TWELFTH INTERNATIONAL COMMAND AND CONTROL RESEARCH AND TECHNOLOGY SYMPOSIUM

ADAPTING C2 TO THE 21ST CENTURY

PAPER ID: I-015

TOPIC: NETWORK-CENTRIC EXPERIMENTATION AND APPLICATIONS

# COTS SOFTWARE FOR THE NET-CENTRIC C2 DECISION SUPPORT AND KNOWLEDGE MANAGEMENT

**PROF. LADISLAV BURITA, VOJTECH ONDRYHAL**

POINT OF CONTACT, AUTHOR: **PROF. LADISLAV BURITA**

UNIVERSITY OF DEFENCE, SENIOR LECTURER-PROFESSOR

65, KOUNICOVA STREET, K-209, 612 00 BRNO, CZECH REPUBLIC

TELEPHONE: +420 973 44 2172 / FAX: +420 973 44 2987

E-MAIL: LADISLAV.BURITA@UNOB.CZ

AUTHOR: **VOJTECH ONDRYHAL**

UNIVERSITY OF DEFENCE, SENIOR LECTURER

65, KOUNICOVA STREET, K-209, 612 00 BRNO, CZECH REPUBLIC

TELEPHONE: +420 973 44 2625 / FAX: +420 973 44 2987

E-MAIL: VOJTECH.ONDRYHAL@UNOB.CZ

# COTS SOFTWARE FOR THE NET-CENTRIC C2 DECISION SUPPORT AND KNOWLEDGE MANAGEMENT

**PROF. LADISLAV BURITA, VOJTECH ONDRYHAL**

## *Abstract*

Decision support systems (DSS) are a class of computer-based information systems including knowledge-based systems that support decision-making activities. The aim of the project in DSS area is finding answers for questions in military personnel career beginning with study at the University of Defence (UoD), followed by military service and finished by civilian positions. Clementine is a data mining software that enables you to quickly develop predictive models using business expertise and deploy them into business operations to improve decision making.

The Semantic Web is a second project that intends to create a universal medium for information exchange by putting data and documents with computer process able meaning on the World Wide Web. Ontology is a specification that we can present in the form of vocabulary, taxonomy, a thesaurus, a conceptual framework, or a theory. Information integration and sharing of information has become one of the key pillars of Network Enabled Capability (NEC).

Research project using commercial software Intelligent Topic Manager that is based on ontology is running at the CIS Department University of Defence. One of project goals is developing the prototype „Information System in the State Security“, implement and verify them in the Institute of Strategic Studies UoD.

## *Keywords*

C2 system, decision support, knowledge management, CIS, COTS SW (SPSS Clementine, MONDECA Intelligent Topic Manager, and TOVEK Tools Analysts Pack), NEC, research, experimentation

## *1 C2 decision support, knowledge management and software applications*

Decision support (DS) and knowledge management (KM) in net-centric command and control (C2) environment present an enormous and complex matter usually including many steps, methods, techniques and tools. It is hardly to expect a unique SW solution for that large subject. This area is dynamic and an application “for ever” we can hardly achieve.

Some SW concepts are very promising, for example service oriented architecture (SOA), but the required services must be well prepared, tested and implemented. In the authors’ opinion, best solution is to find out useful services provided by COTS SW and implement them for C2 support.

In the article, we would like to publish our experience with selected commercial software for decision support, business intelligence and knowledge management. We would like to show how outputs from developed prototypes in administrative and security areas could be applied into command control area in connection with NEC. Our experiments are connected up to coherence transformational maturity level in section Information Integration Services (Information Integration, Semantics Web, SOA, and Interoperability).

## 1.1 Decision support

Decision support systems (DSS) are a class of computer-based information systems including knowledge-based systems that support decision-making activities. DSS belong to an environment with multidisciplinary foundations, including database research, artificial intelligence, human-computer interaction, simulation methods and software engineering.

According to [5], we can classify DSS into the following groups:

- A **model-driven DSS** emphasizes access to and manipulation of a statistical, financial, optimization, or simulation model. Model-driven DSS use data and parameters provided by users to assist decision makers in analyzing a situation; they are not necessarily data intensive.
- A **communication-driven DSS** supports more than one person working on a shared task.
- A **data-driven DSS** or data-oriented DSS emphasizes access to and manipulation of a time series of internal company data and, sometimes, external data.
- A **document-driven DSS** manages, retrieves and manipulates unstructured information in a variety of electronic formats.
- A **knowledge-driven DSS** provides specialized problem solving expertise stored as facts, rules, procedures, or in similar structures.

To the group of DSS we usually include executive information systems (EIS), group decision support systems (GDSS), organizational decision support systems (ODSS), data warehousing and on-line analytical processing (OLAP) and data mining (DM).

## 1.2 Semantic web

Semantic Web presents data and documents that are published on the web. There are not only simple data but also data with the relevant context, with the information about data in machine-readable form. This technology brings for the web a new dimension, by that the data have a sense for computers. It makes possible integration of data from various sources and more significant of the data (automatic) processing. The main condition for the semantic web is production of relevant standards, vocabularies and ontologies.

The idea of Semantic Web comes out from the need to give to the web distinct meaning and make the accessible data understandable for machines. Machines understand data using descriptions of information that are stored on the web specified thru vocabulary and ontology definitions. We consider ontology in this context as a concept specification; it means word categorization in broad context; for example explaining meaning of a word in language dictionary. Web documents and other internet sources are connected with information that for computer represents deduction basis of relations among those information resources.

The very important aspect of the Semantic Web is preparing of appropriate ontologies. This process requires effort of many various communities; for example from health care, assurance or financial and of course it is a problem of military.

Many existing technologies have tried to deal with semantics (Controlled vocabularies, Thesaurus, Index, Classifications Schemes and Taxonomies, Formal Ontologies and Schemas, UML, Data Bases, Metadata). See at the Figure 1 that there is a logical pyramid of constructs and tools in information modelling and knowledge bases. Their short reflection:

- XML (Extended Markup Language): markup language, universal document syntaxes.
- URI (Uniform Resource Identifier): unified information source identification.

- Dublin-Core: Metadata initiative includes set of metadata elements.
- WSDL (Web Service Definition Language): language for web services definition.
- UDDI (Universal Description, Discovery and Integration): registry of web services.
- RDF (Resource Description Framework): information sources description, mechanisms for metadata recording.
- XTM (XML Topic Maps): representation of the information sources structure.
- SKOS (Simple Knowledge Organisation Systems): knowledge systems representation using RDF.

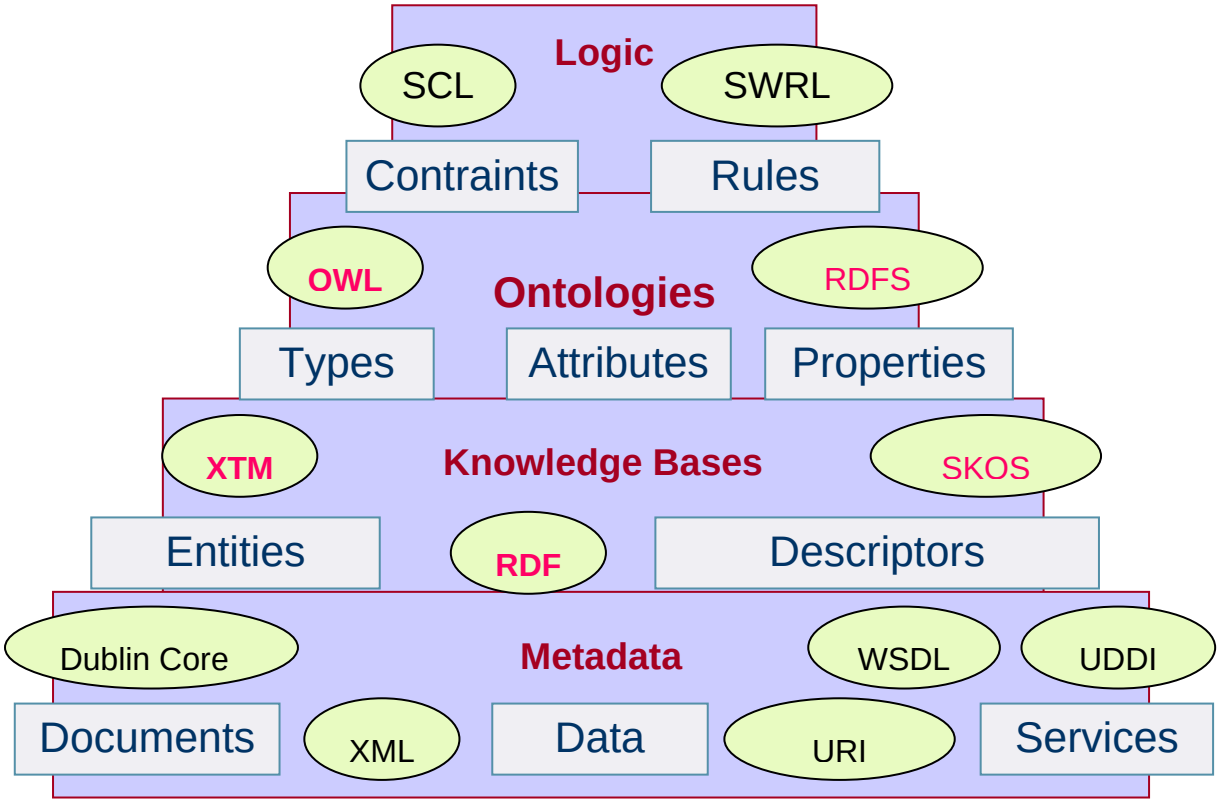



*Figure 1 Constructs and tools pyramid in information modelling and knowledge bases*

## *2 SW possibilities and expectations; directions to the solution*

Managers' expectations of COTS software deployment are usually very high. They may assume that purchased software will easily solve their issues and troubles in an organization. Software is often mistaken for method or process that should exist independently. We present roles participating on software deployment and we discuss possible models of implementation in this part.

### 2.1 User roles

**End users** are users that use software to fulfil elementary repeatable business processes. End users are ordinarily largest group. Application user interface should be very simple, resistant to user faults. Every unnecessary operation or step implemented takes time. Processes cover simple input and update operations, prepared reports and prints.

**Business users** require more sophisticated user interface with advanced options. They work usually on unique and more complex business processes. In some software applications, like Clementine, business user can customize simple business processes for end users.

**IT Users** are specialists that understand technology used for software creation and are familiar with deployment model of current installation, but they are usually not skilled with business processes. The responsibilities of this user group are:

- Data backup and if required data recovery.
- Installations for new end and business users.
- Security model maintenance. It includes mainly user definition and roles assignment.

## 2.2 COTS architecture

Typical COTS architecture consists of four levels, see at the Figure 2. User interfaces, business layer (company knowledge base) including settings and customizations, COTS itself and independent data source. Business customizations and data sources are not part of purchased software; it is companies' responsibility sort out these two layers, even before software is implemented. Business customizations reflects existing or presumptive processes in the organization, usually includes workflow possibilities. Data sources exist in a form of separated databases that we should consolidate.

As was mentioned above, end user interface should be simple. It is wise integrate such interface into existing office applications (e.g. customized email messages, spreadsheet and text files). The integration with office applications lowers expenses for user training; email client or spreadsheet usage is basic skill of nearly every user.

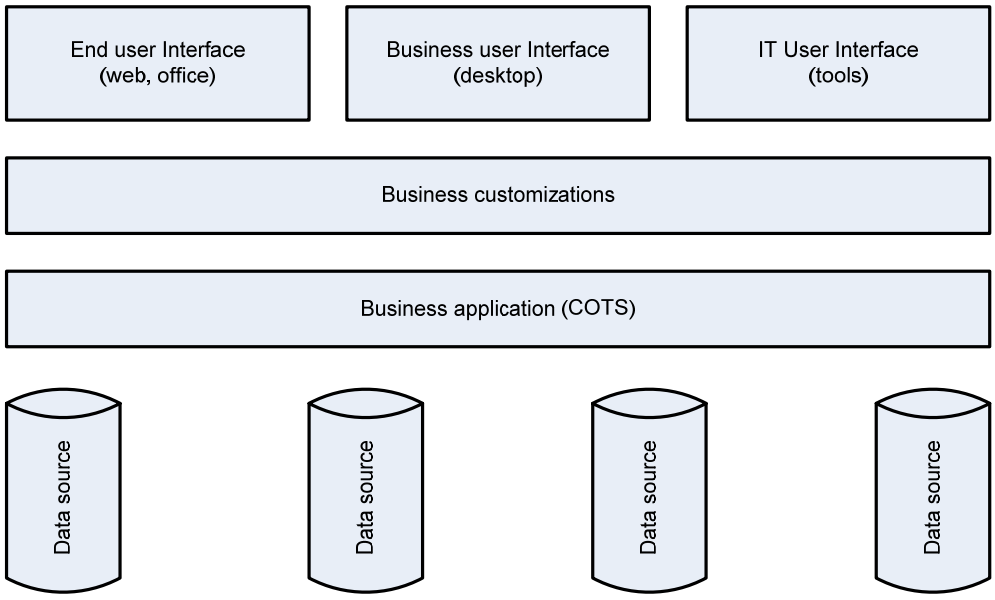



*Figure 2 COTS Architecture*

# *3 Research projects as a background for the experimentation*

The goal for researchers from (not only) University of Defence is to investigate possible technologies and approaches in partial areas. We would like to discuss if it's possible to transfer knowledge from ongoing projects to NEC area of interests (We think so).

Researchers' experience is based on two research projects which has been solved at the University of Defence; "Planning processes in the MoD" [1], and "Communication and information systems (CIS) development and integration in the NATO environment" [2].

First one examines Data Mining Clementine suite [4a] for the decision support in the planning processes and

Experiment one – administrative area (Data Mining) includes steps:

- Data understanding.
- Data preparation.
- Model construction.
- Model validation.
- Deployment.

The second one experiments with ITM (Intelligent Topic Manager) [4b] for the IS development with the effective information retrieval and

Experiment two – state security (Semantic Web) includes steps:

- Thesaurus development.
- Transformation into ontology.
- Ontology validation.
- Document base creation.
- Annotation – automated.
- Knowledge management system development.

## *4 Project in Decision Support area*

The aim of the project is finding answers for questions in military personnel career beginning with UoD study (area, results), followed by military service and finished by civilian positions. There are two information systems (IS) in kind of personnel information available. IS UoD (former MA – Military Academy) includes the study areas and the student results and Personnel IS (PIS) that holds all steps in military service. The third information resource is an external input by Internet or other IS or people knowledge, see Figure 3.

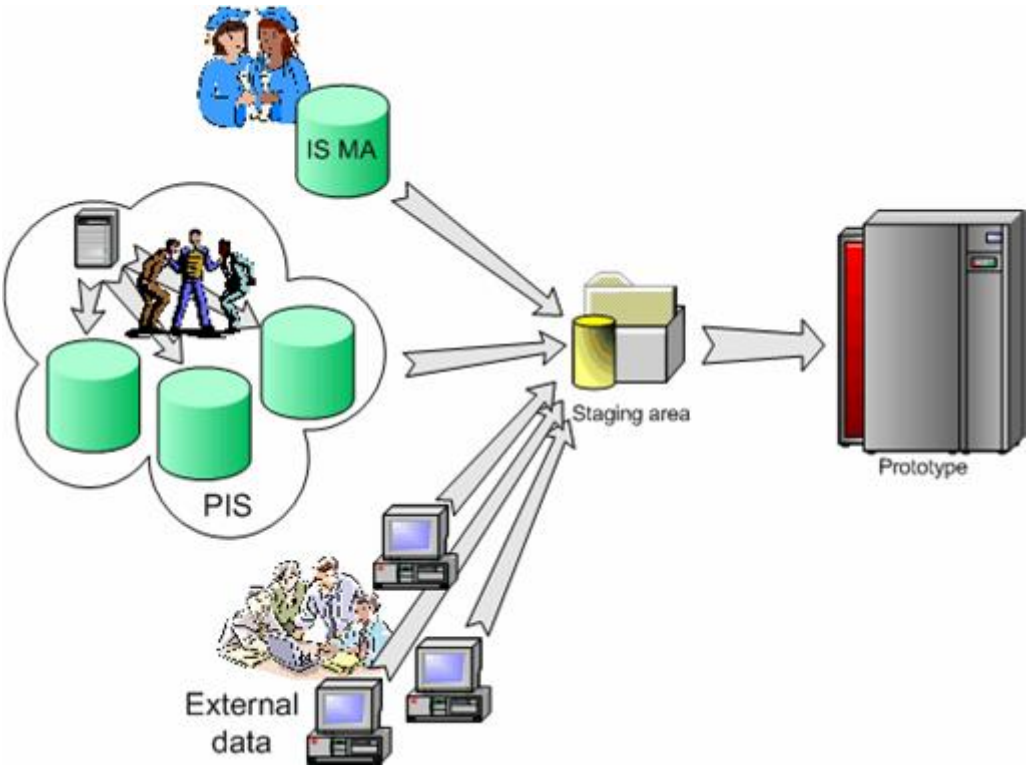



*Figure 3 Personnel project data source architecture*

## 4.1 COTS Clementine

Clementine is a data mining software that enables you to quickly develop predictive models using business expertise and deploy them into business operations to improve decision making. Designed around the industry-standard CRISP-DM model, Clementine supports the entire data mining process, from data to better business results. Clementine helps you discover and predict interesting and valuable relationships within your data. You can use Clementine for decision-support activities. If you have the data and your data contain the right information, Clementine will help you find answers to your questions.

The important features are:

- Automatic version control.
- User authorization and authentication.
- Powerful search tools for locating models.
- Automated data mining processes, including model building, refreshing, and scoring.
- Job control and e-mail notification.

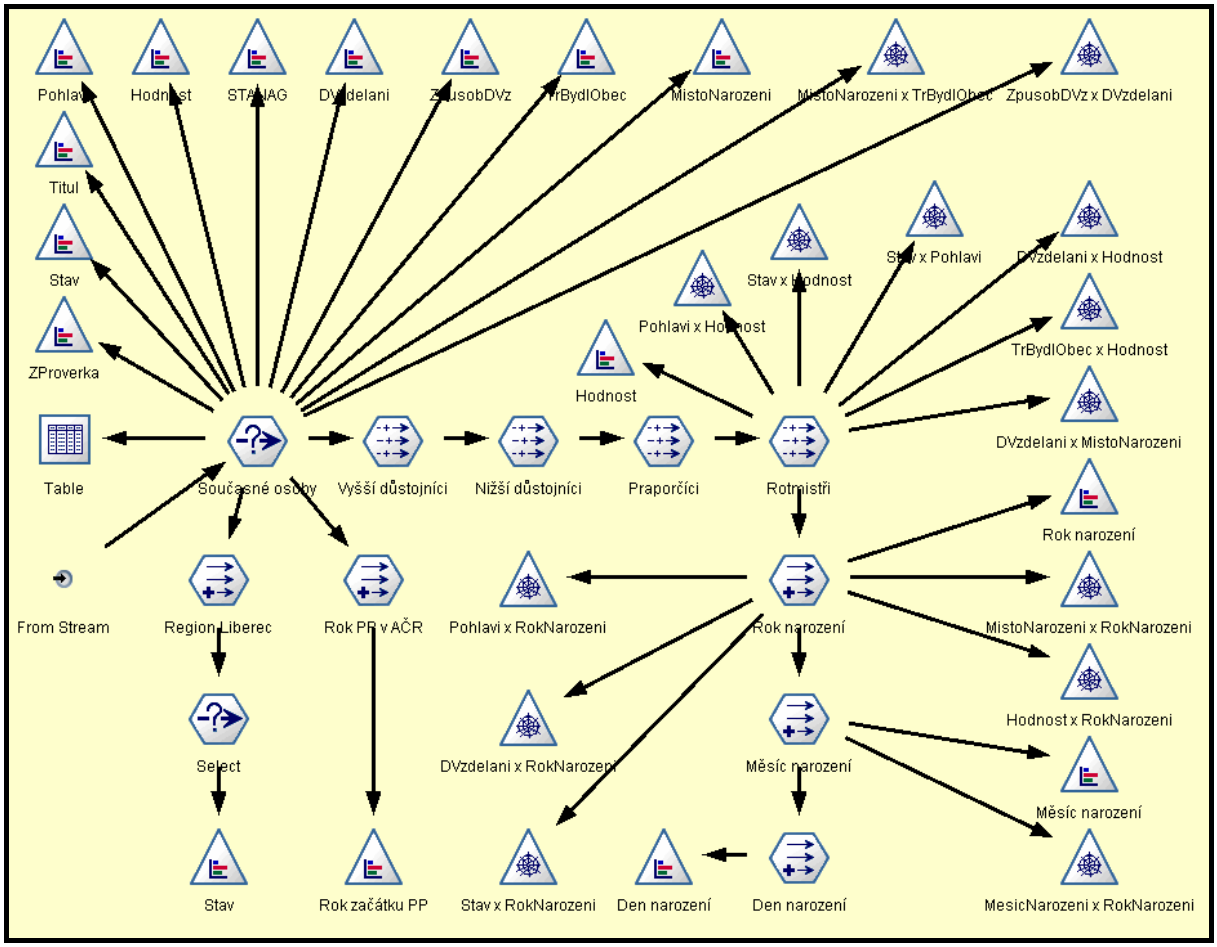



*Figure 4 An example of business user customization made in Clementine*

## 4.2 Solution overview

We transformed all available data sources into enterprise datawarehouse for ongoing analysis. We used Clementine provided by SPSS Company as a tool for required analysis. Clementine is an example of COTS software that requires educated business users. End users cannot use such type of software immediately. Settings and customizations made by business users are required. Business users create simple templates for business operations in form of streams in

Clementine. End users can use results of these streams like reports and plots later. On the Figure 4 there is an example of such stream. This stream analyzes data about persons stored in systems. Stream provides many outputs in table forms (square icon), and plots, tables and webs (triangle icons), see Figure 5.

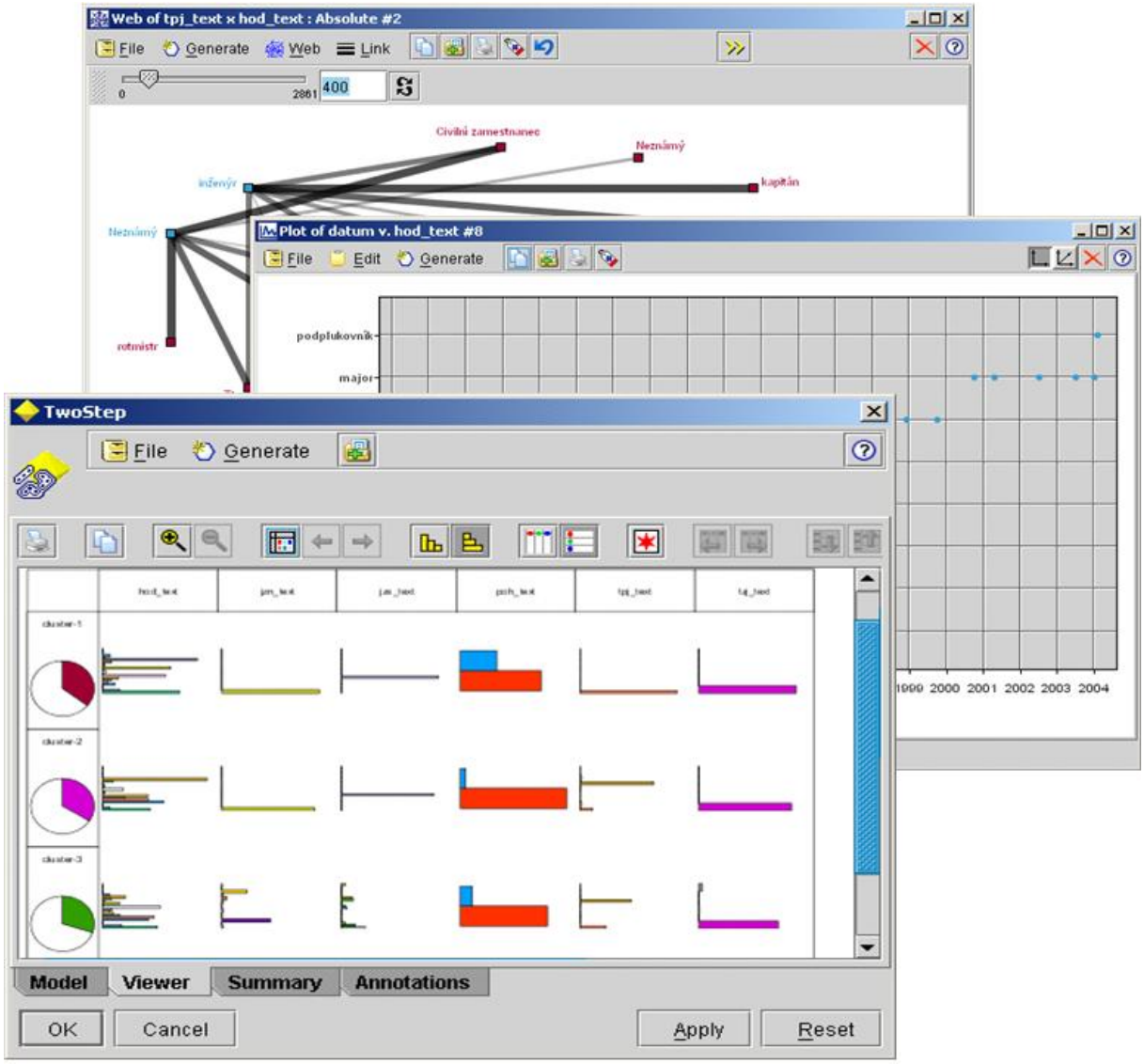



*Figure 5: Output examples for end user*

## *5 Project of Information System in the State Security*

International project of Information System in the State Security area was started at the end of the year 2006. Project was prepared in cooperation with Institute for Strategic Studies (ISS) UoD and companies Mondeca (France) [4b] and Tovek (Czech Republic) [4c]. Project is based on application of commerce software Intelligent Topic Manager (ITM) made by the company Mondeca for the intelligent data organisation and retrieval.

### 5.1 Intelligent Topic Manager

ITM is a unique tool that federates and organizes information and knowledge in a business-specific reference repository for navigation that is more effective and searches.
ITM functionality (see Figure 6) includes:

- Ontology management, thesaurus, taxonomies, knowledge bases.

- Navigation in a business-related representation.
- Multi-criteria searches in bases and content.
- Automatic content annotation and knowledge acquisition.
- Collaborative work to capitalize on knowledge.
- Reuse of content for composition, publishing and distribution.

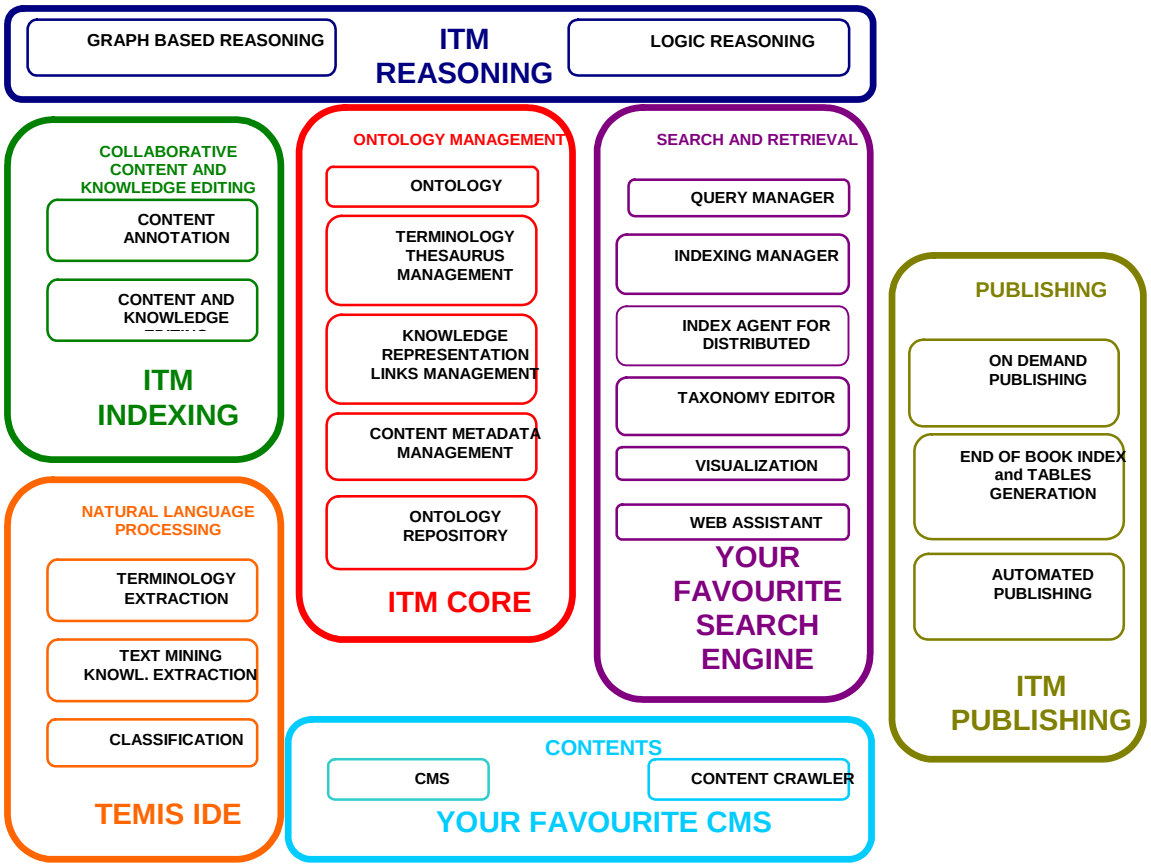



*Figure 6 The ITM functionality with context of other tools*

Companies and administrations face with the challenge of managing processes in which the handling of data, information and knowledge are tightly intertwined. Content Intelligence, Semantic Web and Ontologies are emerging solutions capable of organizing and federating information that is spread throughout the enterprise and making it available to employees, call centres, partners, customers, prospects and citizens.

## 5.2 ITM meta-model and structure

Where ITM is presented as an ontology-driven application and its meta-model can be expressed using OWL to formally define both its topic-map-like meta-classes, and specific data types. An ontology and knowledge base using ITM meta-model can be in both its RDF and XTM representations.

**Simple meta-model**

- A few generic types: Topic, Association, Role, Data Item; defined by specific meta-classes.
- The meta-model can be expressed in RDF-OWL, but also, in a less formal way, using Topic Maps concept.

- The same basic objects are used at every level in a highly recursive way.

**Topics, Associations and Roles**

- ITM semantic network uses a “hyper graph” structure.
- Hyper graph “nodes”: Topics (= vertices) and Associations (= edges), a node is either a Topic or an Association.
- Hyper graph “connectors”: Roles (= incidences). A Role links exactly one Topic to exactly one Association.
- Data Items can be attached to Topics, Associations and Roles and have a logical type, and a physical type.

**Classes, meta-classes and workspaces**

- Every ITM object has at least one declared class (or type).
- Associations, Roles and Data Items have exactly one type.
- Topics have at least one declared class (plus inherited one).
- Recursive meta-model: Classes and types are declared as topics, including meta-classes.
- Classes and instances are not defined in the same workspaces:
    - Basic Ontology workspace defines ITM built-in meta-classes, accessible only by “root administrator”, and released with the system.
    - Client Ontology workspace defines client classes.

## 5.3 Project goals, specification, and results

Current state of the information processing in the ISS could be specified as decentralized and individual. The information obtained and created in the ISS is currently saved in the PC of individual worker. The information is in the form of studies, articles, proceedings, presentations, academic documents and photos.

They come from the Czech Republic and from international sources. The document formats are .jpeg, .gif, .doc, .rtf, .xls, .ppt, .pdf. Information subject classification is consistent with subject of individual group of ISS (security studies, warfare group, and resources - processes).

We suggested for the technical base open software solution (RDMS PostgreSQL and application server JBoss) to achieve compatibility with SW ITM. Final state of the information processing in the ISS should be centralized and integrated. Save consolidated information in accordance to subject of ISS group, central management and integration, intelligent searching.

Project goals:

1. To develop the Prototype „Information System in the State Security“, implement and verify them in the ISS environment.
2. To start research of the Semantic WEB, Ontology, etc. for the C2IS interoperability in NATO environment and to become familiar with those technologies.

The prototype should allow conceptual searching, annotation creating, collaborating on knowledge, subject publishing according to selected criteria, exploitation of ontology and taxonomy.

Project Phases:

- Preparation phase, education in knowledge management, ontology, ITM etc.
- Installation of DBMS PostrgreSQL, AS JBoss, SW ITM.
- Ontology research and preparation.
- Prototype building, implementation and verification.

- Results demonstration and evaluation phase.

Method of thesauri design:

- Preparation of typical ISS document base.
- Thematic vocabulary ad-hoc specification (classes are in the Table 1).
- Analyse of document base (text mining, harvesting), see Table 1, Figure 7.
- Thematic vocabulary corrections.
- Thesauri definition.

Future work:

- Ontology definition.
- Automatic annotation.
- Information retrieval from Internet sources using thesauri and ontology.

| NUM | CLASS | Rel-0,90 | Rel-0,75 | R%-0,90 | R%-0,75 |
|---|---|---|---|---|---|
| 1 | ARMAMENT | 19 | 790 | 0,16 | 6,50 |
| 2 | ARMY | 185 | 3717 | 1,52 | 30,61 |
| 3 | ART OF WAR | 0 | 30 | 0,00 | 0,25 |
| 4 | CAPABILITIES | 42 | 3058 | 0,35 | 25,18 |
| 5 | CONCEPTIONS | 10 | 1570 | 0,08 | 12,93 |
| 6 | CONFLICT | 40 | 1810 | 0,33 | 14,90 |
| 7 | DEFENCE | 256 | 4928 | 2,11 | 40,58 |
| 8 | DOCTRINE | 46 | 689 | 0,38 | 5,67 |
| 9 | ENVIRONMENT | 19 | 1935 | 0,16 | 15,93 |
| 10 | INTERERST | 4 | 572 | 0,03 | 4,71 |
| 11 | MODERNIZATION | 33 | 1318 | 0,27 | 10,85 |
| 12 | ORGANIZATION | 31 | 2735 | 0,26 | 22,52 |
| 13 | POLITICS | 76 | 2540 | 0,63 | 20,91 |
| 14 | POWER | 5 | 513 | 0,04 | 4,22 |
| 15 | RELATION | 18 | 2114 | 0,15 | 17,41 |
| 16 | RESOURCE | 35 | 2602 | 0,29 | 21,42 |
| 17 | RISK | 85 | 1479 | 0,70 | 12,18 |
| 18 | SECURITY | 42 | 2677 | 0,35 | 22,04 |
| 19 | SITUATION | 24 | 3478 | 0,20 | 28,64 |
| 20 | STRATEGY | 47 | 1377 | 0,39 | 11,34 |
| 21 | TERRORISM | 7 | 464 | 0,06 | 3,82 |
| 22 | THREAT | 29 | 1185 | 0,24 | 9,76 |

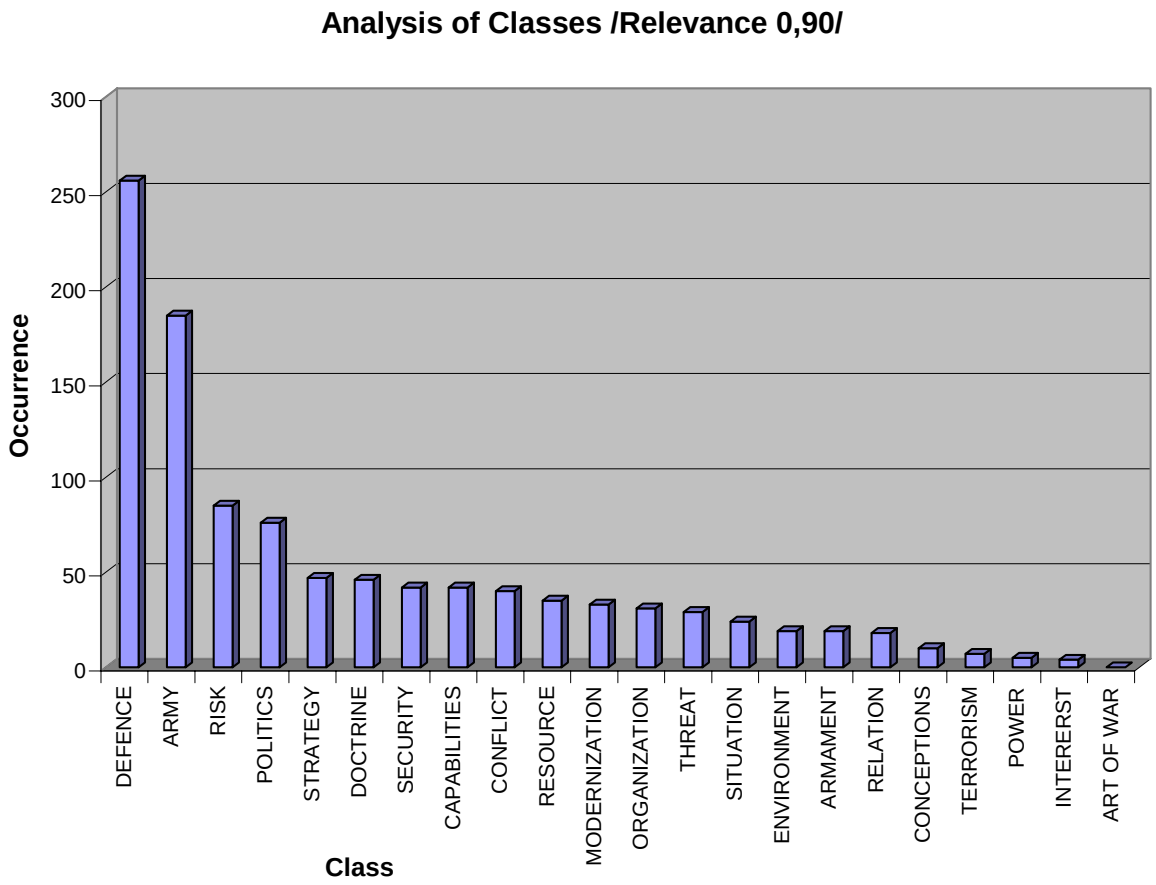
*Table 1 Ontology classes*

Project has started at the end of the year 2006 and we have begun some experimentation attempts. The results will be available at the end the year 2007, but we would like to show the first experiments in the ontology building in the paper.

Analysis of ontology classes was realized using current information sources of the ISS that involve 12 145 documents. All suggested classes we retrieved upon the document based in SW Tovek Tools Analysts Pack [4c]. First track was made by relevance 0,9 of the occurrence class in the document and the second track was made by relevance 0,75. Results are in the Table 1 and Figure 7. Probably some classes are not relevant to the ontology becase they have minimum occurrences (Art of War), but it is the ISS users responsibility at the next step of analysis.

## *6 Lessons learned*

Difficult and complex data analysis can be smartly solved using Clementine. The tool transforms analysis tasks to simple data streams covering requests. Reporting and visualization of data mining analysis results is important step for model understanding. Clementine provides comprehensive set of reporting tools as well.



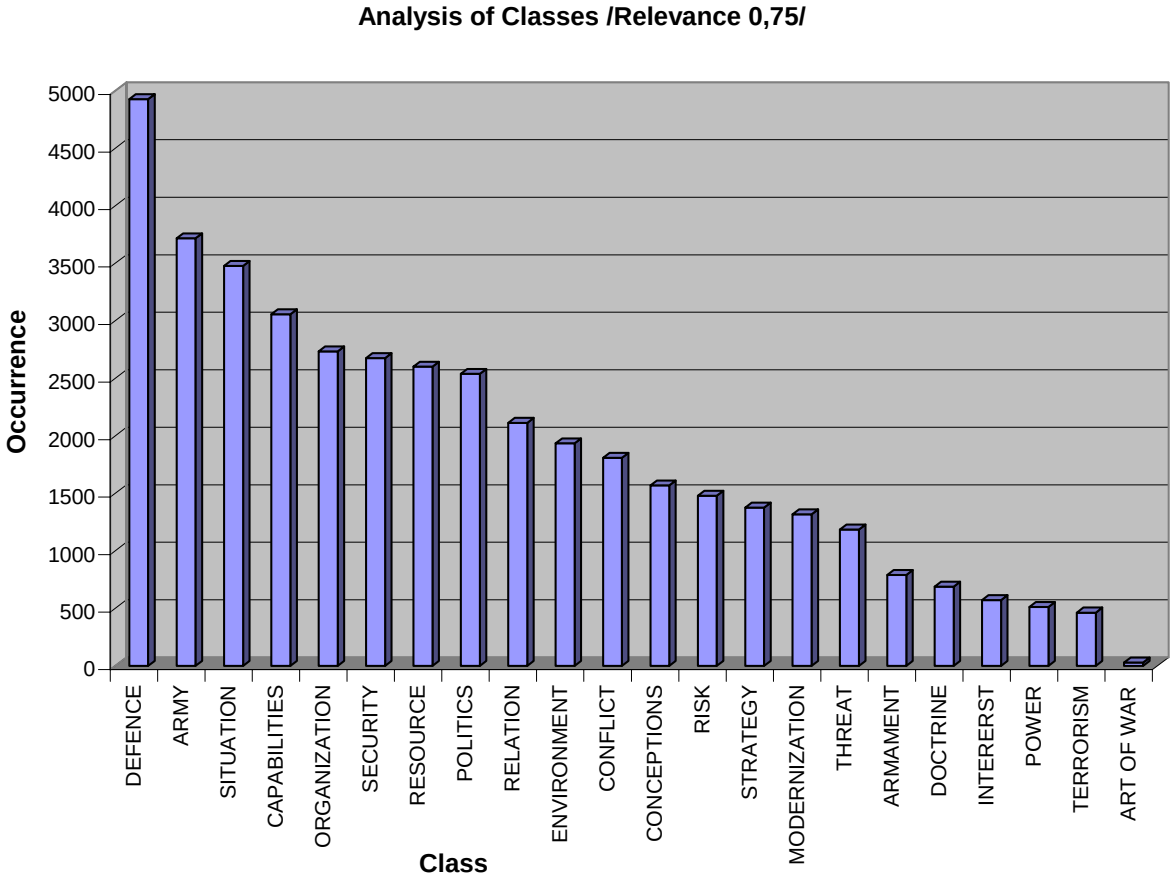




*Figure 7 Analysis of Ontology Classes*

New method for ontology preparation in the security area using analytics and text mining Tovek Tools SW was suggested and IS architecture by ITM was implemented.

Procedures, we took during our research, are transferable and exploitable in NEC C2 area. It is necessary to obtain required data, documents and present processes, and then in the same way apply appropriate COTS Software.

Domain ontology design is required for C2. Then we can exercise data sets and documents according to various situations automatically identified during commanders' decision-making process. We would like to contribute more in this field in the future.

## *References*


[1] Development of suggestion support methods in processes of planning and activity evaluation in MoD. (MoD research project). Brno: UoD, 2006-2007.

[2] Development, integration, administration and security of CIS in NATO environment. Research intent FVT0000403. Brno: UoD, 2004-2008.

[3] NATO Network Enabled Capability Feasibility Study, Version 2. Brussels: NC3A, 2005, 623 pp.

[4] [4a] www.spss.com; [4b] www.mondeca.com; [4c] www.tovek.cz

[5] Power, D. J. Decision support systems: concepts and resources for managers. Westport: Quorum Books, 2002, 251 pp. ISBN 156720497X.