



Slovenská technická univerzita
FAKULTA INFORMATIKY A INFORMAČNÝCH TECHNOLOGIÍ
Ilkovičova 3, 842 16 Bratislava 4



Znalostný manažment na báze technológie .NET

**Posudok prototypu tímu č. 7 – Lucky Number 7
(3. časť)**

*Tím číslo 5: Lubomír Chamraz , Ivan Kišac , Ján Krausko , Michal Kurták , Marián Šimko ,
Michal Šimún*

Vedúci tímu: Mgr. György Frivolt

Študijný odbor / program: Softvérové inžinierstvo / Softvérové inžinierstvo

Ročník, typ štúdia: 1, inžinierske štúdium

Predkladaný dokument sa zoberá posudzovaním práce a výsledkov počas druhého semestra tímu č. 7 pri riešení softvérového projektu “Znalostný manažment na báze technológie .NET” v rámci predmetu Tvorba softvérového systému v tíme v akademickom roku 2006/2007.

V prvej časti posudku je hodnotená časť návrh a implementácie projektu. V druhej časti je opísaná správa o testovaní výsledného produktu.

Časť I – Návrh riešenia a implementácia

V časti Návrh riešenia a implementácia sa autori zaoberajú návrhom a implementáciou výsledného produktu, ktorý je založený na vrstvovej architektúre. Výber tohto typu architektúry hodnotíme pozitívne, nakoľko každá vrstva má presne definovanú vlastnú funkcionálnu. Z dokumentu však nie je veľmi jasné, z koľkých vrstiev táto architektúra pozostáva. Na lepšie porozumenie tohto návrhu by bol vhodný jednoduchý diagram, ktorý by znázorňoval jednotlivé úrovne a ich poradie.

Databázová vrstva

V tejto časti sa autori podrobnejšie zamerali na návrh a implementáciu databázovej vrstvy. Na začiatku kapitoly sa nachádza stručný, ale postačujúci popis funkcionality tejto vrstvy. Sú tu spomenuté aj nefunkcionálne požiadavky na vytváraný systém (použitie externej SQL databázy).

Databázová vrstva slúži na ukladanie a získavanie dokumentov z externej databázy. Tvorí rozhranie medzi vyššou vrstvou (Data management layer) a touto databázou. Preštudovaním dokumentácie sme zistili, že autori pravdepodobne použili rôznu terminológiu, nakoľko opis vrstvy Data management layer sa v dokumentácii nenachádza. Predpokladáme, že táto vrstva je totožná s vrstvou Data provider layer, ktorá je popísaná v kapitole Vrstva poskytovania dát. Rôzna terminológia môže vyvolať chybné predstavy o architektúre systému, a teda aj celej funkcionality.

Návrh databázovej vrstvy je rozdelený do dvoch častí. Prvá časť opisuje návrh dátovej štruktúry, ktorá definuje uchovávaný dokument v databáze. Všetky údajové položky sú popísané veľmi stručne a nie je dostatočne jasné, na čo niektoré položky slúžia (napr. hash). Druhá časť návrhu sa zameriava na opis funkcií, ktoré poskytuje databázová vrstva. Tieto funkcie sú vysvetlené podrobne, a teda vieme presne povedať, čo ktorá funkcia vykonáva. Autori popísali funkcionálnu aj na úrovni implementačného jazyka, kde sú definované hlavičky metód, vstupy a výstupy. Tento popis je v samostatnej časti, pričom sa tu nachádzajú duplicitné popisy funkcií, ktoré boli vysvetlené v predchádzajúcej časti (iná reprezentácia opisu funkcií). Duplicitné popisy funkcií zaťažujú čitateľa dokumentácie.

V poslednej časti autori stručne popisujú implementáciu databázovej vrstvy a prikladajú diagram tried, ktorý názorne zobrazuje štruktúru tried, čo taktiež uľahčuje pochopenie celého návrhu a implementácie.

Vrstva poskytovania dát

Ďalšou vrstvou, ktorú autori navrhli, je vrstva poskytovania dát (Data provider layer). Rovnako ako v kapitole Databázová vrstva, tak aj tu sa autori venovali návrhu a implementácii tejto vrstvy, ako aj popisu vrstvy a jednotlivých požiadaviek.

Hlavnou úlohou vrstvy poskytovania dát je spracovanie údajov zo zadaných dokumentov, ich indexácia a vyhľadávanie. Autori túto vrstvu rozdelili na dva hlavné moduly (parsovací a indexovací modul), ktorých funkciu popísali stručne, ale zato výstižne. Pre lepšie pochopenie štruktúry týchto modulov v rámci vrstvy, ale aj vrstvy v rámci celej architektúry systému, slúži priložený diagram komponentov.

V ďalších podkapitolách je už podrobne popísaný návrh a implementácia obidvoch modulov. Autori sa zamerali už iba na opis najdôležitejších častí, čo je pozitívne, a uviedli aj hlavičky jednotlivých funkcií, vstupy a výstupy (rovnako ako v kapitole Databázová vrstva). V tejto časti (Modul pre indexáciu a vyhľadávanie) sa nachádza aj diagram aktivít, ktorý znázorňuje postup pridania dokumentu do systému. Tento diagram dobre poslúži čitateľovi, ale jeho umiestnenie by bolo asi vhodnejšie v samostatnej kapitole, nie v kapitolách vrstiev, keďže zaznamenáva funkcionálnosť všetkých vrstiev.

Komponent - expertný systém

V tejto kapitole autori predstavujú použitý expertný systém Clips, návrh a implementáciu jeho integrácie do systému, ako aj vytvorené metódy vyhľadávania dokumentov.

Najskôr predstavujú špecifikáciu modulu (ako funkcionálne, tak aj nefunkcionálne požiadavky). Z tejto špecifikácie sa dozvedáme vlastnú úlohu komponentu v systéme. Trochu nešťastným spôsobom bola použitá formulácia nefunkcionálnych požiadaviek. Formulácia "Je zabezpečené uchovávanie bázy znalostí v textovom súbore. Je zabezpečené..." pôsobí máľúco, pretože nevieme, kto a ako danú vec zabezpečil. V tomto prípade zrejme išlo o spôsob uloženia bázy znalostí, ktorý vzhľadom na použitie v štruktúrovanom texte mohol byť vyjadrený jednoduchšie a bez vyvolania zbytočných otázok.

V oblasti samotného návrhu treba oceniť, že autori pri jeho tvorbe pamätali na modularitu a modifikovateľnosť svojho riešenia a návrh rozdelili na všeobecnú časť nezávislú od použitého expertného systému a časť implementujúcu techniku prehľadávania a dopĺňania znalostí v Clipse. Každá časť je popísaná samostatne na základe popisu svojich tried. Na popis architektúry tried použili diagram tried, ktorý výrazne sprehľadnil a uľahčil pochopenie architektúry daných častí.

Súčasťou návrhu je aj popis jednotlivých techník používaných na prehľadávanie dokumentov – technika prehľadávania na základe kľúčových slov a technika prehľadávania na základe už existujúcich kapitol. Každá z týchto techník pracuje s rozličnými pojmami a dátovými štruktúrami. Význam týchto pojmov, aj použitie štruktúr sú dostatočne v každej kapitole popísané, čo umožňuje ľahšie pochopenie predstavovanej techniky. Pochopenie uľahčuje použitie jednoduchých vzorcov a diagramu aktivít predstavujúceho postup činností pri ohodnocovaní systémom Clips.

V závere kapitoly autori opisujú, ako pristupovali k otázke testovania systému. Na úrovni testovania jednotiek použili framework NUnit. Čiže efektívne využili dostupné technológie. Pri testovaní pokryli najmä základné funkcionálne aspekty komponentu.

Plugin do MS Word 2007

Táto kapitola predstavuje prístup k integrácii projektu do programu MS Word. Autori sa rozhodli pre vytvorenie plugin-u pre MS Word 2007. Žiaľ neuvádzajú, či je použiteľný aj v skorších verziách tohto kancelárskeho balíka.

V úvode predstavujú možnosti dostupné s použitím tohto plugin-u, ako aj prostredie, na ktorom bol vyvíjaný. Za týmto úvodom nasleduje ešte súhrn požiadaviek na prostredie pre vývoj plugin-u.

Plugin bol vyvinutý v prostredí .NET Framework. V rámci tohto prostredia vytvorili autori 4 projekty pri vývoji plugin-u. Tieto projekty ďalej popisujú. Každý z nich zohráva určitú úlohu pri vývoji a testovaní plugin-u – webové služby (napr. na testovanie), vytváranie inštalátora, nastavenie bezpečnosti a samotný projekt s plugin-om – GUIKrtko.

O projekte GUIKrtko autori informujú trochu viac ako o podporných projektoch, ale aj tak sa nedozvedáme informácie o návrhu a bližšom fungovaní samotného plugin-u.

V závere sú obsiahnuté informácie o tom, ako je možné celé toto riešenie spustiť.

Formálna stránka

Autori zvýšili zrozumiteľnosť dokumentu oproti dokumentácii prototypu. Vo väčšej miere začali používať aj mimojazykové prostriedky na znázorňovanie architektúry a funkcionality popisovaného systému. Zrozumiteľnosti do značnej miery napomáha používanie diagramov tried a diagramov aktivít. Naďalej sa však pomerne často vyskytujú gramatické chyby či nezrozumiteľné formulácie, ktoré vyrušujú čitateľa a sťažujú tak porozumenie textu. Vzhľadom na to, že v popisoch systému sa často vyskytuje viacero systémovo-špecifických termínov na jednom mieste, neporozumenie formulácie vety sa stáva pri štúdiu dokumentu kritickým.

V zmysle zachovania štandardov písania dokumentov a referencovania v dokumente použitých obrázkov a tabuliek autori nepoužili odkazovanie na tabuľky. Tabuľky sú však logicky umiestnené za textom, ktorý sa na ne odvoláva, takže nedochádza k väčším problémom s ich priradením alebo vyhľadáním. Explicitné použitie odkazov by bolo vhodné na uistenie čitateľa ohľadom referencovanej tabuľky, najmä v prípadoch, kedy formulácie vety nie je jednoznačná.

Časť II – Správa o testovaní produktu

Vytvorený produkt je realizovaný ako plugin do Microsoft Word 2007. Produkt je ovládaný prostredníctvom 3 záložiek – kľúčové slová, podobné kapitoly a konfigurácia. Konfigurácia určuje cestu k serveru, kde sa nachádza báza dát aplikácie. Testy boli počas používateľskej prezentácie vykonané nad serverom, kde bol umiestnený jeden dokument s viacerými kapitolami, ktorý poskytli k dispozícii autori aplikácie.

Aplikácia bola testovaná na úrovni komponentov pomocou nástroja NUnit. Cieľom vykonaných používateľských testov bolo overenie celkovej funkčnosti aplikácie – vyhľadávanie kapitol na základe kľúčových slov, náhľad na odporúčané kapitoly, vloženie kapitol do dokumentu, overenie zmeny relevancie kapitol a získanie podobných kapitol.

Návrh testov

Tab. 1 Návrh testov

Testované funkcie	Vyhľadávanie kapitol, náhľad, vloženie kapitol o dokumentu, overenie zmeny relevancie, získanie podobných kapitol
Opis postupu	Testovanie prebiehalo ručne so snahou pokryť viaceré typy vstupov. Kontrola s požadovanými výstupmi prebiehala vizuálne
Testovacie prípady	TC1, TC2, TC3

Testovacie prípady

V tabuľkách Tab. 2 až Tab. 4 sú uvedené špecifikácie testovacích prípadov. Medzi vstupné požiadavky všetkých prípadov patrí správne nakonfigurovaná aplikácia vložení cesty k serveru.

Tab. 2 Testovací prípad TC1

ID	TC1
Testovaná položka	Vyhľadanie kapitol, náhľad na získané kapitoly, vkladanie kapitol do otvoreného dokumentu
Vstupy	1. Zadanie kľúčových slov 2. Výber kapitoly
Výstupy	1. Umožnenie náhľadu na kapitolu 2. Vloženie kapitoly do dokumentu
Požiadavky na prostredie	Otvorená inštalácia aplikácie Word 2007 s nainštalovaným plugin-om aplikácie, vyhľadávaný dokument vložený do databázy

Tab. 3 Testovací prípad TC2

ID	TC2
Testovaná položka	Viacnásobné vyhľadávanie kapitol a vkladanie do dokumentu s cieľom zvyšovania relevancie vybranej kapitoly
Vstupy	1. Zadanie kľúčových slov 2. Viacnásobný výber kapitol
Výstupy	1. Viacnásobne vložená kapitola do dokumentu 2. Zmenená relevancia kapitol
Požiadavky na prostredie	Otvorená inštalácia aplikácie Word 2007 s nainštalovaným plugin-om aplikácie, vyhľadávaný dokument vložený do databázy, žiadne znalosti v aplikácií

Tab. 4 Testovací prípad TC3

ID	TC3
Testovaná položka	Podobnosť kapitol v dokumente
Vstupy	1. Zadanie kľúčových slov 2. Výber kapitoly do dokumentu 3. Výber podobných kapitol

Výstup	1. Zobrazená podobnosť kapitol
Požiadavky na prostredie	Otvorená inštalácia aplikácie Word 2007 s nainštalovaným plugin-om aplikácie, vyhľadávaný dokument vložený do databázy

Testovacie procedúry

Testovacia procedúra 1

Účel: Testovanie produktu pre testovací scenár TC1.

Kroky procedúry:

1. **Zahájenie:**
Otvorenie Doc Buildera
2. **Procedúra:**
 - 2.1. Zadanie kľúčového slova *software*
 - 2.2. Stlačenie tlačidla Hľadaj
 - 2.3. Vyžiadanie náhľadu vybranej kapitoly (*kapitola dp*) – presunom kurzoru myši nad názov kapitoly
 - 2.4. Kliknutie na vybranú kapitolu
3. **Meranie výsledkov:**
Výsledky sú merané vizuálne na základe očakávaného stavu v dokumente – vloženie textu kapitoly do dokumentu.
4. **Ukončenie:**
Ukončenie programu

Testovacia procedúra 2

Účel: Testovanie produktu pre testovacie scenáre TC2.

Kroky procedúry:

1. **Zahájenie:**
Otvorenie Doc Buildera
2. **Procedúra:**
 - 2.1. Zadanie kľúčového slova – *software*
 - 2.2. Získanie relevancia kapitoly po prvom hľadaní (24%)
 - 2.3. Výber kapitoly *dp* prehľadávaného dokumentu a vloženie do dokumentu
 - 2.4. Viacnásobné vyhľadanie a výber kapitoly *dp*
 - 2.5. Sledovanie zvýšenia relevancia kapitoly (29%)
3. **Meranie výsledkov:**
Výsledky sú merané vizuálne – aplikácia zobrazuje pri jednotlivých výsledkoch aj relevancie kapitol

4. **Ukončenie:**
Ukončenie programu

Testovacia procedúra 3

Účel: Testovanie produktu pre testovacie scenáre TC3.

Kroky procedúry:

1. **Zahájenie:**
Otvorenie Doc Buildera
2. **Procedúra:**
 - 2.1 Zadanie kľúčového slova – *software*
 - 2.2 Výber a vloženie kapitoly do dokumentu
 - 2.3 Požiadavka na zobrazenie podobných kapitol (záložka Podobné kapitoly)
3. **Meranie výsledkov:**
Výsledky sú merané vizuálne na základe zobrazenia podobnosti s ostatnými kapitolami.
4. **Ukončenie:**
Ukončenie programu.

Súhrnná správa o teste

Požadované výsledky	TC1 – získanie relevantných kapitol, umožnenie náhľadu, vloženie textu do dokumentu TC2 – Viacnásobným výberom kapitol zmena relevancie TC3 – Nájdienie podobných kapitol
Dosiahnuté výsledky	TC1 – Pri naplnenej databáze sú výsledky zhodné s očakávanými TC2 – Očakávaná zmena relevancie pre danú kapitolu, avšak podobné kapitoly svoju relevanciu nemenili TC3 – Pri požiadavke na získanie podobných kapitol nastala chyba v aplikácií
Ohodnotenie a návrhy na zlepšenie	Jediným zisteným nedostatkom oproti špecifikácií produktu je vyhľadanie podobných kapitol – TC3

Zhodnotenie

Vytvorený produkt je realizovaný ako plugin do MS Word 2007, čo zlepšuje použiteľnosť a dostupnosť aplikácie. Ďalším pozitívom je interakcia aplikácie s vytváraním dokumentu, kde používateľ môže priamo vložiť text vybranej kapitoly do textu dokumentu. Architektúra aplikácie je založená na klient/server architektúre s využívaním služieb, čo zvyšuje robustnosť a rozšíriteľnosť aplikácie. Autori tiež správne uvažujú nad rozšírením aplikácie na indexáciu a vkladanie iných typov objektov – obrázkov.

V práci sa však vyskytlo aj viacero nedostatkov, ktoré by bolo potrebné v ďalšej práci odstrániť. Vo výslednom produkte sme nenašli žiadne roly používateľov alebo personalizované profily, ktoré sú opisované v dokumente. Taktiež autori nezdôvodnili vybranú reprezentáciu dát a tiež neuviedli dôvod, prečo nepoužívali ontológie určené práve na reprezentáciu znalostí. V práci nám tiež chýba vytvorenie relácií medzi jednotlivými kľúčovými slovami, ktoré by sa mohli využiť pri odporúčaní kapitol. Počas testovania boli zistené niektoré nedostatky, kde najväčším nedostatkom bolo nefunkčnosť podobnosti kapitol. Testovanie podobnosti kapitol nebolo opísané ani autormi v dokumentácii projektu.