

# SW nástroje pro tvorbu znalostní báze

Tomáš Vejlupek

Tovek, spol. s r.o.,  
Chrudimská 2, 130 00, Praha 3  
vejlupek@tovek.cz

**Abstrakt.** Cílem příspěvku je představit řešení nabízené firmou Tovek pod názvem ARMS, „Analytický, Rešeršní a Monitorovací Systém“, a popsat metodiku jeho využití k tvorbě znalostní báze pro „Competitive Intelligence“.

**Klíčová slova:** Competitive Intelligence, vizualizace informací, znalostní báze, poznatková báze, informační báze, Tovek Tools, Analyst's Notebook, iBase

## 1 Úvod

Všeobecně se přijímá tvrzení, že informace jsou nejcennějším zdrojem konkurenceschopnosti firem. To je ovšem pravda pouze v případě, že jsou ve firmě vytvořeny předpoklady k tomu, aby informace bylo možné transformovat na znalosti využitelné pro rozhodování při určování firemních cílů a strategie a taktiky jejich dosahování v konkurenčním prostředí.

Pojem informace nelze oddělit od subjektu (uživatele). Nezávisle mohou existovat pouze data. Z toho plyne jeden z největších problémů využívání informací pro rozhodování v organizacích. Na jedné straně stojí ti, kteří informace zpracovávají či vytváří (*knowledge owners*), a na druhé straně ti, kteří by tyto informace měli využívat (*decision making owners*). To, co je informací na jedné straně, může být pouhým šumem na druhé straně.

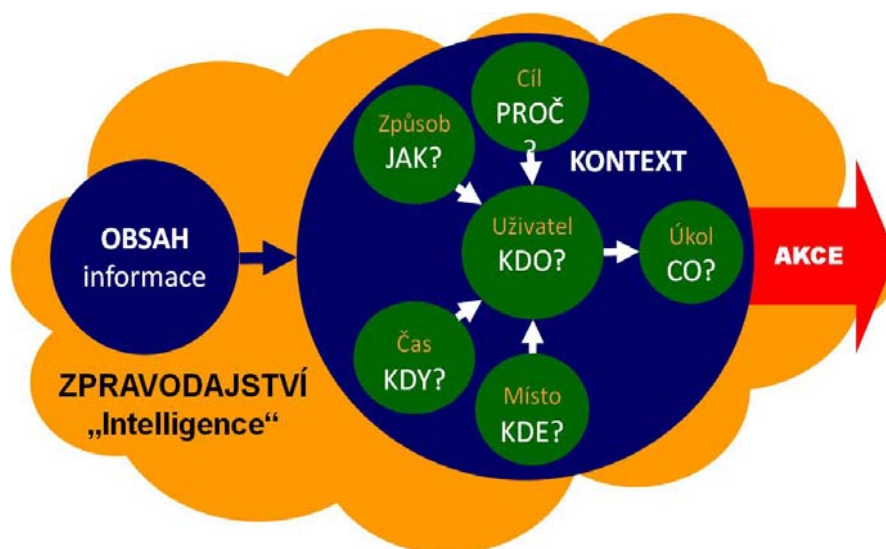
Příjemce dokáže předávanou informaci pochopit a využít pouze tehdy, pokud má pro něj srozumitelnou formu. Jde jednak o formu sdělení (psaný text, obrázek, ústní sdělení, atd.) a jednak o formu poskytnutí kontextu sdělovaného obsahu. Forma prezentace informací se v dnešní době stává stejně důležitým atributem jako jejich vlastní obsah a rychlost doručení.

Velmi známé rčení: „*Jeden obrázek nahradí tisíc slov*“; je důkazem toho, že nejnázornější formou zpracování informací je forma vizuální. Forma prezentace informace je důležitá pro získání potřebné pozornosti příjemce, aby se obsahem informace vůbec začal zabývat, resp. aby do zpracování informace zapojil své mentální kapacity. Na důležitosti proto nabývají nástroje pro vizualizaci informací.

Méně známý již je výrok ředitele institutu pro budoucnost Paula Saffa: „*It is not content, but context that will matter most a decade or so from now. The scarce resource will not be stuff, but point of view*“ [2]. Kontext informace je to, co příjemci umožňuje pochopit její obsah, tj. zařadit jej do svého myšlenkového modelu reality („*mindset*“) a překonat případnou kognitivní disonanci [1].

Informace se pro organizaci tedy stávají výhodou (zbraní) až v okamžiku, kdy existuje proces organizovaného poskytování informací, jehož výsledkem jsou informační produkty („*Intelligence*“), dávající informace o konkurenčním okolí organizace do kontextu s jejími cíli a strategií (*So What?*), resp. do kontextu s individuálními odpovědnostmi jejich manažerů (*What Is In For Me?*). Z hlediska zpracování obsahu informací je nutné rozlišit dvě kategorie kontextů:

- *Objektivní kontext*, který lze do jisté úrovně vytvářet z obsahu informace automatizovaně (metody extrakce pojmů):
  - konkrétní: osoby, organizace, věci, místa, události, zdroje
  - abstraktní: témata, kategorie.
- *Subjektivní kontext*, který může být pouze výsledkem interpretace obsahu informací analytikem, resp. uživatelem:
  - procesní (cíle, klíčové oblasti, úkoly, pravidla, postupy, atd.),
  - vztahový (konkurent, dodavatel, odběratel, riziko, atd.)



**Obrázek 1:** Atributy využitelné informace – zpravodajství (Intelligence)

Poskytování informací v nevhodné formě vede buď k informačnímu stresu příjemců, nebo v krajním případě až k odmítání nových informací a upřednostňování vlastního, nutně zkresleného úsudku (obrazu reality).

Problematikou organizovaného informování lidí, kteří ve firmách rozhodují, za účelem podpory rozhodování souvisejícího se získáváním konkurenčních výhod firmy, se zabývá obor nazývaný „*Competitive Intelligence*“ (CI) [8,9].

Základem CI je tzv. zpravodajský cyklus, který má čtyři hlavní fáze:

- identifikace informačních potřeb (tvorba sdíleného kontextu)
- sběr informací (vyhledávání, formalizace, kategorizace)
- analýza (a syntéza) informací (interpretace informací)
- komunikace informací (doručení / sdílení a prezentace)

## 2 Co je ARMS

ARMS je akronymem slov „Analytický, Rešeršní a Monitorovací Systém“. Systém tvoří kombinace několika profesionálních nástrojů k vyhledávání a analýze informací a jejich propojení na různé interní a externí informační zdroje.

Účelem systému je především zvýšení efektivity a kvality práce lidí, jejichž úkolem je vytvářet nové informace pro podporu rozhodování. ARMS umožňuje:

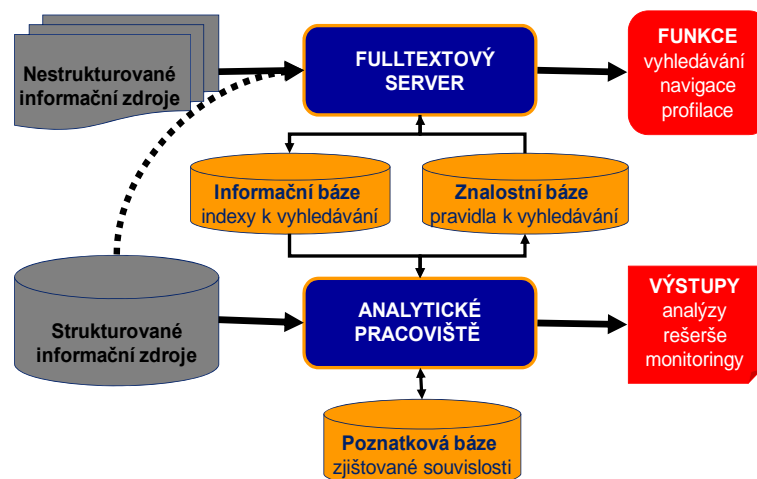
- vytvořit rychle názorný obraz problému, který je předmětem rozhodování,
- poskytovat rychle nové informace k problému v potřebném kontextu,
- uchovávat získávané poznatky pro rozvoj znalostí o dané oblasti.

Jak vyplývá z názvu, ARMS má tři hlavní funkce:

- *rešeršní funkce* umožňují z různorodých informačních zdrojů rychle vyhledat informace k danému problému,
- *analytické funkce* umožňují vyhledané informace utřídit, extrahovat z nich relevantní poznatky a v názorné formě je prezentovat,
- *monitorovací funkce* umožňují aplikovat znalost o problému a o potřebách rozhodování k vyhledávání a kategorizaci nových informací.

Klíčovou funkcí ARMS je ale systematické vytváření přidané hodnoty, která tvoří obsah třech datových úložišť:

- *informační báze* tvořená indexy umožňujícími jednotné a efektivní prohledávání různorodých nestruturovaných informačních zdrojů,
- *poznatková báze* tvořená databází entit (osoby, organizace, věci, místa, události, proces, pravidla, témata) a jejich atributů a vztahů,
- *znalostní báze* tvořená strukturovanými dotazy pro vyhledávání a kategorizaci informací v informační bázi.



Obrázek 2: Fyzická architektura ARMS

### 3 Produkty tvořící ARMS

Funkce ARMS jsou realizovány pomocí produktů, které byly původně určeny pro analytiku zpravodajských služeb a policejní vyšetřovatele. Možnosti vyhledávání a analýzy informací jsou výsledkem těch nejnáročnějších potřeb a jdou daleko za možnosti běžného vyhledávání nebo jednoduché vizualizaci souvislostí, které dnes nabízí řada firemních aplikací pro podporu rozhodování.

Odůvodněnost použití těchto profesionálních nástrojů pro podporu rozhodování v konkurenčním boji je naprosto zřejmá, neboť stejně tak jako zpravodajské služby, musí být i firmy schopny vyhodnocovat velké množství informací z různých zdrojů a činit rozhodnutí na základě kusých faktů a slabých signálů.

Pro vytěžování strukturovaných dat (databáze a poznatky), jsou použity produkty Analyst's Notebook, iBase a iBridge od fy. i2 [7].

Pro vytěžování nestrukturovaných (textových) dat, jsou použity produkty Tovek Tools a Tovek Server, které jsou založeny na technologii fy. Verity (Autonomy). Propojení analýzy nestrukturovaných a strukturovaných dat je založeno na modulu Tovek Tools plug-in pro Analyst's Notebook.

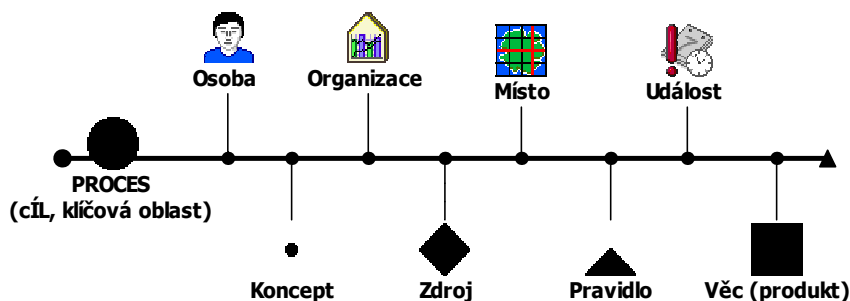
- *Tovek Server*: Indexuje obsah různorodých informačních zdrojů a poskytuje webové služby pro vyhledávání, filtrování a kategorizaci informací.
- *Tovek Tools*: Umožňuje přesné vyhledávání informací v souborech textových dat (indexovaných lokálně nebo prostřednictvím Tovek Serveru) a nabízí unikátní metody obsahové a kontextové analýzy vyhledaných informací.
- *Analyst's Notebook*: Umožňuje efektivní analýzu a názornou prezentaci velkého množství faktů a souvislostí založenou na různých formách jejich vizuální prezentace a na funkcích pro vyhledávání duplicit, shluků, periodicit, atd.
- *iBase*: Umožňuje systematické shromažďování a uspořádávání velkého množství různorodých poznatků a jejich následnou analýzu pomocí Analyst's Notebook (vztahy, souslednost), GIS plug-in (poloha) či Dataminer (atributy).
- *iBridge*: Umožňuje propojení Analyst's Notebook na obsah databází. Nabízí se propojení na databáze ekonomických vazeb firem CreditInfo a Čekia.

### 4 Poznatková báze ARMS pro „Competitive Intelligence“

V rámci projektu "Znalostní báze pro vyhledávání a rozvoj synergických podnikatelských seskupení", který realizovala firma TOVEK společně s ČVUT v Praze, byly produkty tvořící ARMS použity pro zmapování struktury klastru OMNIPACK, jeho produktů, informačních potřeb a relevantních informačních zdrojů. Na základě toho pak byla vytvořena znalostní báze umožňující k vytipovaným tématům vyhledávat aktuální informace.

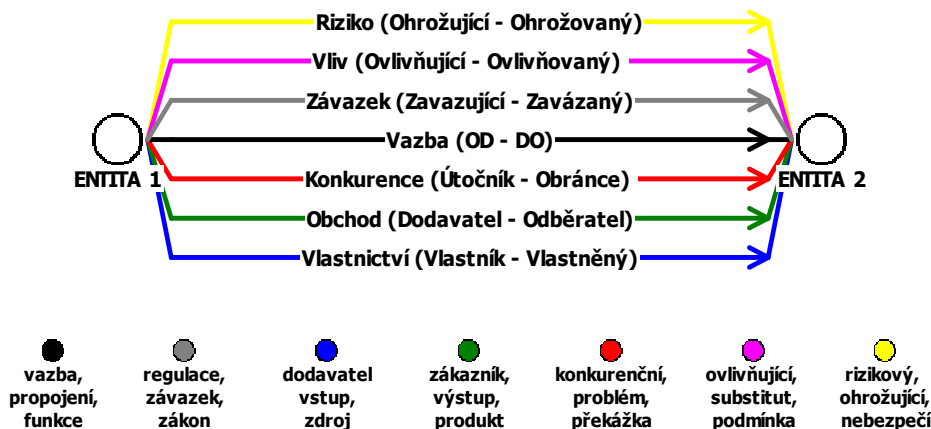
V návaznosti na zkušenosti z tohoto projektu pak byla vytvářena obecnější metodika pro podporu „Competitive Intelligence“. Metodika vychází z obecných principů zpravodajství jako je zpravodajský pentagram a specificky pak z metod analýzy konkurenčního prostředí jako je: 5PFs analýza (Five Porters Forces), STEEP analýza (Society, Technology, Economy, Environment, Policy) [3]. Základem této metodiky je struktura poznatková báze, která umožňuje organizovat poznatky získávané zpracováním různých informací. Poznatková báze definuje:

- 9 typů entit (osoby, organizace, věci, místa, události, procesy, pravidla, témata, zdroje),
- 7 typů vztahů (obecná vazba, vlastnictví, dodavatel-odběratel, konkurence, regulace, vliv, hrozba).



Obrázek 3. Poznatková báze vytvořená v iBase

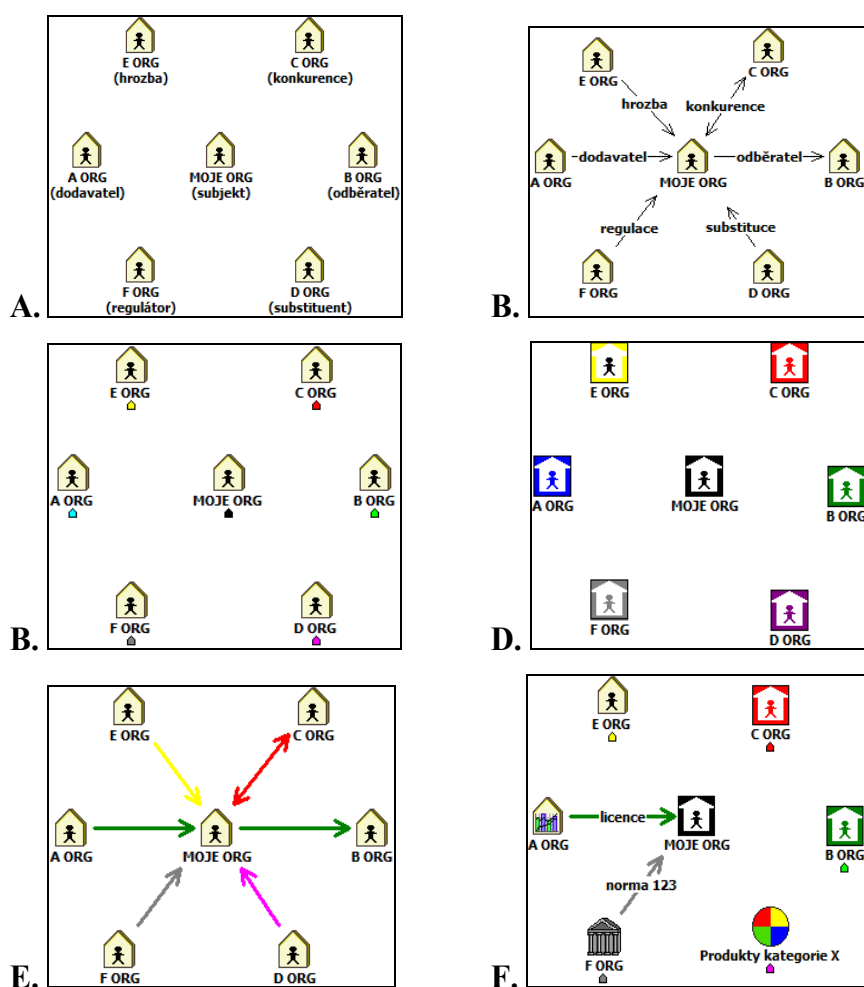
Klíčovým významem poznatkové báze je především zachycování souvislostí mezi entitami. Ty mohou být vyjádřeny buď explicitní vazbou, nebo specifickým (barevným) atributem, který reprezentuje vztah k subjektu, pro nějž je poznatková báze vytvářena. Pro potřeby „Competitive Intelligence“ byly zvoleny tyto typy vazeb a atributů.



Obrázek 4. Explicitní a implicitní vyjádření vztahů v poznatkové bázi

Jak vyplývá z označení „poznatková báze“, je účelem této databáze zaznamenávat poznatky, tj. elementární části informací, které analytik, resp. skupina analytiků, zpracovává. Poznátky tvoří především vztahy mezi entitami, představující model reality. Následně tak lze z poznatkové báze vytvářet nové informace, které jsou tvořeny poznatky vyhledanými a propojenými v kontextu aktuální potřeby.

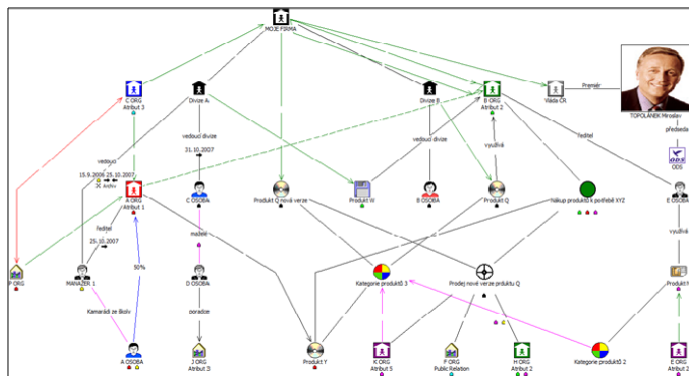
Hlavní silou analytických nástrojů ARMS je především vizuální prezentace poznatků, která umožňuje poznatky různým způsobem třídit, seskupovat, filtrovat a znázorňovat tak, aby analytikův mozek byl schopen odvodit využitelné závěry a aby je konečný uživatel byl schopen pochopit a využít. Je také na analytikovi, aby z řady možností, které systém k reprezentaci poznatků nabízí, zvolil ten nejnázornější.



Obrázek 5. Různé způsoby záznamu vazeb mezi organizacemi

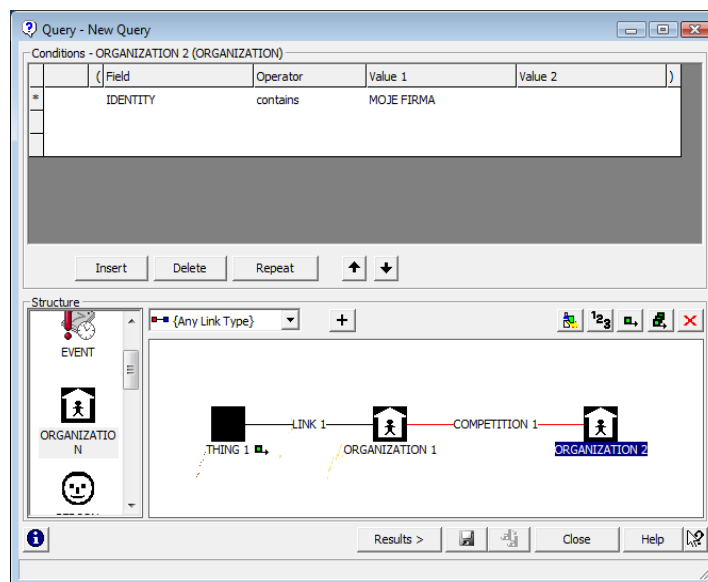
Obsah poznatkové báze, který je relevantní k řešení určitého problému, je typicky reprezentován vztahovým diagramem.

Pro uspořádání entit v diagramu je možné volit některé z automatických rozložení (vějířové, skupinové, hierarchické, s minimem křížení vazeb), které pak lze dále upravovat tak, aby vynikly podstatné souvislosti. Mezi možné úpravy diagramu patří zvětšování ikon a tloušťky a formy čar, změna způsobu grafické reprezentace entit (rámeček, textový blok, fotografie), barevné stínování, vložení podkladu (bitmapa). Důležitá je zejména možnost změnit reprezentaci entit na časové linky a na události mezi nimi, tak že lze sledovat také souslednosti událostí a vazeb.



Obrázek 6. Vztahový diagram

Postupem času se v poznatkové bázi nashromáždí velké množství poznatků a iBase proto disponuje řadou funkcí, které poznatky umožňují snadno prohledávat, filtrovat, shlukovat do skupin s určitými charakteristikami, a které také umožňují vyhledávat duplicity a shodné či podobné záznamy spojovat.



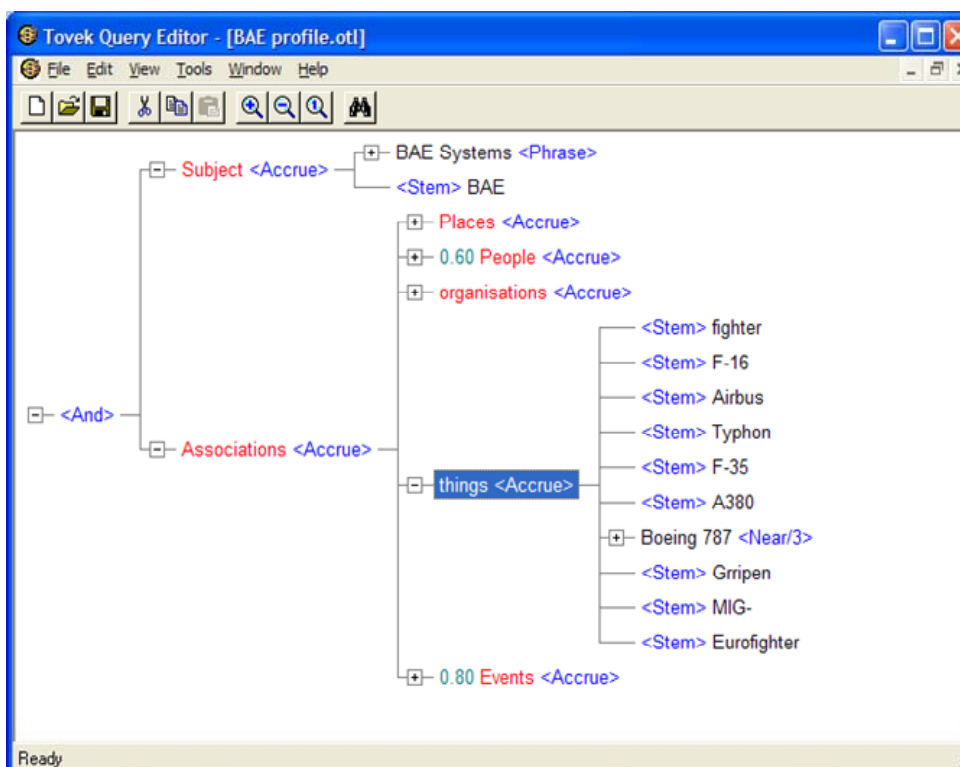
Obrázek 7. Vizualní dotazy v poznatkové bázi



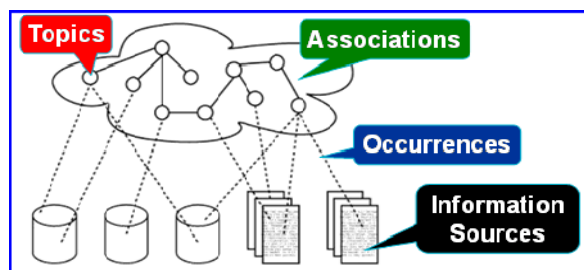
## 4.2 Tvorba znalostní báze

Poznatková báze je především nástrojem analytika. Znalostní báze je produktem analytika pro koncové uživatele. Představuje znalost analytika v podobě expertních dotazů, pomocí kterých je běžný uživatel, resp. Tovek Server, schopen vyhledávat a kategorizovat informace relevantní k definovaným tématům.

Expertní dotazy mohou být organizovány v podobě „Topic Maps“ [6], kde představují „dynamickou vazbu na informace“ (occurrence).



Obrázek 10. Expertní dotaz (Topic)



Obrázek 11. Znalostní báze v podobě mapy témat (Topic Map)

## 5 Závěr

ARMS v rukách dobrého analytika a v napojení na potřebné informační zdroje může představovat skutečnou informační zbraň, pomocí které mohou manažeři firmy trvale získávat informace, které jím umožňují získávat konkurenční výhody [4].

Prototyp řešení byl vyvinut v rámci grantu AV ČR na příkladě informačních potřeb klastru OMNIPACK. Další rozvoj probíhá v rámci projektu „Porter“ ve spolupráci Tovek, SCIP CZECH [9] a VŠE Praha [10].

## Poděkování

Příspěvek byl vytvořen za podpory grantového projektu 1ET210170503 s názvem „Znalostní báze pro vyhledávání a rozvoj synergických podnikatelských seskupení“ řešený v programu „Informační společnost“ Akademie věd ČR.

## Reference

1. Heuer, Richards Jr. *Psychology of Intelligence Analysis*. CIA, Langley (1999). ISBN 1-929667-00-0.
2. Saffo Paul. Institute for the Future. *It's the Context, Stupid*. [www.wired.com/wired/archive/2.03/context.html](http://www.wired.com/wired/archive/2.03/context.html)
3. Alessandro Comai; Joaquin Tena Millan. *Mapping & Anticipating the Competitive Landscape*. Emecom Ediciones, Barcelona (2006) . ISBN-13 978-84-935178-7-8; ISBN-10 84-935178-7-9
4. Straka Lubomír. Znalostní podpora síťování firem v klasrtu Omnipac. Sborník konference „Znalostní báze pro podporu řízení podnikatelských seskupení 2007“, Praha (2007). ISBN 978-80-01-03891-8.
5. Tovek, spol. s r.o. [www.tovek.cz](http://www.tovek.cz)
6. i2 Group. [www.i2group.com](http://www.i2group.com)
7. Bouvet / Ontopia. [www.ontopia.net](http://www.ontopia.net)
8. Society of Competitive Intelligence Professionals. [www.scip.org](http://www.scip.org)
9. SCIP CZECH. [www.scip-czech.cz](http://www.scip-czech.cz)
10. Tomáš Kliegr, Jan Rauch, Vojtěch Svátek, Jiří Šplíchal, Tomáš Vejlupek. Omnipack Case Study. *Sborník konference Znalosti 2008*. Bratislava (2008).

## Annotation:

### *SW Tools for Knowledge Base*

This paper introduce Tovek's solution „Analytical Research & Monitoring System” based on tools for visual analysis and presentation of information and a methodology of its usage to support process of „Competitive Intelligence“.