

# Studentská Vědecká Konference 2010

## EXPERIMENTÁLNÍ LÉKAŘSKÝ INFORMAČNÍ SYSTÉM - ZPRACOVÁNÍ HETEROGENNÍCH NESTRUKTUROVANÝCH DAT

Miroslav KRÁL<sup>1</sup>

### 1 ÚVOD

Datové standardy pro přenos a uchovávání medicínských informací představují velmi zajímavou oblast z hlediska archivace a získávání důležitých informací, které mohou zachránit život nejednoho pacienta.

Tuto situaci si uvědomili i lékaři z Fakultní nemocnice Plzeň (dále FNPL), od kterých vzešel požadavek na vytvoření experimentálního informačního systému. Ten by lékařům umožňoval archivovat vybraná data pacientů a z těchto dat získávat výsledky, které by jim byly zobrazeny v přehledné podobě.

V rámci mé diplomové práce jsem se zabýval především navržením architektury systému, vytvořením databázového modelu pro uchovávání nestrukturovaných medicínských dat a programů pro import těchto dat do vytvořené databáze.

### 2 ARCHITEKTURA SYSTÉMU

Zjednodušený pohled na architekturu systému poskytuje obrázek 1, který obsahuje všechny stěžejní části systému. Nedílnou součástí získávání dat z FNPL je anonymizace těchto dat, která probíhá na straně nemocnice a bez ní by k samotnému přenosu dat na univerzitu nemohlo dojít. Zde se jedná především o hrozbu v podobě porušení zákona o ochraně osobních údajů.

Jednou z hlavních myšlenek systému je zpracovávat data, uložená v databázi, externími programy a získat tak potřebné výsledky. Každý výsledek musí být označen takzvaným *uživatelským příznakem*, který je definován lékařem a na základě těchto příznaků bude možno mezi uloženými výsledky vyhledávat. Takovýmto uživatelským příznakem může být například váha pacienta.

### 3 DATABÁZOVÝ MODEL A PROGRAMY PRO IMPORT DAT

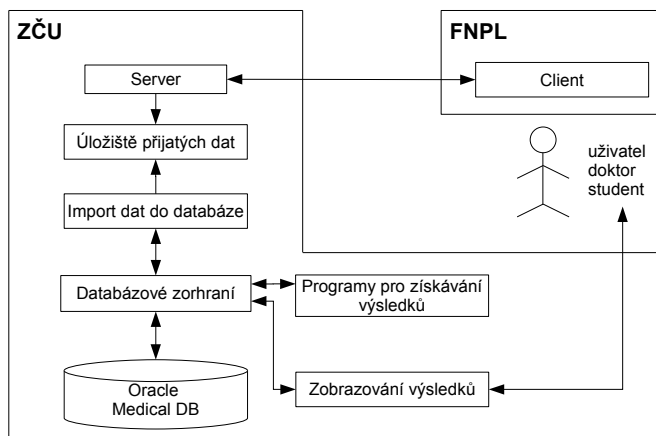
Samotné realizaci předcházela analytická fáze, ve které bylo potřeba prozkoumat možnosti konkrétních datových standardů pro přenos a archivaci medicínských dat. Těmito standardy jsou DASTA a DICOM.

Prvně jmenovaný standard slouží k přenosu textových informací v podobě XML souborů a je výsledkem dlouhodobého vývoje Ministerstva zdravotnictví České republiky a spolupracujících organizací. Druhý jmenovaný je mezinárodní standard pro přenos obrazových informací a k nim přidružených metadat. Doplnkově budou do databáze také ukládána data z mezinárodního registru SITS<sup>2</sup>. Ten shromažďuje informace o pacientech, které postihla mozková mrtvice.

Již od počátku bylo cílem navrhnout databázový model tak, aby odděloval dvě hlavní oblasti a to oblast pro archivaci medicínských dat a oblast pro uchovávání výsledků získa-

<sup>1</sup>Bc. Miroslav Král, student navazujícího studijního programu Aplikované vědy a informatika, obor Inženýrská informatika, specializace Softwarové inženýrství, e-mail: miika@students.zcu.cz

<sup>2</sup>Safe Implementation of Thrombolysis in Stroke



Obr. 1: Architektura systému

ných z těchto dat. Model musí být jednoduše rozšiřitelný z hlediska budoucího přidávání nových datových standardů nebo struktur uchovávajících medicínská data.

Programy pro import dat do databáze musí zajisti získání konkrétních informací z přijatých souborů a jejich uložení spolu s celým obsahem souboru do databáze. Další nedílnou činností je správné vzájemné provázání přijatých dat. Zde se jedná zejména o přílohy DASTY v podobě DICOM a SITS souborů souvisejících s vyšetřením pacienta. V tomto případě nešlo předpovědět pořadí příchodu jednotlivých souborů, a proto musely být ošetřeny všechny možné případy.

#### 4 ZÁVĚR

Momentálně je možné importovat přijatá anonymizovaná data do části databáze určené k jejich archivaci. Ověření funkcionality a správného návrhu druhé části databáze bude možné až ve chvíli, kdy budou existovat externí programy, které budou získávat výsledky. Tato část skýtá do budoucna největší potenciál v oblasti možných zlepšení a vytváření jednotlivých programů a stejně tak optimalizace jednotlivých komponent systému při nasazení do reálného provozu.

Největším problémem se stále jeví získání odpovídajícího množství dat. Tento problém je způsoben především dříve zmíněnými legislativními nároky.

**Poděkování:** Příspěvek byl podpořen grantovým projektem *Intelligent system for early diagnosis and treatment of cerebrovascular diseases*. Rád bych poděkoval Doc. Dr. Ing. Janě Klečkové, vedoucí mé práce, a Ing. Petru Včelákovi, se kterým jsem konzultoval navržená řešení.