



SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
Fakulta informatiky a informačných technológií

Posudok dokumentácie k softvérovému systému

Znalostný manažment na báze technológie .NET

Tím č. 2 – 1337: Pol druhej a niečo

Bc. Marko Divéky
Bc. Marian Gablovský
Bc. Matúš Korman
Bc. Jakub Kožíšek
Bc. Katarína Sirotová

23. november 2007

E-mail: tp0708.02@gmail.com

ÚVOD

Tento dokument predstavuje posudok tímu číslo 2 na dokumentáciu tímu číslo 19, ktorého témou je „Znalostný manažment na báze technológie .NET.“

Posudzovaný dokument obsahuje analýzu problému, špecifikáciu požiadaviek a návrh riešenia pre znalostný softvérový systém, ktorého úlohou je uchovávanie, spracovávanie, kategorizovanie a flexibilné vyhľadávanie v dokumentoch rozličných formátov.

Posudok sa zaoberá formálnou i obsahovou stránkou posudzovanej dokumentácie, podľa častí ktorej jej aj členený.

FORMÁLNA STRÁNKA

Z formálneho hľadiska posudzovaný dokument neobsahuje gramatické chyby, no na viacerých miestach chýbajú čiarky, respektíve sú vo vetách zle umiestnené, napr. vo vete na str. 20: *„Výhodou využitia takýchto prepojení je fakt že sa nám môže podariť nájsť dokument ktorý síce priamo nesúvisí s nami vyhľadávaným výrazom, ale odkazujú sa naň dokumenty ktoré s našou požiadavkou priamo súvisia a teda je veľká pravdepodobnosť, že takto nájdený dokument bude relevantný.“*

V niektorých vetách je poprehadzovaný slovosled, prípadne sa v nich nachádzajú slová v zlom tvare. Ako príklad uvádzame časť vety zo str. 21: *„...priradená hodnota určujúca jeho dôležitosť, a tak ovplyvniť výsledky určenie relevancie dokumentu.“*

Syntaktické a štylistické chyby tohto druhu v značnej miere prispievajú k zníženiu celkovej formálnej stránky a čitateľnosti dokumentácie.

V posudzovanom dokumente sa taktiež viackrát vyskytuje pojem “užívateľ,” ktorého správnym ekvivalentom v kontexte používania softvérových systémov je pojem “používateľ.”

OBSAHOVÁ STRÁNKA

V tejto kapitole posudku sa nachádzajú postrehy a pripomienky týkajúce sa obsahovej stránky. Tie sú roztriedené podľa častí, ktorých sa dotýkajú. Názvy týchto častí sú identické s názvami častí posudzovaného dokumentu.

0 Úvod

Úvodná kapitola posudzovanej dokumentácie je číslovaná od nuly, čo pokladáme za veľmi neobvyklý a netradičný prístup.

0.2 Účel dokumentu. V podkapitole s názvom “Účel dokumentu” autori namiesto samotného účelu dokumentu opisujú skôr účel projektu a tímu.

1 Analýza

Analýza problémovej oblasti sa skladá z niekoľkých kapitol, ktorým chýba jasná logická nadväznosť. Pre čitateľa je preto obtiažnejšie sledovať myšlienku, ktorá chcela byť analýzou vyjadrená.

1.1 Úvod do manažmentu znalostí. Súčasťou podkapitoly “Úvod do manažmentu znalostí” je zoznam disciplín, z ktorých manažment znalostí čerpá. Položku “organizačná kultúra” by sme nezaradili medzi disciplíny, nakoľko sa jedná len o faktor, resp. súhrn pravidiel alebo smerníc, ktorý je platný v rámci organizácie.

1.2 Analýza znalostného systému vytvoreného tímom Lucky Number 7. V analýze existujúcej práce tímu Lucky Number 7 je, podľa nášho názoru, kladený veľký dôraz na úroveň implementácie analyzovaného systému. Tento fakt posúva danú časť analýzy na príliš nízku úroveň abstrakcie, a tým jej informačná hodnota v kontexte dokumentu klesá. Na úkor analýzy tried a funkcií odporúčame sústrediť sa na návrhové vzory a princípy, pomocou ktorých je možné opisované rozhrania charakterizovať.

Diagramy popisujúce architektúru systému vytvoreného tímom Lucky Number 7 nehovoria veľmi o jeho funkcionalite, ktorú by bolo vhodné popísať slovne za príslušnými diagramami. Na druhej strane, iba strohé vymenovanie metód v diagramoch tried nie je veľmi vravné.

Autori ďalej v opise návrhu dátovej štruktúry projektu tímu Lucky Number 7 uvádzajú položku *hash*, ktorý vraj jednoznačne identifikuje každý súbor. Toto tvrdenie považujeme za odvážne, pretože na základe princípu hašovania ako jednej zo základných kryptografických funkcií, sa nedá spoľahnúť na jednoznačnosť výstupu tohto procesu pre dva rôzne vstupy (v tomto prípade súbory) napriek tomu, že pravdepodobnosť vzniku kolízie je malá.

Z analýzy minuloročného projektu celkovo nie je veľmi jasné, aká bola výsledná funkcionalita tohto softvérového systému. Taktiež nie je jasné, ktoré z jeho častí sa rozhodol tím číslo 19 použiť, teda integrovať do navrhovaného softvérového systému.

1.3. Možnosti spracovania dokumentov. Očíslovanie a matematické naformátovanie jednotlivých vzorcov na výpočet váh kľúčových slov v rámci časti "Vektorový model" by zlepšilo jej čitateľnosť.

1.4 Vyhľadávanie informácií v dokumentoch, indexovanie. Množinu pomenovaní indexujúceho a vyhľadávajúceho systému Lucene.Net na str. 13 by bolo vhodnejšie spomenúť v časti úvodu „Použité skratky,“ nakoľko v časti dokumentácie venujúcej sa analýze ich výskyt pôsobí rušivo.

Popis tabuľky č. 2 sa nezhoduje s jej obsahom, pretože v tejto tabuľke je zobrazený prehľad možných požiadaviek pre vyhľadávanie v systéme Lucene.Net a nie porovnanie akýchsi modelov, ako je uvedené v popise tabuľky.

1.6 Modelovanie používateľov v znalostnom manažmente. V časti “Manuálne definovaná časť modelu používateľa” sú vymenované najdôležitejšie informácie o používateľovi už vzhľadom na navrhovaný systém. Tie by bolo vhodné zaradiť skôr do časti týkajúcej sa návrhu, než do časti zaoberajúcou sa analýzou.

Napriek spomenutej, prevažne negatívnej kritike, oceňujeme výkladový charakter textov tejto kapitoly dokumentácie, ktoré svojou podstatou získajú záujem aj u nezainteresovaného čitateľa. Obsahovo je analýza problematiky po informačnej stránke pomerne vyčerpávajúca.

2 Špecifikácia

Táto kapitola posudzovanej dokumentácie je prehľadne spracovaná a všetky zadané prípady použitia sú detailne popísané. Na základe zmienok o nasadení navrhovaného systému vo firemnom prostredí sme ale očakávali, že špecifikácia systému bude vo väčšej miere odrážať tento predpoklad.

2.1 Špecifikácia požadovaného riešenia. Diagram prípadov použitia navrhovaného systému naznačuje, že pridávať dokumenty môže len administrátor systému a používateľ má právo iba v dokumenty prezerať, resp. vyhľadávať. Nechávame autorom na zváženie, či nie je vhodné umožniť aj používateľom pridávať súbory do navrhovaného systému.

2.3 Prípady použitia. Tabuľkám opisujúcich prípady použitia chýbajú popisy, ktoré treba uviesť aj keď je implicitne jasné, čo sa v týchto tabuľkách nachádza.

Opis prípadu použitia „UC4 Prezeranie dokumentu“ naznačuje, že používatelia sa k prehliadaniu dokumentov dostanú iba cez vyhľadávanie, a teda prezeranie dokumentov v

rámci kategórií nie je možné. Nechávame na zváženie autorom v dokumente uviesť, či táto skutočnosť je zámerná alebo nie je.

V opise prípadu použitia „UC7 Priradenie kategórie dokumentu“ je spomenuté, že jeden dokument môže patriť iba do jednej kategórie. Podľa nás ale existujú aj také dokumenty, ktoré môžu patriť aj do viacerých kategórií.

2.4 Model údajov. V logickom modeli údajov navrhovaného systému sú identifikované len dve entity – jedna predstavuje *používateľa* a druhá *dokument*. Takýto model rozhodne nemôže stačiť pre navrhovanú funkcionálnosť systému, nakoľko už v tejto etape špecifikácie a návrhu je evidentné, že treba do modelu údajov zahrnúť aj triedy pre podporu napr. kategorizácie dokumentov, spolu s príslušnými väzbami na ostatné triedy.

2.5 Nefunkcionálne požiadavky. Nefunkcionálne požiadavky na systém sú formulované veľmi stručne. Unikli nám vlastnosti, na základe ktorých má byť navrhovaný systém lepším a vhodnejším než existujúce riešenia. Inými slovami povedané, z nefunkcionálnych požiadaviek nie je jasné, v čom je navrhovaný systém inovatívny, napr. z hľadiska nových prístupov k vyhľadávaniu v dokumentoch alebo z hľadiska použitých algoritmov a postupov.

Nechávame na zváženie autorov posúdiť, či nie je výhodnejšie a efektívnejšie postaviť navrhovaný systém ako webovú aplikáciu namiesto samostatnej „stand-alone“ aplikácie.

3 Návrh

Tretia kapitola posudzovanej dokumentácie obsahuje okrem architektúry navrhovaného systému a popisom každého modulu aj fyzický model údajov. Čiastkové procesy navrhovaného systému a ich dynamika sú v rámci modulov detailne popísané.

3.1 Architektúra systému. Architektúra systému znázorňuje „Vyhľadávací modul“ ako modul poskytujúci dodatočnú funkcionálnosť systému (teda modul, bez ktorého je navrhovaný systém schopný fungovať), no na str. 31 je vyslovene napísané, že tento modul „...*bude pevnou súčasťou celého systému.*“ Odhadujeme, že spomenutý modul skutočne tvorí základ celého systému, nakoľko je v dokumentácii vyhľadávanie súborov chápané ako jedna zo základných funkcií navrhovaného systému, a teda chyba je na strane diagramu s architektúrou.

Taktiež sa nám zdá korektné umiestniť v rámci architektúry spomínaný „Vyhľadávací modul“ nad moduly pre grafové algoritmy, kategórie dokumentov a sociálne siete, nakoľko práve modul zabezpečujúci vyhľadávanie využíva poznatky týchto troch modulov, nie naopak.

3.2 Fyzický model. Špecifikácia fyzického modelu údajov je skôr na úrovni logického modelu údajov. V diagrame chýbajú nielen údajové typy atribútov zobrazených tried, ale aj samotné atribúty tvoriace prepojenie medzi nimi. Ako už bolo spomenuté pri logickom modeli údajov v predchádzajúcej kapitole, tento údajový model, spolu s jeho fyzickou interpretáciou, rozhodne nemôže stačiť pre navrhovanú funkcionálnosť systému a je nutné do neho doplniť ďalšie potrebné triedy.

3.3 Vyhľadávací modul. Obr. 14 znázorňujúci prídanie dokumentu do databázy by lepšie zapadal do dokumentácie otočený o 90° tak, aby v ňom neboli texty zvislé, ale vodorovné. V texte pod obrázkom sa nachádza veta končiaca s výkričníkom, ktorá nie je prípustná v technickej dokumentácii. V tomto prípade ale pravdepodobne ide o preklep.

V popise vyhľadávacieho modulu na str. 31 až 33 je spomenuté, že tento modul slúži okrem vyhľadávania aj na pridávanie a vymazávanie dokumentov. Navrhujeme, aby sa táto funkcionálnosť presunula do samostatného nového modulu určeného na manipuláciu s dokumentmi.

V názve podkapitoly „Metódy zavolať inými modulmi“ uvedenej na str. 33 sa v druhom slove nachádza preklep.

3.4 Modul pre kategórie dokumentu. Časť „Vyhľadávanie dokumentov“ uvedená v popise modulu „Modul pre kategórie dokumentu“ je vhodnejšie presunúť do podkapitoly opisujúcej modul slúžiaci na vyhľadávanie, teda „Vyhľadávací modul“ z toho dôvodu, že samotné vyhľadávanie dokumentov prebieha práve v tomto module, pričom nájdené dokumenty sa usporiadajú na základe ich kategórií, a to zrejme zabezpečí vyhľadávací modul na základe údajov poskytnutých od modulu zabezpečujúceho kategorizáciu dokumentov.

Časť „Výpočet relevancie dokumentu na základe kategórie“ hovorí, že „...Tieto dokumenty, ako aj ich relevancie sa potom ďalej spracovávajú,“ no bližšie nehovorí kde a ani akým spôsobom.

3.5 Modul pre grafové algoritmy. V časti „Komunikácia s okolím“ sa píše, že „Vyhľadávací modul“ opisovanému modulu poskytuje zoznam viacerých dokumentov, ktoré obsahujú najväčší počet kľúčových slov zadaných používateľom. Časť „Činnosť modulu“ ale uvádza, že daný modul „...pošle požiadavku na získanie ID dokumentu ktorý má najväčší počet podobných kľúčových slov.“ Prvá časť hovorí o viacerých dokumentoch, no druhá o jednom dokumente, čo vyvoláva nejednoznačnosť funkcionality opisovaného modulu.

Obr. 18 a 19 znázorňujúce grafy činností majú namiesto koncového stavu znázornený (druhý) začiatkový stav, nedodržiavajúc pravidlá korektnej UML notácie.

3.6 Modul pre sociálne siete. Podkapitola opisujúca „Modul pre sociálne siete“ obsahuje zoznam viacerých atribútov, resp. položiek tvoriacich profil každého používateľa navrhovaného softvérového systému. Tieto atribúty sa ale nevyskytujú ani v logickom, ani vo fyzickom modeli údajov.

Opis funkcionality uvedený v tejto podkapitole ďalej naznačuje, že používatelia majú možnosť pridávať dokumenty do bázy dát, čo podľa diagramu prípadov použitia uvedenom na str. 24 môže vykonávať striktne iba administrátor systému.

V predchádzajúcej časti dokumentácie je na strane 18 v časti “Manuálne definovaná časť modelu používateľa” napísané, že sa predpokladá použitie navrhovaného systému v prostredí podniku. Na str. 38 sa hovorí o prihlasovaní, nie je však konkrétne špecifikované, kedy sa používateľ do navrhovaného systému prihlasuje. Celkový dojem nám hovorí, že používateľ sa bude prihlasovať pred samotným vyhľadávaním dokumentov.

Bolo by vhodné pri prihlasovaní do systému brať do úvahy možnosť, že každý zamestnanec vo firme je prihlásený na počítači, na ktorom práve pracuje. Navrhovaný systém by už toto prihlásenie mohol spracovávať a akceptovať.

? Zhodnotenie

Posudzovanej dokumentácií chýba slovný záver, resp. kapitola, ktorá obsahuje slovné zhodnotenie dokumentácie i celkovej práce na projekte.

4 Použitá literatúra

V posudzovanej dokumentácií chýba odkaz na 7. a 12. zdroj uvedený ako použitá literatúra. Zdroje s poradovými číslami 11 a 12 nemajú uvedené dátumy citácie ich Internetových adries.

CELKOVÉ ZHODNOTENIE

Napriek vytýkaným chybám je podľa nášho názoru posudzovaná dokumentácia vypracovaná na dobrej úrovni a má potenciál viesť k úspešnému a funkčnému systému znalostného manažmentu. Pevne veríme, že naše pripomienky pozitívne prispievajú k celkovej kvalite a funkčnosti výsledného produktu.