуальної сам. Основними користувачами САМ є лікарі. У таблиці "Doctors" (див. рис. 4) знаходяться поля для зберігання інформації про користувачів (лікарів) системи. Поле "role_id" відповідає за присвоєну роль користувача в системі, наприклад, лікуючий лікар, експерт, адміністратор. Логін та пароль створюються тільки за участю адміністратора. Функції реєстрації без відома адміністратора не передбачено.

Лікарі та пацієнти мають спільний набір полів, таких як "Прізвище", "Ім'я", "По батькові", тому для оптимізації даних ці поля розташовані в окремій таблиці "Users", як показано на рис. 4. У таблицях "Doctors" та "Patients" зберігається тільки ідентифікатор запису в таблиці "Users".

Поле "doctor_id" в таблиці "Patients" призначене для зберігання ідентифікатора лікаря, що заніс його в систему та є відповідно його лікуючим лікарем.

Зовнішній вигляд головного вікна розробленої САМ наведено на рис. 5. На боковій панелі реалізовано механізм налаштування доступу до віддаленої бази даних MySQL та віддаленого FTP-сервера для зберігання зображень для кожного дослідження.

Приклад вікна для перегляду пацієнтів наведено на рис. 6 (а). На рис. 6 (б) наведено приклад вікна для додавання нового пацієнта. У блоці "Перегляд пацієнта" передбачені такі функції: перегляд результатів досліджень, сортування досліджень за датою, постановка діагнозу, перегляд персональної інформації.

Модуль опрацювання та аналізу зображень є ключовим у розробленій САМ. Результати кожного дослідження зберігаються в БД, а гістологічні чи цитологічні зображення завантажуються на віддалений FTP-сервер. Такий підхід передбачений для зручного обміну зображеннями між користувачами системи. Наприклад, ек-

сперт, що перебуває за межами закладу, після авторизації та аутентифікації може переглядати зображення для консультації з лікарями. Ще однією перевагою розробленої САМ є простота інтерфейсу, а саме лікарям не потрібно знати всі аспекти роботи FTP-сервера. Зображення автоматично завантажуються на комп'ютер, з яким працює лікар після вибору пацієнта та дослід. Для забезпечення конфіденційності інформації директорії із зображеннями шифруються. Також лікарям не обов'язково знати про ПІБ пацієнтів, окрім лікуючого

лікаря. Приклад графічного інтерфейсу вікна для опрацювання зображень наведено на рис. 7.

Базовими критеріями оцінки біомедичних зображень є кількісні та якісні оцінки. Таблиця БД "cell" (див. рис. 4) вміщує поля для зберігання характеристик зображення. Поле "image_id" зберігає ідентифікатор кожного досліджуваного зображення. Графічний інтерфейс для опису кількісних та якісних характеристик наведено на рис. 8.

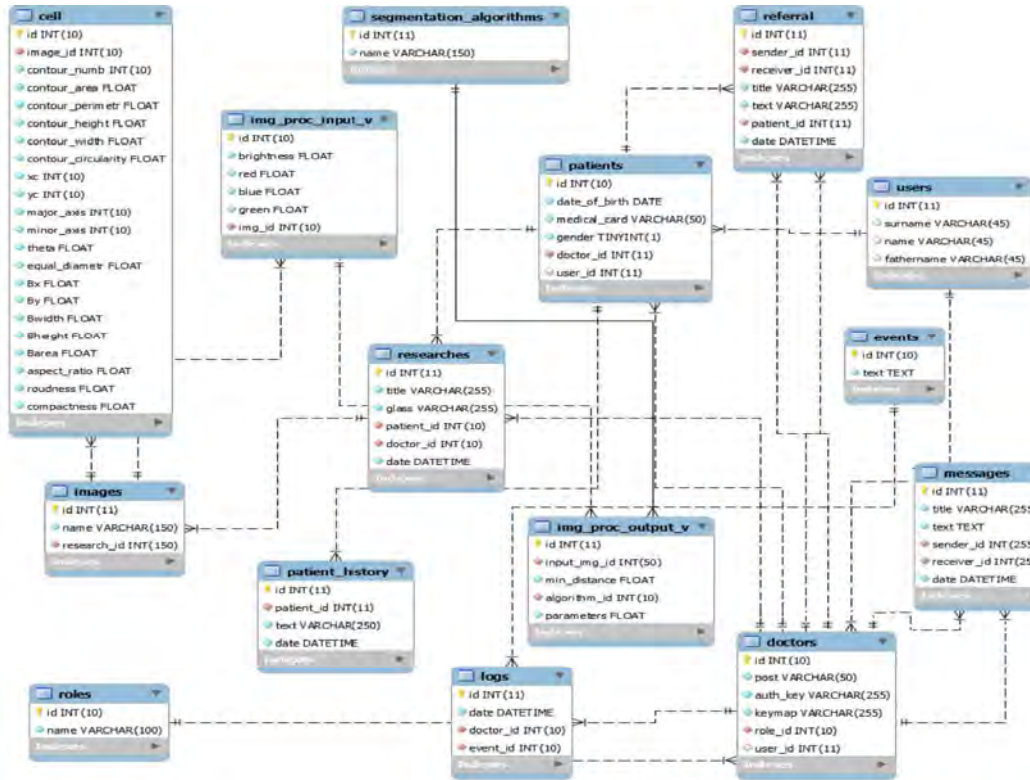


Рис. 4. Даталогічна модель бази даних

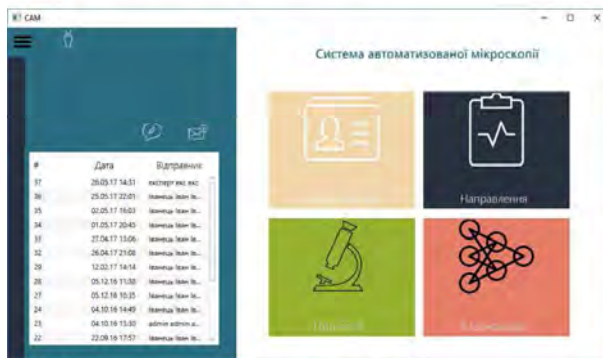


Рис. 5. Головне вікно розробленої САМ

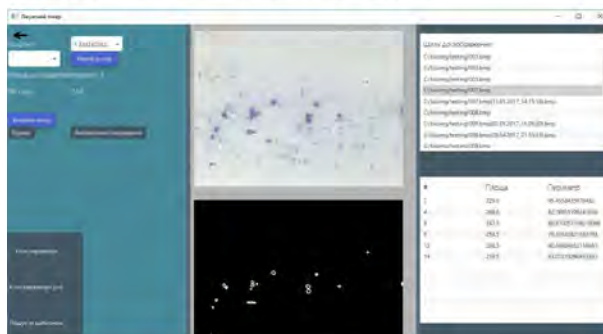


Рис. 7. Модуль опрацювання зображень

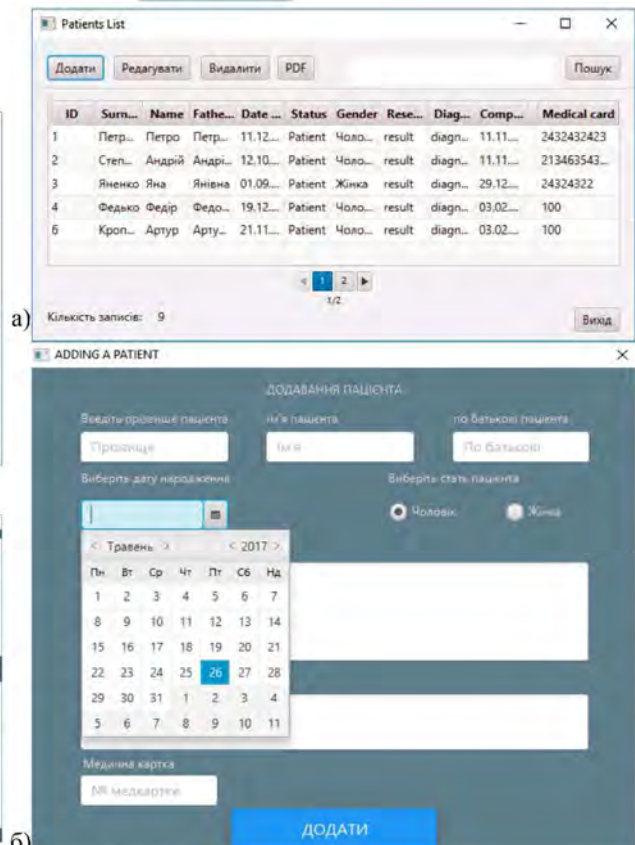
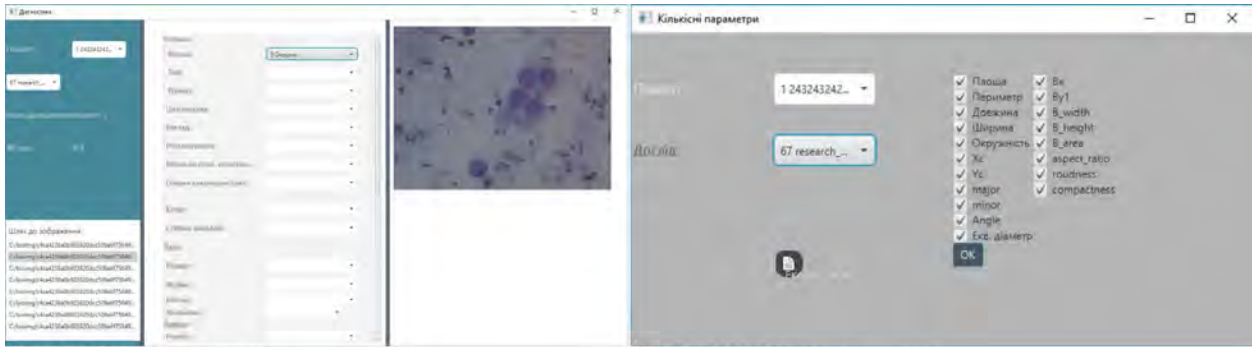


Рис. 6. Модуль "Пацієнти"



а) якісні характеристики
 б) кількісні характеристики

Рис. 8. Опис кількісних та якісних характеристик зображення

```

mysql> select * from bioimage.researches;
+----+-----+-----+-----+-----+
| id | name   | num_glass | patient_id | user_id |
+----+-----+-----+-----+-----+
| 60 | research1 | 325690 | 2 | 5 |
+----+-----+-----+-----+-----+
1 row in set (0.00 sec)

mysql> insert into bioimage.researches (name, num_glass, patient_id, user_id) value ('research2', 325690, 2, 5);
Query OK, 1 row affected (0.06 sec)

mysql> insert into bioimage.researches (name, num_glass, patient_id, user_id) value ('research3', 325690, 2, 5);
Query OK, 1 row affected (0.02 sec)

mysql> insert into bioimage.researches (name, num_glass, patient_id, user_id) value ('research4', 325690, 2, 5);
Query OK, 1 row affected (0.01 sec)

mysql> insert into bioimage.researches (name, num_glass, patient_id, user_id) value ('research5', 325690, 2, 5);
Query OK, 1 row affected (0.07 sec)

mysql> select * from bioimage.researches;
+----+-----+-----+-----+-----+
| id | name   | num_glass | patient_id | user_id |
+----+-----+-----+-----+-----+
| 60 | research1 | 325690 | 2 | 5 |
| 69 | research2 | 325690 | 2 | 5 |
| 70 | research3 | 325690 | 2 | 5 |
| 71 | research4 | 325690 | 2 | 5 |
| 72 | research5 | 325690 | 2 | 5 |
+----+-----+-----+-----+-----+
5 rows in set (0.01 sec)

mysql>
    
```

а)

б)

Рис. 9. master- slave реплікація на прикладі таблиці "researches"

Перевагою розробленої САМ є наявність механізму обміну повідомленнями між лікарями не відходячи від робочого місця та не використовуючи сторонні програмні засоби. Модуль "Направлення" дає змогу спростити роботу лікарів. Усі записи зберігаються в БД. Головними полями у вікні створення нового направлення є такі: Кому (вибір лікаря, до якого здійснюється направлення), Номер мед. Картки (номер медичної картки пацієнта), Текст (текст направлення, уточнення тощо). Структуру таблиці БД "referral" наведено у табл. 1.

Табл. 1. Структура таблиці "referral"

Поле	Тип даних	Пояснення
id	int (10)	Ідентифікатор
sender_id	int (100)	Ідентифікатор відправника
receiver_id	int (100)	Ідентифікатор отримувача
date	Datetime	Дата створення повідомлення
title	varchar (100)	Заголовок
text	varchar (100)	Текст повідомлення
patient_id	int (100)	Ідентифікатор пацієнта

Додавання запису в таблицю "researches". У даній таблиці міститься список проведених досліджень. Приклад SQL запити для додавання нового запису про дослідження:

```

INSERT INTO researches (name, num_glass, patient_id, user_id) VALUES ('research6', '241764', 5, 9);
    
```

Вміст таблиці наведено на рис. 9. Рис. 9 (а) та 9 (б) відображають вміст таблиці на серверах master та slave відповідно. Внаслідок запити тільки до master сервера отримаємо оновлену таблицю з новим записом на обох серверах.

Порівняльний аналіз. До відомих на ринку САМ відносять такі: MEKOC-Ц2, TissueFAXS, AnalySIS Five, BioVision, ВидеоТесТ Морфо 5.2, BioImageXD, Ariol, ImageJ, analySIS FIVE, Motic Images Advanced 3.2, ДинаМорф, Motic ВидеоТесТ Морфо 5.2, Cell D. Порівняльний аналіз САМ за критерієм роботи з БД наведено у табл. 2.

Табл. 2. Порівняльний аналіз САМ

Критерій	Imagej	Bi-oVi-sion	Диа-Морф	Axio-Vis-ion	Ami-ra	ME-KOC-ЦГ	AMS-Di-agnos-is	Розроб-лена САМ
Наявність БД	-	-	+	+	-	-	+	+
Master-slave реплікація	-	-	-	-	-	-	-	+
Master – master реплікація	-	-	-	-	-	-	-	+

Висновки. Розроблено структуру БД для оброблення та аналізу біомедичних (цитологічних, гістологічних) зображень. Розроблена САМ передбачає можливість одночасної роботи великої кількості лікарів та персоналу, тому спроектована структура БД дає змогу налагодити, спростити та пришвидшити взаємодію між користувачами. Одночасна робота з одним SQL-сервером кількох користувачів може значно уповільнити роботу системи. Тож використання master-slave реплікації дає змогу розподілити навантаження між кількома серверами.

Порівняно з аналогами розроблена система дає змогу автоматизувати та спростити роботу лікарів.

Перелік використаних джерел

- Alper, S. B., Stevermer, J. J., White, D. S., & Ewigman, B. G. (2001). Answering family physicians clinical questions using electronic medical databases. *Journal of Family Practice*, 50(11), 960–965.
- Berezkyi, O. M., Verbovyi, S. O., & Pitsun, O. Y. (2016). Systemy avtomatyzovanoi mikroskopii: stan ta perspektyvy rozvytku. *Visnyk Khmelnytskoho natsionalnoho universytetu*, 2(235), 61–68. [in Ukrainian].
- Berezsky, O., Dubchak, L., & Pitsun, O. (2017). Access distribution in automated microscopy system. In *The Experience of Designing and Application of Cad Systems in Microelectronics, CADSM2017*. Poliana-Svaliava, 360 p.
- Chang, F., Dean, J., Ghemawat, S., & Hsieh, W. C. (2008). Bigtable: A Distributed Storage System for Structured Data. *ACM Transactions on Computer Systems (TOCS)*, 26(2), 1–26. <https://doi.org/10.1145/1365815.1365816>
- Delgado, M., SâNchez, D., MartiN-Bautista, M. J., & Vila, M. A. (2001). Mining association rules with improved semantics in medical databases. *Artificial Intelligence in Medicine*, 21(1), 241–245. [https://doi.org/10.1016/S0933-3657\(00\)00092-0](https://doi.org/10.1016/S0933-3657(00)00092-0)
- Goldacre, M., Kurina, L., Yeates, D., Seagroatt, V., & Gill, L. (2000). Use of large medical databases to study associations between diseases. *QJM*, 93(10), 669–675. <https://doi.org/10.1093/qjmed/93.10.669>
- Pacitti, E., Özsu, M. T., & Coulon, C. (2003) Preventive Multi-master Replication in a Cluster of Autonomous Databases*. In: *Kosch H., Böszörményi L., Hellwagner H. (Eds) Euro-Par 2003 Parallel Processing*, (Vol. 2790). Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-45209-6_48
- Thomson, A., Diamond, T., Weng, S.-C., & Ren, K. (2012). Calvin: Fast Distributed Transactions for Partitioned Database Systems. *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*. Scottsdale, Arizona, USA, May 20–24. <https://doi.org/10.1145/2213836.2213838>
- Vucetic, M., Hudec, M., & Vujošević, M. (2013). A new method for computing fuzzy functional dependencies in relational database systems. *Expert Systems with Applications*, 40(7), 2738–2745. <https://doi.org/10.1016/j.eswa.2012.11.019>
- Wedashwara, W., Mabu, S., Obayashi, M., & Kuremoto, T. (2006). Combination of genetic network programming and knapsack problem to support record clustering on distributed databases. *Expert Systems with Applications*, 46, 15–23. <https://doi.org/10.1016/j.eswa.2015.10.006>
- Wiesmann, M., Pedone, F., Schiper, A., Kemme, B., & Alonso, G. (2000). Understanding replication in databases and distributed systems. *Proceedings of the The 20th International Conference on Distributed Computing Systems (ICDCS 2000)* (pp. 23–31), 464 p.

О. Н. Березький¹, О. Й. Пицун¹, С. О. Вербовий¹, Т. В. Дацко²

¹ Тернопольский национальный экономический университет, г. Тернополь Украина

² Тернопольский государственный медицинский университет им. И. Я. Горбачевского, г. Тернополь, Украина

РАЗРАБОТКА РЕЛЯЦИОННОЙ БАЗЫ ДАННЫХ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ АВТОМАТИЗИРОВАННОЙ МИКРОСКОПИИ

Спроектирована даталогическая модель базы данных интеллектуальной системы автоматизированной микроскопии для анализа гистологических и цитологических изображений и проведен обзор современных технологий проектирования распределенных баз данных. Разработанная структура базы данных предназначена для хранения и обмена информацией между пользователями системы и состоит из двух базовых составляющих: mysql-сервера базы данных и FTP-сервера для хранения изображений, полученных от микроскопа. Учитывая, что согласно законодательства нельзя масштабировать оригинальное изображение, на FTP-сервер осуществляется значительная нагрузка. Однако такой механизм дает возможность значительно упростить работу врачей, поэтому нет необходимости знать всех технических деталей работы серверов. Большинство существующих систем для обработки любого типа изображений обладают только средствами для анализа и не имеют в своем составе механизмов для хранения и обмена информацией. Предложенная структура базы данных дает возможность реализовать механизм обмена информацией между несколькими пользователями. Применение "master-slave" репликации дает возможность распределить нагрузку между существующими серверами и осуществить резервное копирование данных, что является незаменимым в медицинских системах. Спроектированная даталогическая модель обеспечивает минимальное дублирование информации и дает возможность расширять систему.

Ключевые слова: репликация; FTP-сервер; даталогическая модель; гистологические и цитологические изображения; master-slave репликация.

O. M. Berezkyi¹, O. Yo. Pitsun¹, S. O. Verbovyi¹, T. V. Datsko²

¹ Ternopil National Economic University, Ternopil, Ukraine

² I. Ya. Horbachevsky Ternopil State Medical University, Ternopil, Ukraine

RELATIONAL DATABASE OF INTELLIGENT AUTOMATED MICROSCOPY SYSTEM

Currently in medicine researchers pay much attention to designing databases for information systems that facilitate the work of doctors. Basically, the structure of a relational database allows conveniently generating reports and statistics on patients and their diagnoses. Most existing automated microscopy systems for image analysis do not have some kind of database or have limited functionality. In this paper, based on an analytical approach to the review of existing solutions for the design databases of information systems and based on the theory of database design created a model of a relational database for automated microscopy system that allowed developing adaptive functionality for different types of users. Developed intelligent automated microscopy system includes graphical interfaces and functionality for these types of users: doctor, doctor – diagnostician, expert, and administrator. A feature of this system is a mechanism of information exchange between doctors. For example, the doctor conducts image processing, and the doctor – diagnostician carries out a detailed analysis of the same set of images. In case of doubtful situations, doctors can contact the expert and ask his opinion. The system is designed so that doctors do not need to know all the technical nuances associated with the operation of the database and FTP – server to store images. Stored data securing is an important concern, especially in medicine. This system is suggested to use the mechanism of data replication type "mater-slave". This approach allows distributing the load between servers and backing up data. That is, when damaged one of the servers, the system automatically switches to the other. Given the fast growing popularity of non-relational database and increasing amounts of information about patients and their research raises the urgent task of developing software modules using non-relational databases (mongoDB) to speed read / write data.

Keywords: replication; FTP-server; datalogic model; histological and cytological images, master-slave.