



**Tovek Tools is a set of desktop applications for information retrieval and analysis and creating different kinds of overviews and summaries. It is suitable for processing even large volumes of unstructured data from miscellaneous information sources.**

---

### Description

Tovek Tools is based on a powerful technology for searching for information. It enables to search and analyse local information sources - file systems, e-mails or ODBC databases, as well as a special data such as speech-to-text transcriptions or data from mobile phones. It can be used as a **Tovek Server** client and work with its existing sources.

---

### Benefits and Advantages

#### Accurate and complete search

Tovek Tools uses powerful full-text technology ensuring easy, reliable and very powerful searching even in large volumes of data. Complex query language enables to precisely define a query. Advanced functions are continuously improved based on professional user requests and suggestions.

#### Combination of several information sources

Using Tovek Tools you can create custom sources as well as connect to remote Tovek Servers and search, analyse and combine search results to one output.

#### Easy document search in different languages

A fulltext query can consist of several sub-queries in different languages which are used on documents in the relevant language. The search results are combined into one common output for further work.

#### Wide range of analysis

The search results can be then analysed using context and content analyses and clearly visualised. The user can thus easily discover relationships or identify new trends and topics. Document content and structured sources can be also visualised in a chart form.

#### It displays outputs in well-arranged form

The search or analysis results can be displayed in the form of well-arranged output print-outs and stored in different formats for further use.

---

### Typical Areas of Use

Tovek Tools is typically used in intelligence, in R&D departments when investigating and discovering fraud or mapping out the competitive environment. In the commercial sector it's mainly used in media agencies, at the information providers or in the finance sector in banks and insurance companies. In the government sector it's mainly used in government and public institutions, at ministries and in the law enforcement, defence and intelligence community.



### Searching

Search of documents which comply with set query conditions. This functionality enables to use complex query language:

- ▶ evaluation of relevance to retrieved documents and their sorting
- ▶ multilingual search
- ▶ automatic document annotation generating
- ▶ using fragments of free-text in the query text
- ▶ search for similar documents

Searching using Tovek query language

Search criteria definition - a query

List of retrieved documents ordered according to relevance to the query

Document text with highlighted searched words and extracted entities

Information source selection

Score	Source	Index	Date	Title
83	demo.en	83	2012-12-20	Javelin thrower Spotakova
83	demo.en	83	2012-12-20	Javelin thrower Spotakova
71	demo.en	71	2012-10-28	Czech tennis official invites
70	demo.en	70	2012-10-10	Czech hockey goalie Hasek d
68	demo.en	68	2012-12-11	World Hockey Champ in 201
67	demo.en	67	2012-12-06	Jagr not to play for Czechs
67	demo.en	67	2012-11-07	Czech tennis team unchanged
67	demo.en	67	2012-10-03	Czech-Spain Davis Cup final
67	demo.en	67	2012-08-24	Czech President decorates successful
67	demo.en	67	2012-11-15	Results of Czech ice hockey Elita team

### Query Editor

Query Editor is designed for well-arranged hierarchically structured queries. You can precisely specify search criteria and weights for each query part and so specify what is most important and affect the resulting retrieved document order.

Structured query in Query Editor

Structured query in Query Editor

Query Editor interface showing a structured query tree and search results.

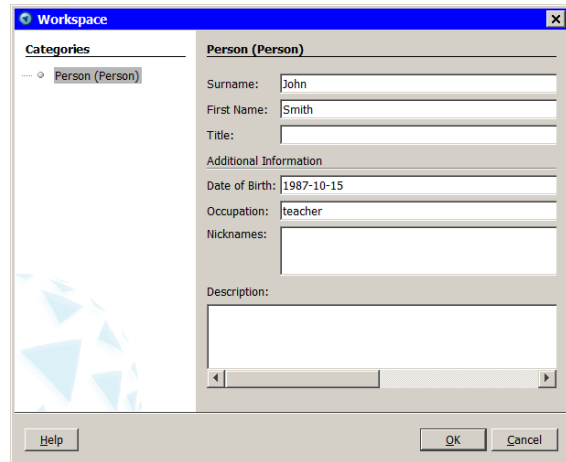
Score	Title	Date	de
64	Czech govt approves action plan for rare diseases	2012-08-30	
40	Beds reduction, computerisation best Czech health care steps	2012-11-28	
40	Microsoft has best Czech health care computerisation proposal	2012-11-21	
40	Czech healthcare unions call on minister to resign	2012-11-14	
40	Czech healthcare unions call on minister to resign	2012-11-14	

### Knowledge Base and Forms

Knowledge base is used for storing queries and other items (forms) and their easy sharing between Tovek Server users. Depending on license, local Knowledge Base used by one user and shared Knowledge Bases of particular Tovek Servers used by more users can be available. Work possibilities vary according to the settings. Work of more users during the actualisations and queries saving into the Knowledge Base is ensured. Some queries can be marked as topics and called by names in the queries.



Forms are intended for users who don't need to know query language and query creation. It's for users who fill in only specific details (for example name of person and its date of birth.)

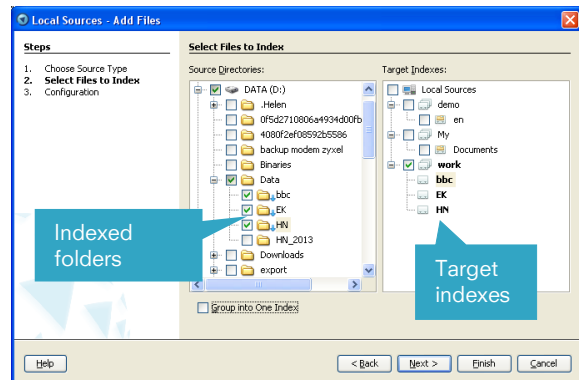


### Source Management

Source management enables connection to already existing sources. You can index different file types, database records and e-mails directly in the environment of Tovek Tools. Fast and reliable filters supporting all common data formats are used for indexing, including speech- to text recognition or data from mobile phones.

So the user can fast search information from different sources from one interface.

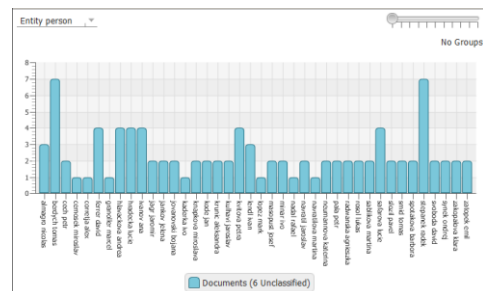
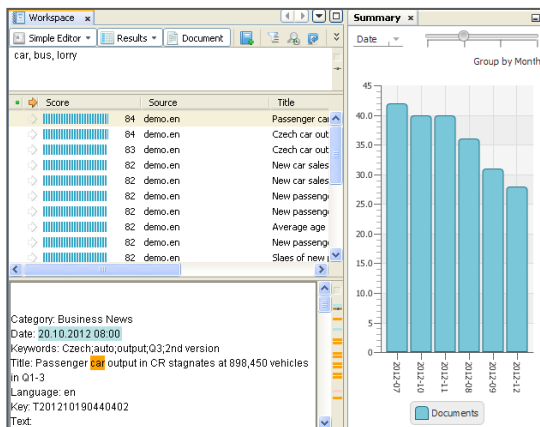
When indexing in Tovek Tools Analyst Pack, the entity extraction can be selected, i.e. automatic marking e.g., persons, telephone numbers, towns etc. in text, similarly to indexation by Tovek Server.



### Summary

The Summary enables the creation of graphs according to document fields, e.g., according to date or found entities and so get complex overview of found document characteristics.

A Summary example according to date and extracted person entities



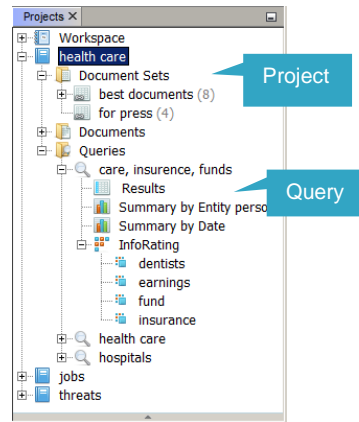


### Projects

You can work with several queries at once in Tovek Tools. The queries can be saved in individual projects.

The following can also be saved to projects:

- ▶ lists of selected documents, e.g. for further export
- ▶ document texts to which the user has access even if he/she doesn't have the source data available
- ▶ summary
- ▶ context analyses (InfoRating)
- ▶ charts



### Exports

Tovek Tools offers several well-arranged output variants in several common formats (txt, html, xml) including summary and InfoRating for retrieved document delivery.

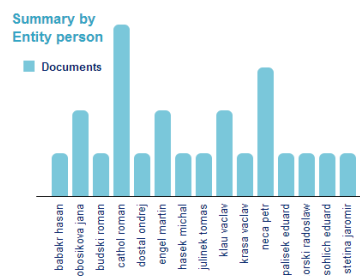
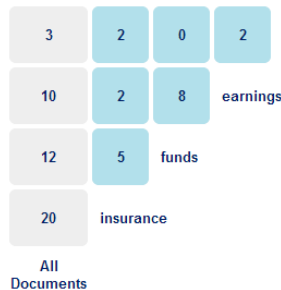
A part of selected document export in html format with Summary and InfoRating context analysis

EXPORT 2014-07-27 14:32:42

Title	Date	dentist	earnings	funds	insurance
Results of Social Democrat governors poor - Czech PM Necas	2012-08-28	0	40	42	
Beds reduction, computerisation best Czech health care steps	2012-11-28	0	0	40	
Microsoft has best Czech health care computerisation proposal	2012-11-21	0	0	40	
Czech doctors' salaries, insurance payments not to rise in 2013	2012-11-01	43	48	0	100

List of selected documents with chosen fields

Results: 4  
Context Analysis



#### Beds reduction, computerisation best Czech health care steps

Title: Beds reduction, computerisation best Czech health care steps  
 Date: 2012-11-28  
 dentist: 0  
 earnings: 0  
 funds: 40  
 insurance: 100  
 Entity person: vich michael

Category: Daily News  
 Date: 28.11.2012 08:00  
 Keywords: Czech;health;social;economy  
 Title: Beds reduction, computerisation best Czech health care steps  
 Language: en  
 Key: T201211270268601  
 Text:

Beds reduction, computerisation best Czech health care steps

Prague, Nov 27 (CTK) - Experts consider the abolition of redundant acute beds in hospitals and the proposed computerisation the most contributing measures to boost health care in the Czech Republic, Czech Health Care Forum (CZF) NGO executive head Michael Vich told reporters today.

Document texts with highlighted found words

Smart Monitoring enables documents export in format in which the documents are thematically divided into categories. It is possible to switch between these categories.



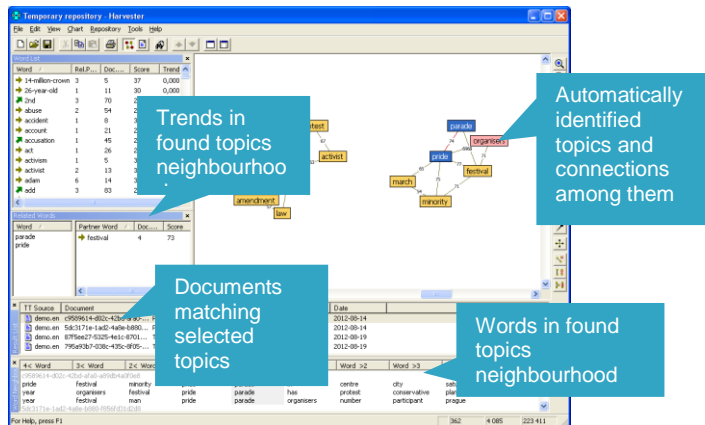
### Harvester

Harvester enables users to perform content analysis of unstructured texts especially in order to:

- ▶ determine topics automatically in analysed documents
- ▶ select documents describing similar topics
- ▶ identify new topics in document files

Topics are identified according to statistical analysis of appropriate word pair and trio occurrences. Varieties of them can be visualized.

Content analysis automatically identifies new topics from unstructured texts



Trends in found topics neighbourhood

Automatically identified topics and connections among them

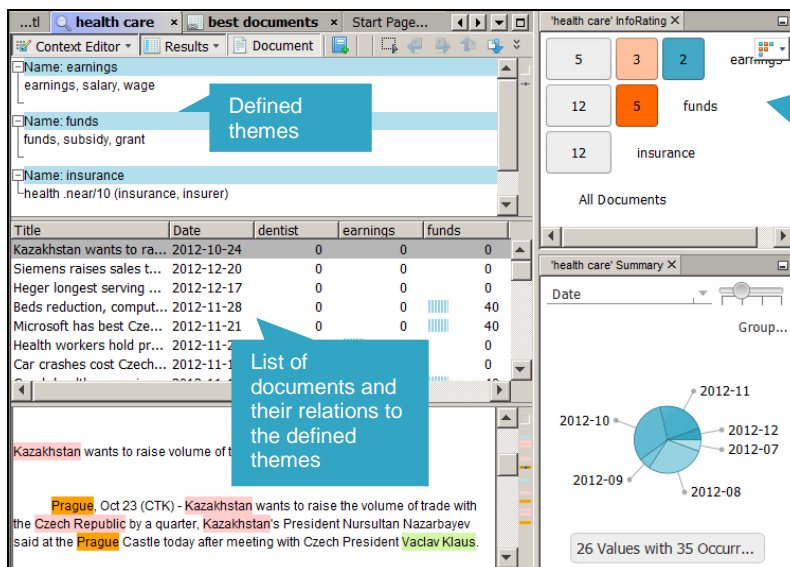
Documents matching selected topics

Words in found topics neighbourhood

### InfoRating

InfoRating provides context analysis of retrieved documents which finds and displays relationships between document content and defined themes - queries. The result can be displayed in the form of a well-arranged cross matrix or a graph. Users can also use the InfoRating to identify relationships among defined themes or to create structured reports with cross references.

InfoRating context analysis shows relationships between retrieved document contents and defined themes



Defined themes

Cross matrix showing relationships between retrieved document contents and defined themes

List of documents and their relations to the defined themes



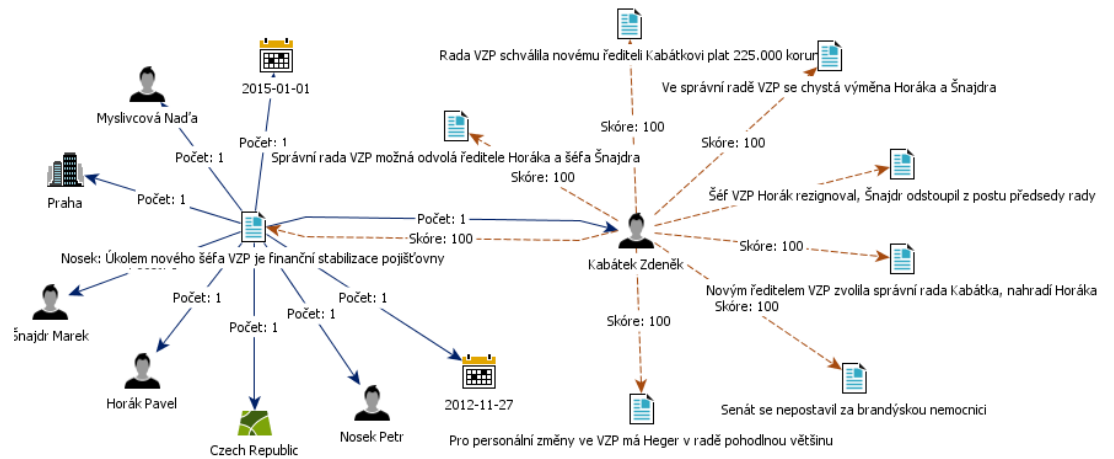
The screenshot displays the Tovek Tools interface with several components:

- Top Left:** A window titled "health care' Summary X" showing a pie chart of "Persons extracted for a topic" with labels: srsen vlastimil, vich michael, voleman milos, engel martin, klirova marie, and sohlich eduard.
- Top Right:** A window titled "health care' Summary X" showing a bar chart "Group by Month" with categories: earnings, funds, and insurance. The x-axis shows months from 2012-07 to 2012-12.
- Bottom Left:** A list of "Documents matching selected words" with columns for titles and counts for "lékař", "lůžka", and "sestra".
- Bottom Center:** A network graph titled "Chart of topic and document group relationships" showing nodes for "lůžka", "lékař", "sestra", and "úhradová vyhláška" with connecting lines and numerical values.
- Bottom Right:** A window titled "Shrnutí pro 'zdravotnictví' X" showing a pie chart with dates 2012-09-25, 2012-11-21, and 2012-07-12.

### Visualisation

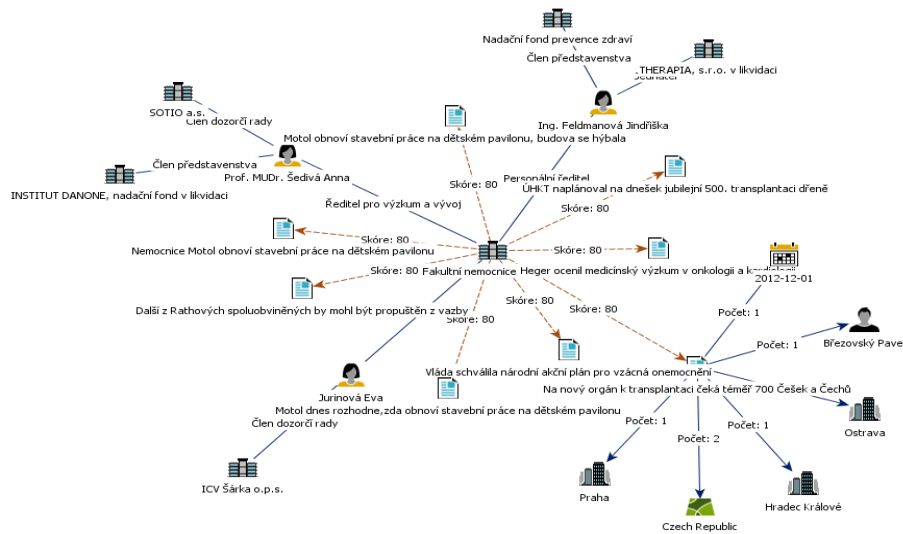
Based on entity extraction it's possible to visualise content of selected documents in a chart form. It also enables creation of special structured sources where objects and relationships between them are defined directly. Both ways can be combined. Created charts can be modified and also exported according to the user needs.

### Visualisation of found documents





Visualisation of structured source in combination with documents



It is also possible to find shortest path between visualised objects which can come from different sources. What's more, it enables specifying type of the link between the objects in order to find shortest and most relevant path.

### Fulltext Plug-in for IBM i2 Analyst's Notebook

Using the module Fulltext Plug-in for IBM i2 Analyst's Notebook users can visualize and analyse textual information in IBM i2 Analyst's Notebook - the visual analysis software. Fulltext Plug-in offers active connection between IBM i2 Analyst's Notebook chart items and information stored in text files. It also enables to use fulltext search methods within the IBM i2 Analyst's Notebook environment.

Connections between chart items in IBM i2 Analyst's Notebook and information stored in fulltext files

The screenshot shows the IBM i2 Analyst's Notebook interface. On the left, a network diagram displays nodes like 'Tatra', 'VOLKSWAGEN', and 'SKODA AUTO' connected by lines with numerical weights (e.g., 0.49, 0.83, 0.79, 0.43). Callouts point to 'Links to other subjects', 'Related documents', and 'Analysed subject'. On the right, a document preview window shows text with highlighted words and extracted entities, with a callout 'Document text with highlighted searched words and extracted entities'.

### Contact

TOVEK, spol. s r.o.  
Chrudimska 1418/2, 130 00 Prague 3, Czech Republic  
Tel.: +420 222 728 129, Fax: +420 222 728 103  
E-mail: [info@tovek.com](mailto:info@tovek.com), Web: [www.tovek.com](http://www.tovek.com)