



Tovek Tools Analyst Pack je desktopovou aplikací určenou k vyhledávání informací, tvorbě různých typů analýz a vytváření přehledů a rešerší. Jsou vhodné pro práci i s velkým objemem textových dat z různorodých informačních zdrojů.

Popis

Tovek Tools využívají špičkovou technologii pro vyhledávání informací. Umožňují prohledávat a analyzovat lokální informační zdroje - souborový systém, elektronickou poštu nebo ODBC databáze, stejně jako i různá speciální data jako např. přepisy hlasových záznamů nebo data z mobilních telefonů. Mohou být také provozovány jako klient **Tovek Serveru** a pracovat s jeho existujícími zdroji.

Přínosy a výhody

Přesné a úplné vyhledávání

Tovek Tools využívají špičkovou fulltextovou technologii, která zaručuje snadné, spolehlivé a velmi výkonné vyhledávání i ve velkých objemech dat. Přesnou formulaci dotazu umožňuje komplexní dotazovací jazyk. Pokročilé funkce jsou průběžně vylepšovány na základě požadavků a připomínek profesionálních uživatelů.

Kombinace více informačních zdrojů

Pomocí Tovek Tools lze vytvořit vlastní zdroje i připojit se k různým vzdáleným Tovek Serverům a prohledávat, analyzovat a kombinovat výsledky vyhledávání do jednoho výstupu.

Snadné vyhledávání dokumentů v různých jazycích

Fulltextový dotaz se může skládat z více dílčích dotazů v různých jazycích, které jsou použity na dokumenty v příslušném jazyce. Výsledky vyhledávání jsou sloučeny do jednoho společného výstupu, se kterým lze dále pracovat.

Široké možnosti analýzy

Výsledky vyhledávání mohou být dále analyzovány pomocí kontextové a obsahové analýzy a přehledně graficky znázorněny. Uživatel tak může snadno odhalovat souvislosti nebo identifikovat nové trendy a témata. Obsah dokumentů a strukturované zdroje lze také vizualizovat ve formě diagramů vztahů.

Zobrazují výstupy v přehledné formě

Výsledky vyhledávání nebo analýz mohou být zobrazeny v přehledné formě různých výstupních sestav a ukládány v různých formátech pro další využití.

Typické oblasti využití

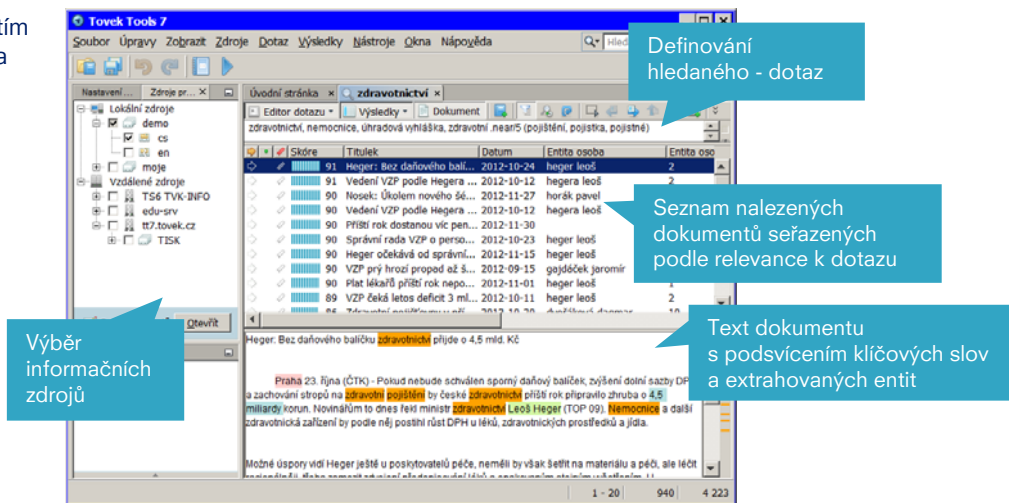
Tovek Tools se nejčastěji využívají ve zpravodajství, ve výzkumu a vývoji, při vyšetřování a odhalování podvodů nebo mapování konkurenčního prostředí. V komerčním sektoru se využívají zejména v mediálních agenturách, u poskytovatelů informací nebo ve finančním sektoru v bankách a pojišťovnách. Ve veřejné správě se využívají hlavně ve vládních a veřejných institucích, na ministerstvech, v policejních složkách a v bezpečnostní a zpravodajské komunitě.



Vyhledávání

- Hledání dokumentů, které vyhovují zadaným podmínkám dotazu. Použitá technologie umožňuje využít komplexní dotazovací jazyk:
- ▶ ohodnocení relevance vyhledaných dokumentů a jejich seřazení
 - ▶ vyhledávání v různých jazycích současně
 - ▶ automatické vytvoření anotace dokumentů
 - ▶ použití fragmentů volného textu ve formulaci dotazu
 - ▶ vyhledávání podobných dokumentů

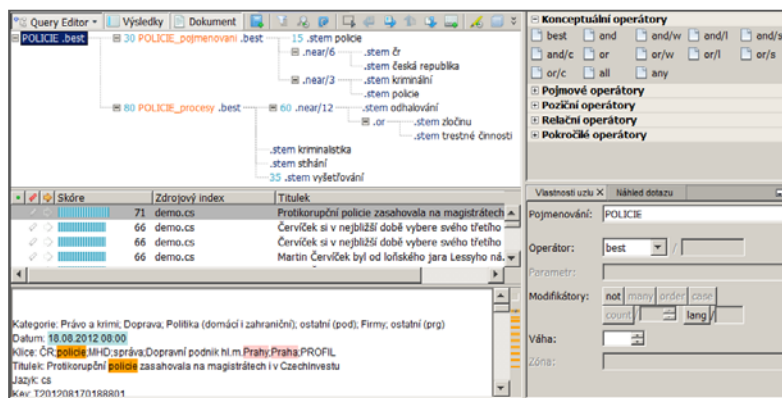
Vyhledávání s využitím dotazovacího jazyka Tovek



Query Editor

Query Editor je určen pro přehlednější vytváření hierarchicky strukturovaných dotazů. Lze přesně zadat podmínky hledání a pomocí vah jednotlivých částí dotazu určit co je nejdůležitější a ovlivnit tak výsledné pořadí nalezených dokumentů.

Strukturovaný dotaz v Query Editoru



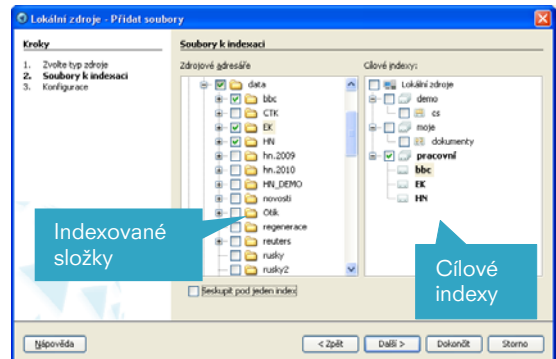
Správa zdrojů

Správa zdrojů umožňuje připojení k již existujícím zdrojům. Přímou v prostředí Tovek Tools lze indexovat různé typy souborů, databázových záznamů a e-mailových zpráv. Při indexaci se využívají rychlé a spolehlivé filtry podporující všechny běžné formáty dat včetně prepisů hlasových záznamů nebo dat z mobilních telefonů.



Uživatel tak může z jednoho rozhraní rychle vyhledávat informace z různorodých zdrojů.

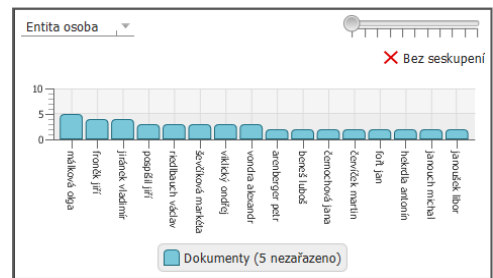
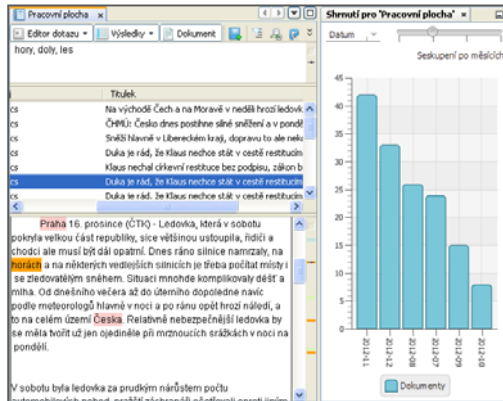
Při indexaci v Tovek Tools Analyst Pack lze zvolit extrahování entit, tj. automatické označení např. osob, telefonních čísel, měst apod. v textu, obdobně jako při indexaci na Tovek Serveru.



Shrnutí

Příklad shrnutí dle data a extrahovaných entit osob

Shrnutí umožňuje vytvořit grafy podle různých polí dokumentů, např. podle data nebo nalezených entit a získat tak celkový přehled o vlastnostech nalezených dokumentů.

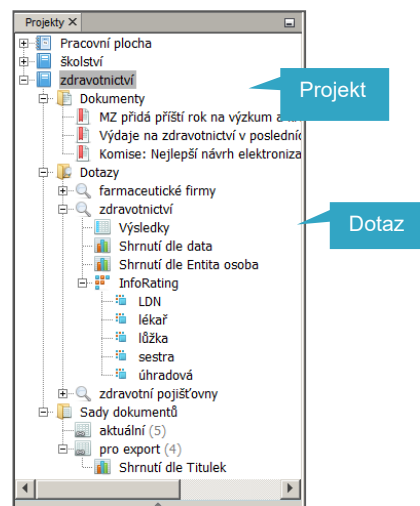


Projekty

V Tovek Tools lze najednou pracovat s více dotazy, které mohou být přehledně uloženy v jednotlivých projektech.

Spolu s nimi mohou být uloženy:

- ▶ seznamy vybraných dokumentů např. pro pozdější export
- ▶ texty dokumentů, ke kterým má uživatel přístup i pokud již nemá k dispozici zdrojová data
- ▶ shrnutí
- ▶ kontextové analýzy (InfoRating)
- ▶ diagramy



Exporty

Pro předání vyhledaných dokumentů nabízí Tovek Tools různé varianty přehledných výstupů v řadě běžných formátů (txt, html, xml) včetně shrnutí a Info Ratingu.



Části exportu vybraných dokumentů ve formátu html se shrnutím a kontextovou analýzou InfoRating

EXPORT 2014-05-02 16:57:12

Titulek	lékař	lůžka	úhradová	Datum
Plat lékařů příští rok neporoste, nezvýší se ani úhrady za péči		70	0	84 2012-11-11
Správní rada VZP o personálních nejednala, spor byl o peníze		40	0	80 2012-10-10
Příští rok dostanou víc peněz centra a domácí péče		0	84	80 2012-11-11
Nejprůnosnější podle expertů je rušení lůžek a elektronizace		0	95	0 2012-11-28

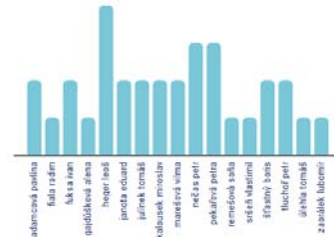
Počet výsledků: 4

Seznam vybraných dokumentů se zvolenými poli

Kontextová analýza



Shrnutí dle Entita osoba



Příští rok dostanou víc peněz centra a domácí péče

Titulek: Příští rok dostanou víc peněz centra a domácí péče
 lékař: 0
 lůžka: 84
 úhradová: 80
 Datum: 2012-11-30

Kategorie: Zdravotnictví; Peněžnictví, banky, spořitelny, fondy
 Datum: 30.11.2012 08:00
 Klíče: ČR;zdraví;pojišťovny;úhrady
 Titulek: Příští rok dostanou víc peněz centra a domácí péče
 Jazyk: cs
 Key: T201211290684801
 Text:

Texty dokumentů se zvýrazněnými nalezenými slovy

Příští rok dostanou víc peněz centra a domácí péče

Praha 29. listopadu (ČTK) - Víc peněz pro centra specializované péče, kde se léčí například rakovina nebo roztroušená skleróza, podle návrhu vyhlášky, podle níž by měly příští rok platit zdravotní pojišťovny za léčení. Návrh vyhlášky poskytl ČTK ministerský úřad. Návrh vyhlášky poskytl ČTK ministerský úřad. Návrh vyhlášky poskytl ČTK ministerský úřad.

"Úhradová vyhláška je koncipována jako restriktivní. Drobnou restrikci a uskromnění musí každý velký systém typu ve zdravotnictví," uvádí se v návrhu.

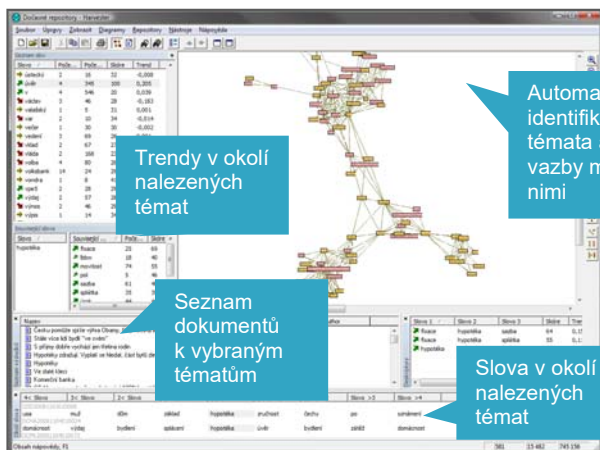
Harvester

Harvester umožňuje uživatelům provádět obsahovou analýzu nestructurovaných textů zejména pro účely:

- ▶ automatizovaného určení témat dokumentů
- ▶ výběru dokumentů s podobnými tématy
- ▶ identifikaci nových témat v souborech dokumentů

Témata jsou identifikována na základě statistické analýzy výskytu blízkých dvojic a trojic vhodných slov, jejichž varianty spojení je možné graficky znázornit.

Obsahová analýza automaticky identifikuje nová témata z nestructurovaných textů



Trendy v okolí nalezených témat

Automaticky identifikovaná témata a vazby mezi nimi

Seznam dokumentů k vybraným tématům

Slova v okolí nalezených témat



Inforating

InfoRating umožňuje uživateli provádět kontextovou analýzu nalezených dokumentů, která najde a znázorní souvislosti mezi obsahem vyhledaných dokumentů a definovanými tématy. Výsledek je znázorněn ve formě přehledné matice nebo grafu. InfoRating mohou uživatelé využít i pro zjišťování vztahů mezi jednotlivými tématy nebo pro tvorbu strukturovaných rešerší s křížovými referencemi.

Kontextová analýza InfoRating, která znázorňuje souvislost mezi obsahem vyhledaných dokumentů a definovanými tématy

Definovaná témata

Titulek	lékař	lůžka	sestra	úhradov...
Heger: Bez daňového balí...	40	80	0	80
Zákon o zdravotních pojiš...	0	80	0	80
Správní rada VZP o perso...	40	0	0	80
Výdaje na zdravotnictví v ...	0	0	0	0
Ministerstvo vyhlásilo sout...	40	0	0	0
MZ přidá příští rok na výzk...	0	0	0	0
Odbory a zaměstnavatelé ...	0	0	0	0
Nejprůběžnější podle expe...	0	0	0	0

Seznam dokumentů a jejich příslušnost k definovaným tématům

Průsečky dotazů znázorňují souvislost mezi obsahem vyhledaných dokumentů a definovanými tématy

180	7	7	25	lékař
75	7	3		lůžka
7	0			sestra
17				úhradová vyhláška

Všechny dokumenty

Shrnutí pro 'zdravotnictví' X

data Entita osoba

Datum

Seskupení po měsících

2012-10, 2012-11, 2012-12, 2012-07, 2012-08, 2012-09

133 hodnot s 672 výskity

Osoby extrahované k tématu

nečas petr, nosek petr, roda jiri, roda jiri, rihová ludmila, sejka michal, štrání boris, žitkovič dagmar, fiata petr, fiata stanislav, gajdášek jeromír, heger leoš, hegera leoš, horák pavel, kabátěk zdeněk, krása václav, moravec václav

Témata v čase

Seskupení po měsících

2012-07, 2012-08, 2012-09, 2012-10, 2012-11, 2012-12

lékař, lůžka, sestra, úhradová vyhláška

Graf souvislostí témat a skupiny dokumentů

lékař, doktor, lůžka, sestra, úhradová vyhláška

Libye by prý stála o stáze ... 70 80 91

Asi ve třetině nemocnic v ... 46 89 80

Menší nemocnice jsou spo... 40 97 80

Kategorie: Zdravotnictví, Makroekonomika, Firmy
Datum: 12.07.2012 08:00
Klíče: ČR,Libye,zdraví,nemocnice,lekař,hospodářství,firmy
Titulek: Libye by prý stála o stáze svých lékaři v ČR i o české doktor
Jazyk: cs
Key: T201207110609601
Text:

Shrnutí pro 'zdravotnictví' X

data Entita osoba

Datum

Bez seskupení

2012-09-25, 2012-11-21, 2012-07-12

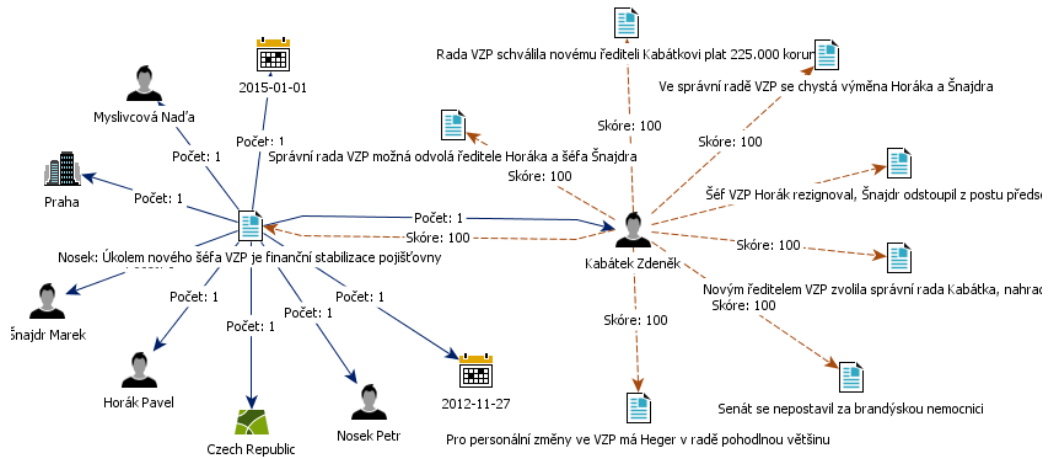
Dokumenty ke zvolené trojici témat



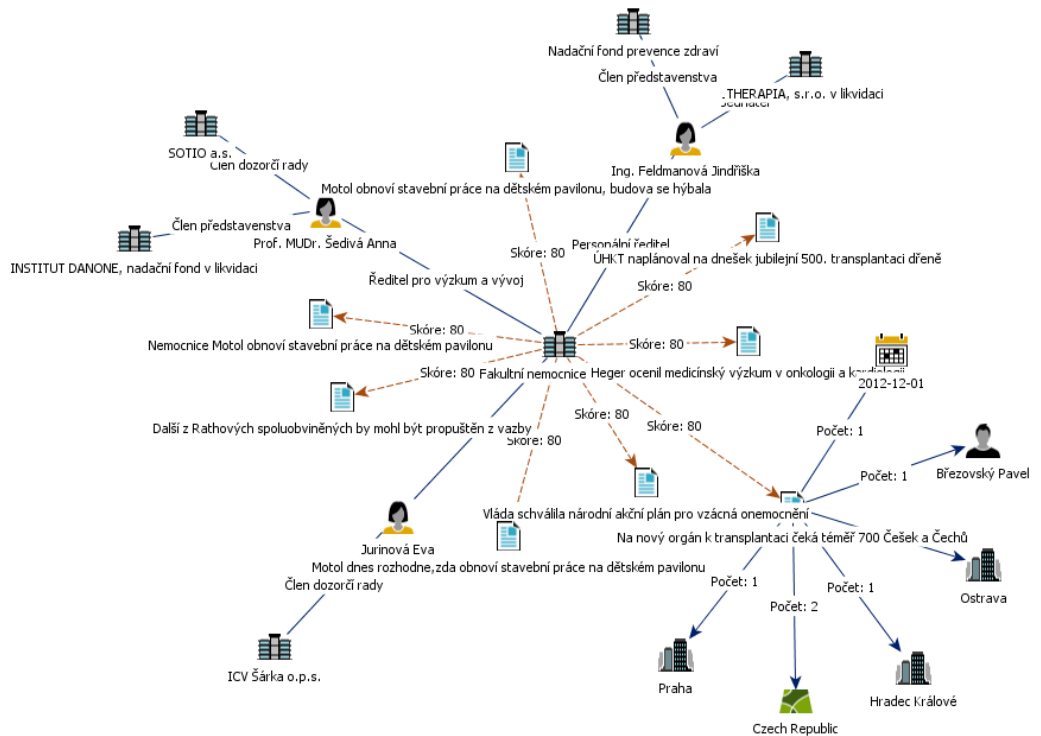
Vizualizace

Na základě extrahovaných entit lze zobrazovat obsah vybraných dokumentů ve formě diagramu vztahů. Umožňuje také vytvářet speciální strukturované zdroje, kde jsou přímo definovány objekty a vztahy mezi nimi. Oba způsoby se dají kombinovat. Vytvořené diagramy lze podle potřeby upravovat a také exportovat.

Vizualizace vyhledaných dokumentů



Vizualizace strukturovaného zdroje v kombinaci s dokumenty





Fulltext Plug-in pro IBM i2 Analyst's Notebook

S pomocí modulu Fulltext Plug-in pro IBM i2 Analyst's Notebook mohou uživatelé zobrazit a analyzovat textové informace v IBM i2 Analyst's Notebooku, který je špičkovým produktem pro vizuální analýzu. Fulltext Plug-in nabízí aktivní propojení mezi prvky diagramu v IBM i2 Analyst's Notebooku a informacemi uloženými v textových souborech. Umožňuje také využívat fulltextové metody vyhledávání přímo v prostředí IBM i2 Analyst's Notebooku.

Propojení mezi prvky diagramu v IBM i2 Analyst's Notebooku a informacemi uloženými v textových souborech

The screenshot displays the IBM i2 Analyst's Notebook interface. On the left, a network diagram shows nodes representing entities and documents, connected by lines with numerical weights. Callouts point to specific elements: 'Související dokumenty' (Related documents), 'Vazby na další subjekty' (Links to other subjects), 'Analyzovaný subjekt' (Analyzed subject), and 'Text dokumentu s podsvícením klíčových slov a extrahovaných entit' (Text document with highlighting of key words and extracted entities). On the right, a text document is open, showing a snippet of text with key words highlighted in yellow and extracted entities in blue.

Kontakt

TOVEK, spol. s r.o.
Chrudimská 1418/2, 130 00 Praha 3
Tel.: +420 222 728 129, Fax: +420 222 728 103
E-mail: info@tovek.cz, Web: www.tovek.cz