EuroTermBank

Collection, consolidation and dissemination of multilingual terminology resources through cooperation of terminology institutions

Andrejs Vasiljevs
Tilde, Latvia

eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

Terminology resources in new EU countries

- Many institutions involved in terminology development with insufficient cooperation
- Subject fields fragmented in different publications, data collections and files
- Different structure, incompleteness, inconsistency
- Different approaches in terminology development
- Large variety of incompatible storage formats (inc. many resources on paper only)
- Majority of resources are bilingual

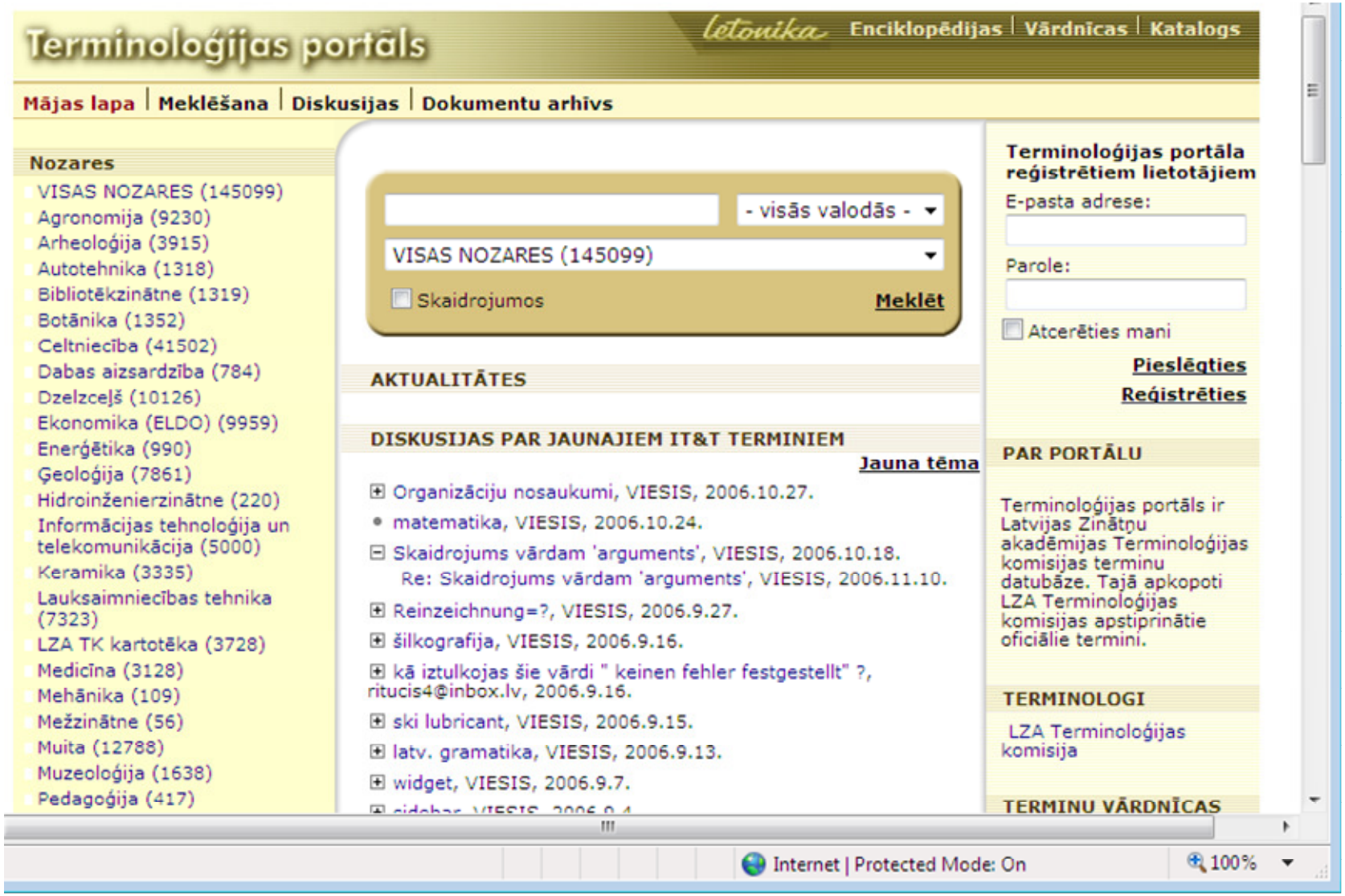
eContent

EuroTermBank

Examples of activities to consolidate national terminology

- termnet.lv - Terminology portal of Terminology Commission of Latvia
- Terminology portal of Lithuanian State Language Commission

eContent



termnet.lv - Terminology portal of Terminology Commission of Latvia

Terminology portal of Lithuanian State Language Commission

EuroTermBank

EuroTermBank goal is to collect, harmonise and disseminate terminology resources in new EU member states, transferring experience from other EU terminology projects and accumulating competencies and efforts of the accessed countries.

eContent

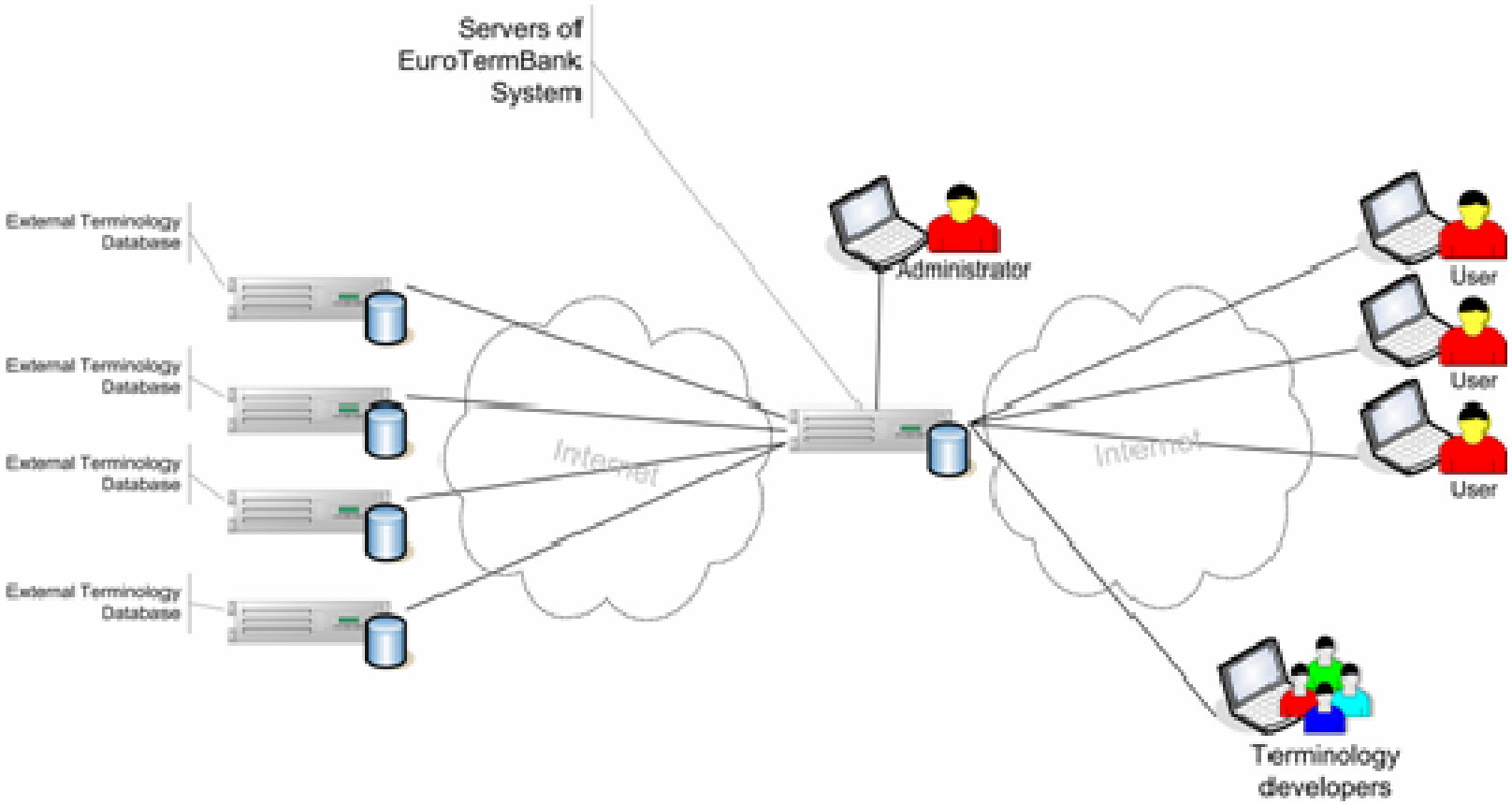
EuroTermBank

Focus areas of the EuroTermBank

- Development of methodology recommendations for new EU member states, aligned with the best practices in EU
- Establishment of network of national and international terminology institutions
- Consolidation, harmonization and centralization of fragmented terminology resources
- Creation of online terminology data bank with links to various external term databases and banks

eContent

EuroTermBank – Single Access Point to Multilingual Terminology



EuroTermBank

European Union eContent Programme Project

- Part of EU eContent programme aimed to stimulate the development and use of European digital content on the global networks and to promote the linguistic diversity in the information society
- Project partners are from Germany, Denmark, Estonia, Latvia, Lithuania, Hungary and Poland
- Partners comprise universities, private companies and institutions of state coordinated/regulated terminology

eContent

EuroTermBank Participants & Roles

- Center for Language Technologies, University of Copenhagen – Denmark
  - Assessment of the best practice
  - Development of methodology recommendations
- University of Applied Sciences Cologne – Germany
  - Development of organisational and legal frameworks
  - Networking of terminology institutions
- The Institute of the Lithuanian Language) – Lithuania
  - Collection of national resources
  - Awareness, dissemination, marketing and exploitation
- Tilde – Latvia
  - Coordinator
  - Project management
  - System development and implementation

eContent

EuroTermBank Participants & Roles

- Latvian Academy of Sciences – Latvia
  - General content selection, acquisition and processing
  - Collection of national resources
- MorphoLogic – Hungary
  - Requirements analysis and system design
  - Collection of national resources
- Department of Estonian and Finno-Ugric Linguistics, University of Tartu – Estonia
  - Assessment, evaluation and system elaboration
  - Collection of national resource
- Information Processing Centre – Poland
  - Collection of national resource
  - Awareness and dissemination on national level

eContent

EuroTermBank

Users & Benefits for Users

- Targeted to both terminology users and terminology creators
- Single and free online access point to comprehensive terminology resources in different languages and fields
- Effective dissemination of terminology resources
- Involvement of general public in terminology development and dissemination through online discussions
- Facilitation of professional terminology development through authorized online database tools and discussions
- Sophisticated search mechanisms taking into account language specifics

eContent

EuroTermBank

Usage Scenarios

- Translation
- Technical writing
- Reference/research
- Terminology development, lexicography
- Online data source for dictionary software and translation tools

eContent

EuroTermBank

Terminology Work Recommendations

Provide information and recommendations for all aspects of terminology work based on *standards* and *best practice*

- compilation of terms, conceptual analysis, creation of term entries
- Architecture, design and data structure of term base

eContent

EuroTermBank

Terminology scenarios

- The composition of goals and conditions that typically co-exist suggested 3 terminology scenarios
  - Local
  - National
  - International
- These 3 scenarios are different in crucial respects and also share some goals and conditions

eContent

EuroTermBank

International Scenario – some Goals and Conditions

- Coordination
  - Approval/dismissal and harmonization of terms
- Multilingual terminology collection
- Usually well-organized infrastructure
- Broad domain coverage
- All types of language professionals represented
- Quality is high priority
- Exchangeability is high priority
- Availability is high priority

eContent

EuroTermBank

National Scenario – some Goals and Conditions

- Coordination and regulation
  - Approval/dismissal and harmonization of terms
- Mono- or bi-lingual terminology collection
- Usually well-organized infrastructure
- Broad domain coverage
- All types of language professionals represented
- Quality is high priority
- Exchangeability is high priority
- Availability is high priority

eContent

EuroTermBank

Local Scenario – some Goals and Conditions

- Translation
- Bi- or multi-lingual terminology collection
- Terminologists often not part of terminology team
- Focused domain coverage
- Tight time frames co-exist with – and put limitations on requirements for – high quality
- Exchangeability is high priority
- Availability is not a priority

EuroTermBank

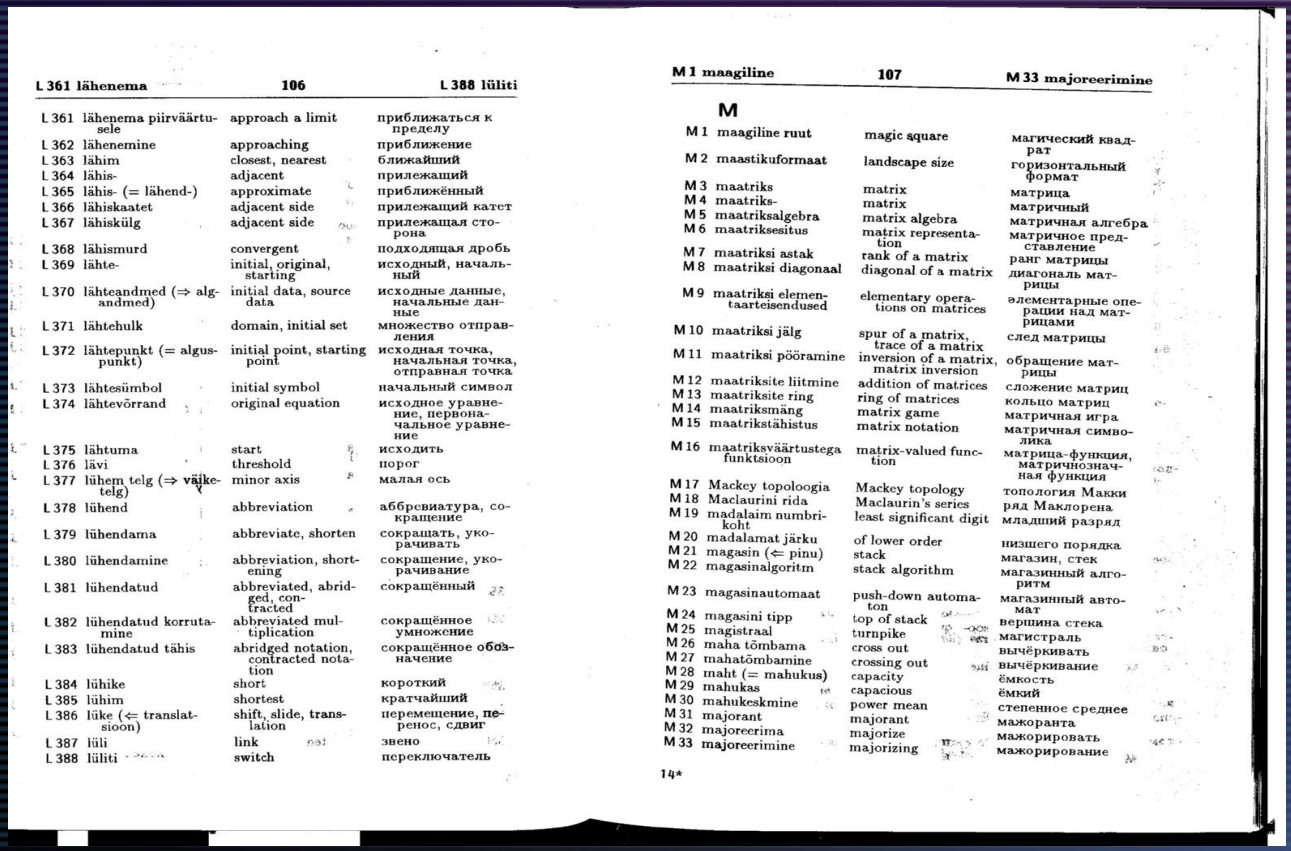
Consolidation of terminology resources

- Identification of existing terminology resources
- Evaluation and Selection of Resources
- Content acquisition (contracts with owners)
- Content digitalisation and processing
- Modification to harmonize and meet structural and technical requirements

eContent

EuroTermBank

EE-EN-RU Mathematical terms in printed form



L 361 lähenema — 106 — L 388 lüliti

L 361	lähenema piirväärtusele	approach a limit	приближаться к пределу
L 362	lähenemine	approaching	приближение
L 363	lähim	closest, nearest	ближайший
L 364	lähis-	adjacent	прилежащий
L 365	lähis- (= lähend-)	approximate	приближённый
L 366	lähiskaatet	adjacent side	прилежащий катет
L 367	lähiskülg	adjacent side	прилежащая сторона
L 368	lähismurd	convergent	подходящая дробь
L 369	lähte-	initial, original, starting	исходный, начальный
L 370	lähteandmed (⇒ algandmed)	initial data, source data	исходные данные, начальные данные
L 371	lähtehulk	domain, initial set	множество отправления
L 372	lähtepunkt (= alguspunkt)	initial point, starting point	исходная точка, начальная точка, отправная точка
L 373	lähtesümbol	initial symbol	начальный символ
L 374	lähtevõrrand	original equation	исходное уравнение, первоначальное уравнение
L 375	lähtuma	start	исходить
L 376	lävi	threshold	порог
L 377	lühem telg (⇒ väike-telg)	minor axis	малая ось
L 378	lühend	abbreviation	аббревиатура, сокращение
L 379	lühendama	abbreviate, shorten	сокращать, укорачивать
L 380	lühendamine	abbreviation, shortening	сокращение, укорачивание
L 381	lühendatud	abbreviated, abridged, contracted	сокращённый
L 382	lühendatud korrutamine	abbreviated multiplication	сокращённое умножение
L 383	lühendatud tähis	abridged notation, contracted notation	сокращённое обозначение
L 384	lühike	short	короткий
L 385	lühim	shortest	кратчайший
L 386	lüke (⇐ translatsioon)	shift, slide, translation	перемещение, перенос, сдвиг
L 387	lüli	link	звено
L 388	lüliti	switch	переключатель

M 1 maagiline — 107 — M 33 majoreerimine

M

M 1	maagiline ruut	magic square	магический квадрат
M 2	maastikuformaat	landscape size	горизонтальный формат
M 3	maatriks	matrix	матрица
M 4	maatriks-	matrix	матричный
M 5	maatriksalgebra	matrix algebra	матричная алгебра
M 6	maatriksesitus	matrix representation	матричное представление
M 7	maatriksi astak	rank of a matrix	ранг матрицы
M 8	maatriksi diagonaal	diagonal of a matrix	диагональ матрицы
M 9	maatriksi elementaarteisendused	elementary operations on matrices	элементарные операции над матрицами
M 10	maatriksi jälg	spur of a matrix, trace of a matrix	след матрицы
M 11	maatriksi pööramine	inversion of a matrix, matrix inversion	обращение матрицы
M 12	maatriksite liitmine	addition of matrices	сложение матриц
M 13	maatriksite ring	ring of matrices	кольцо матриц
M 14	maatriksmäng	matrix game	матричная игра
M 15	maatrikstähistus	matrix notation	матричная символика
M 16	maatriksväärtustega funktsioon	matrix-valued function	матрица-функция, матричнозначная функция
M 17	Mackey topoloogia	Mackey topology	топология Макки
M 18	Maclaurini rida	Maclaurin's series	ряд Маклорена
M 19	madalaim numbrikoht	least significant digit	младший разряд
M 20	madalamat järku	of lower order	низшего порядка
M 21	magasin (⇐ pinu)	stack	магазин, стек
M 22	magasinalgoritm	stack algorithm	магазинный алгоритм
M 23	magasinautomaat	push-down automaton	магазинный автомат
M 24	magasini tipp	top of stack	вершина стека
M 25	magistraal	turnpike	магистраль
M 26	maha tõmbama	cross out	вычёркивать
M 27	mahatõmbamine	crossing out	вычёркивание
M 28	maht (= mahukus)	capacity	ёмкость
M 29	mahukas	capacious	ёмкий
M 30	mahukeskmine	power mean	степенное среднее
M 31	majorant	majorant	мажоранта
M 32	majoreerima	majorize	мажорировать
M 33	majoreerimine	majorizing	мажорирование

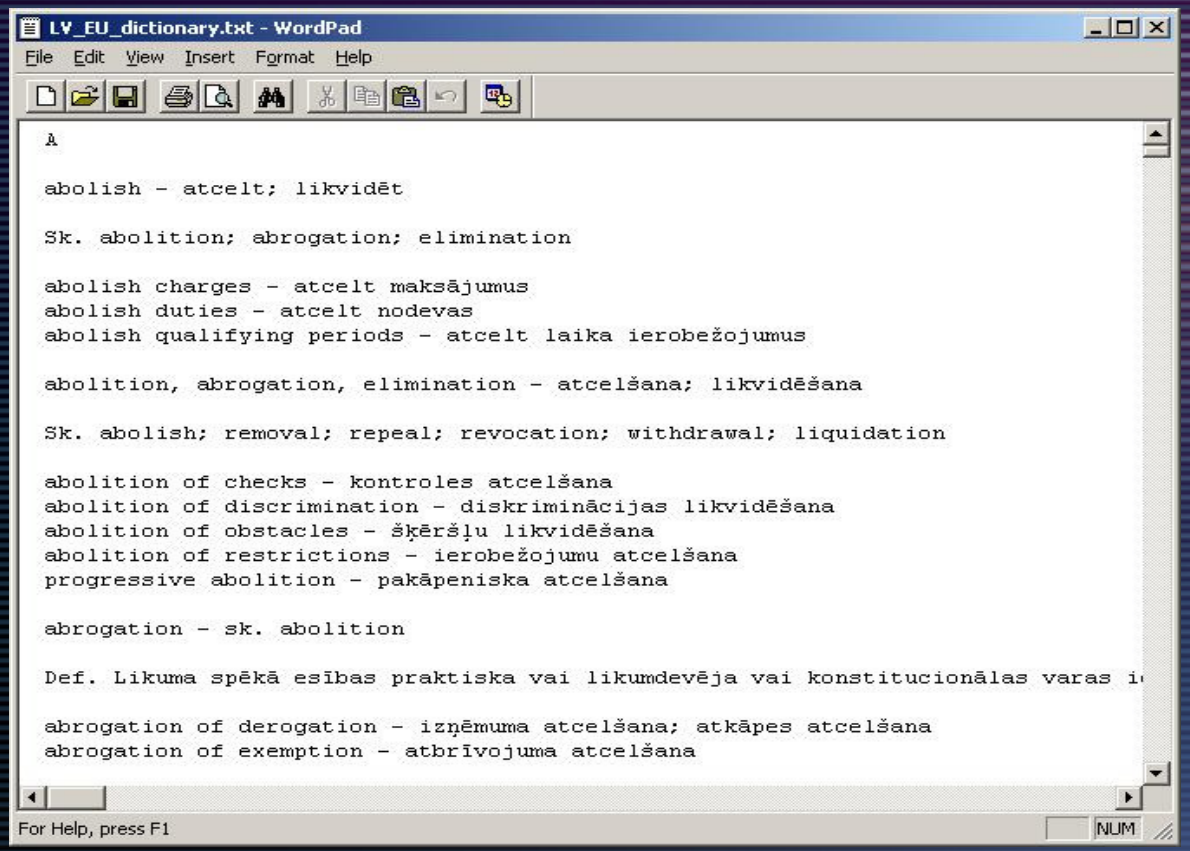
14*

eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

EN-LV Term collection in TXT format

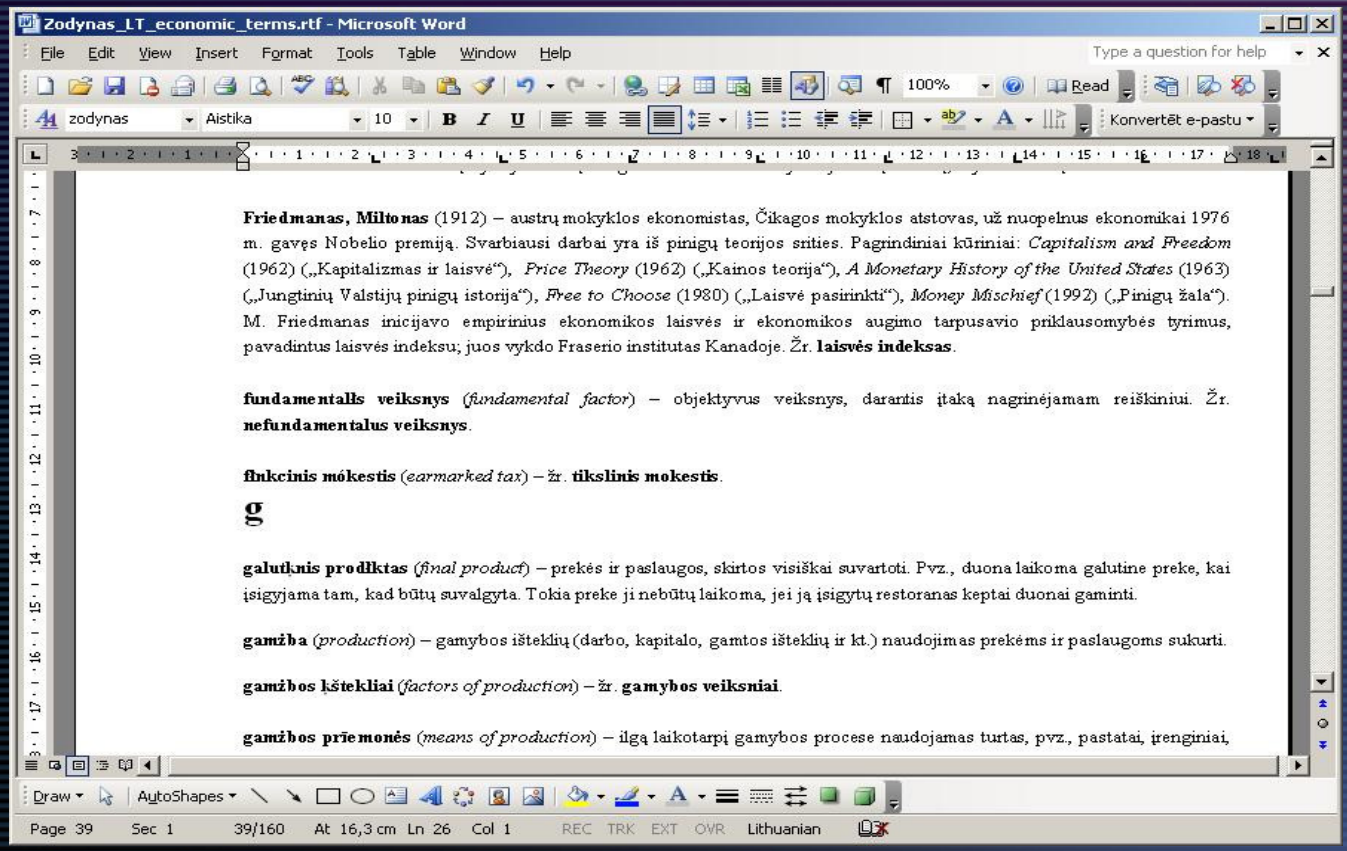


eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

LT Explanatory dictionary of Economy terminology in Word format

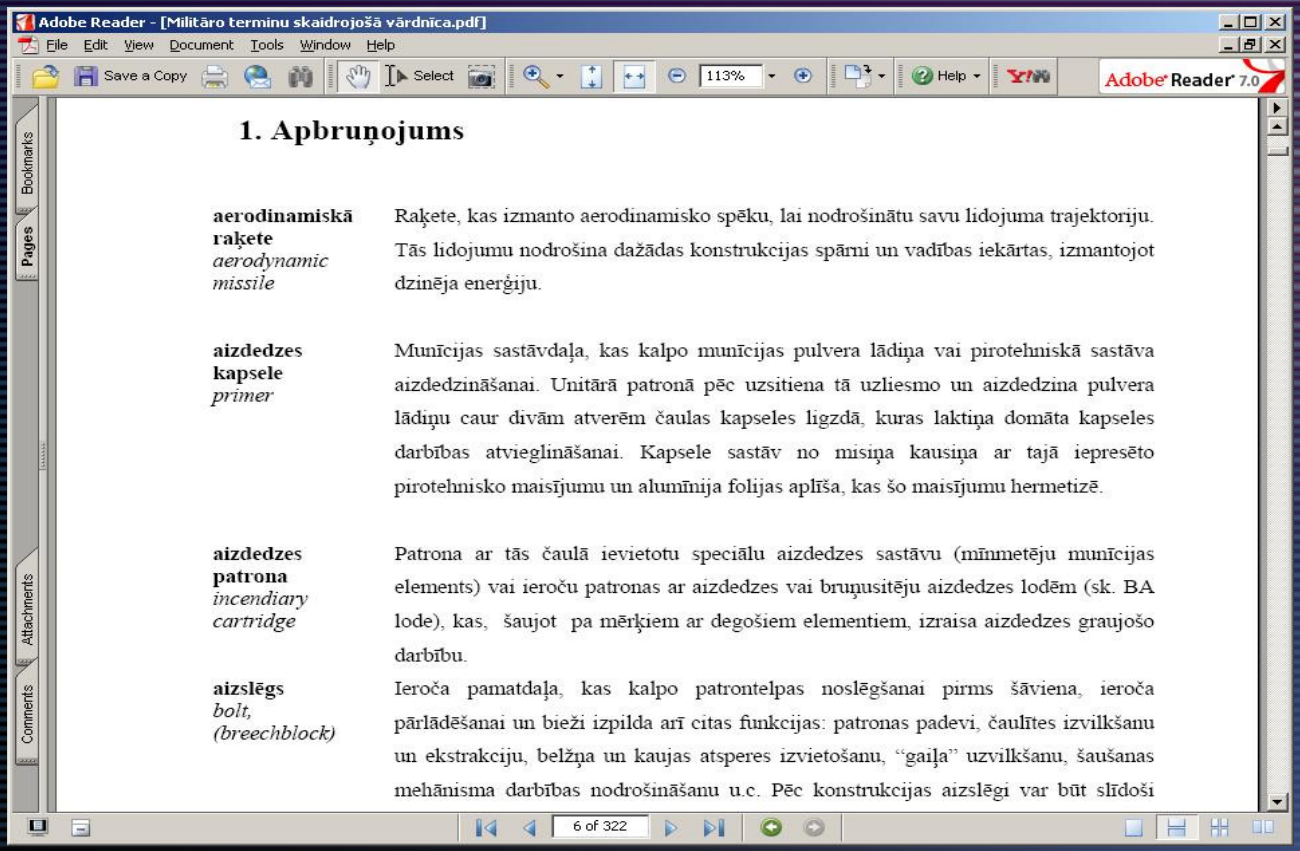


eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

LV Military terminology in PDF format

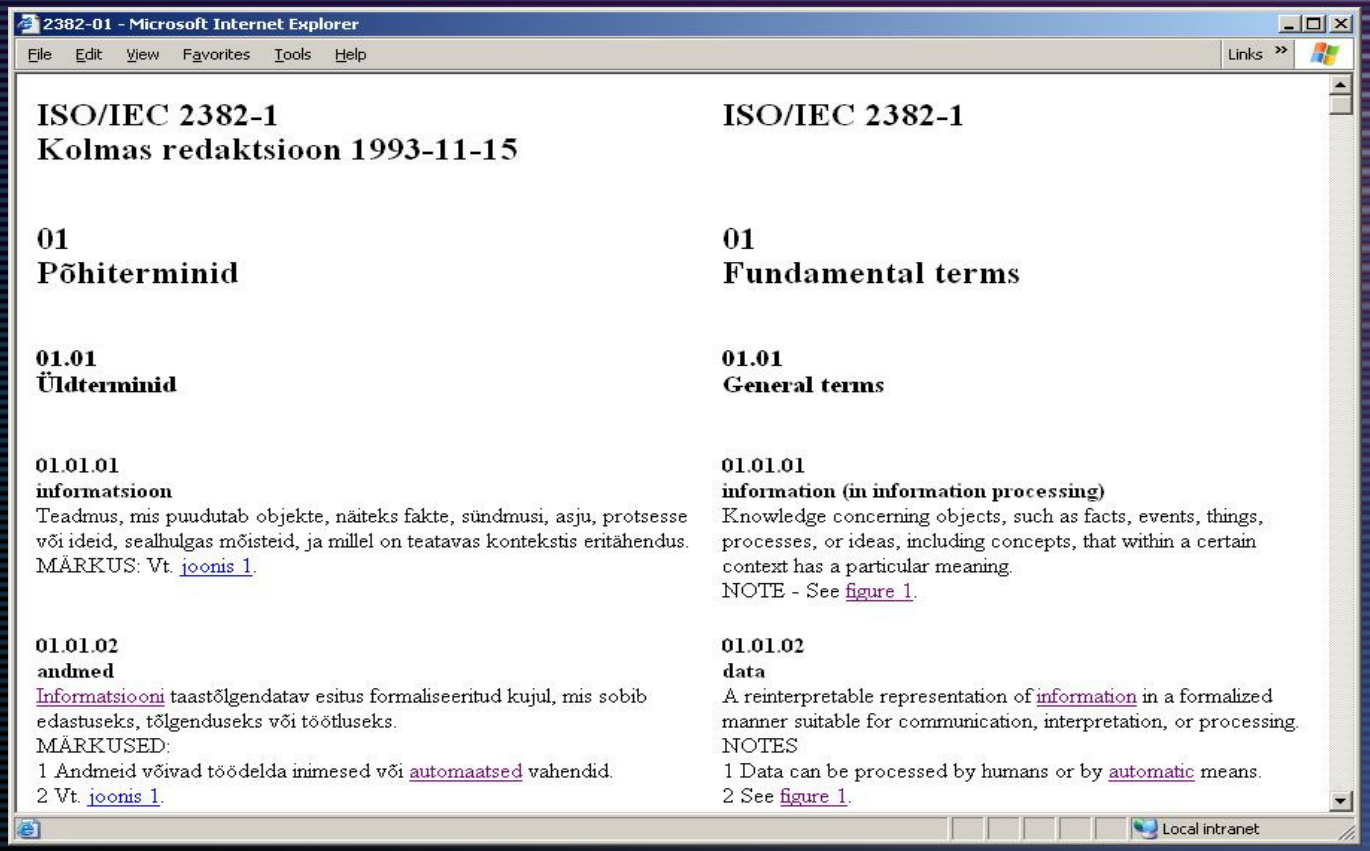


eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

EE ISO IT Standard terminology in HTML

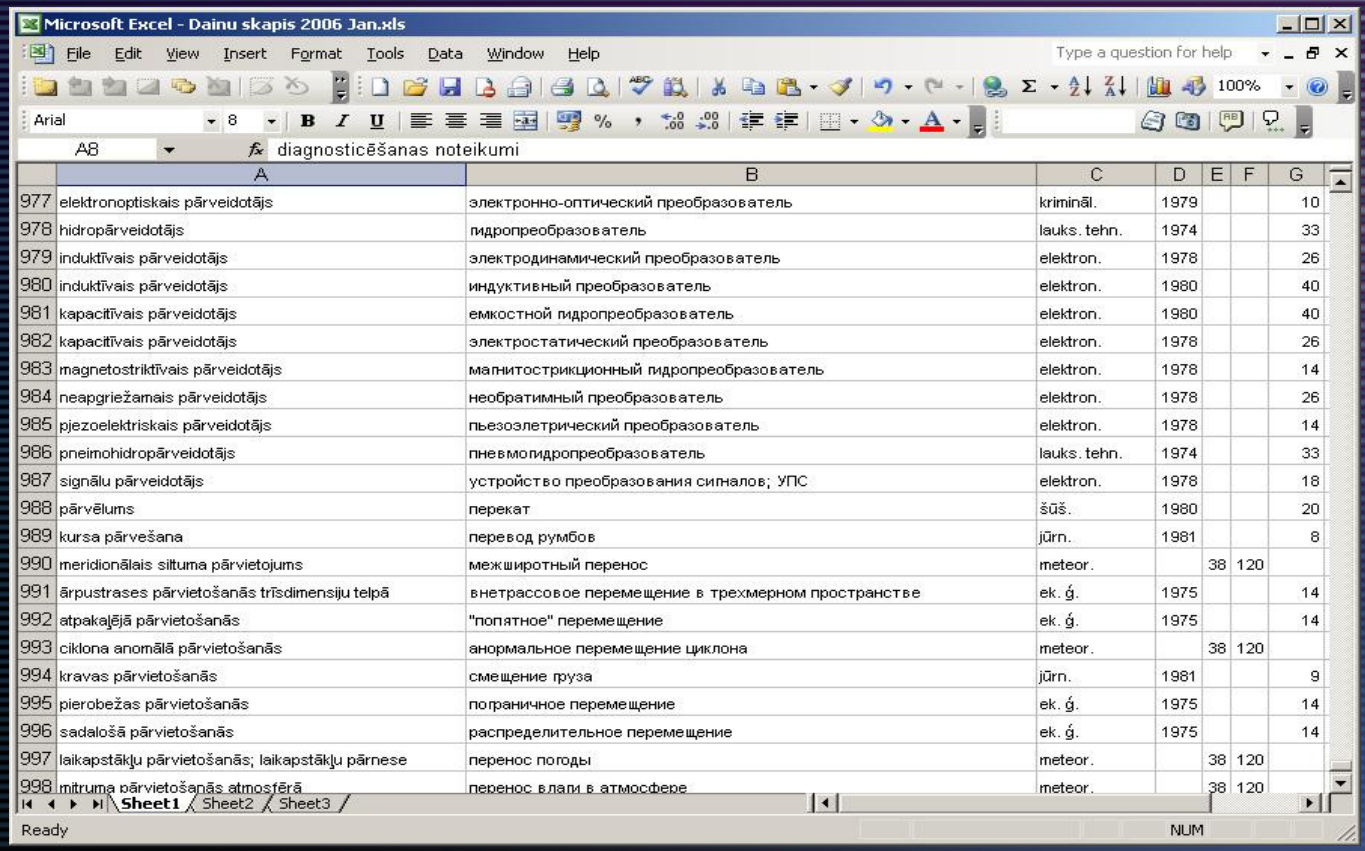


eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

LV-RU Multi-Domain term list in Excel


Microsoft Excel - Dainu skapis 2006 Jan.xls

File Edit View Insert Format Tools Data Window Help

Type a question for help

Arial 8

A8 diagnosticēšanas noteikumi

	A	B	C	D	E	F	G
977	elektronoptiskais pārveidotājs	электронно-оптический преобразователь	krimināl.	1979			10
978	hidropārveidotājs	гидропреобразователь	lauks. tehn.	1974			33
979	induktīvais pārveidotājs	электродинамический преобразователь	elektron.	1978			26
980	induktīvais pārveidotājs	индуктивный преобразователь	elektron.	1980			40
981	kapacitīvais pārveidotājs	емкостной гидропреобразователь	elektron.	1980			40
982	kapacitīvais pārveidotājs	электростатический преобразователь	elektron.	1978			26
983	magnetostriktīvais pārveidotājs	магнитострикционный гидропреобразователь	elektron.	1978			14
984	neapgriežamais pārveidotājs	необратимный преобразователь	elektron.	1978			26
985	pjezoelektriskais pārveidotājs	пьезоэлетрический преобразователь	elektron.	1978			14
986	pneimohidropārveidotājs	пневмогидропреобразователь	lauks. tehn.	1974			33
987	signālu pārveidotājs	устройство преобразования сигналов; УПС	elektron.	1978			18
988	pārvēlums	перекат	šūš.	1980			20
989	kursa pārvešana	перевод румбов	jūrn.	1981			8
990	meridionālais siltuma pārvietojums	межширотный перенос	meteor.		38	120	
991	ārpustrases pārvietošanās trīsdimensiju telpā	внетрассовое перемещение в трехмерном пространстве	ek. ģ.	1975			14
992	atpakaļējā pārvietošanās	"попятное" перемещение	ek. ģ.	1975			14
993	ciklona anomālā pārvietošanās	анормальное перемещение циклона	meteor.		38	120	
994	kravas pārvietošanās	смещение груза	jūrn.	1981			9
995	pierobežas pārvietošanās	пограничное перемещение	ek. ģ.	1975			14
996	sadalošā pārvietošanās	распределительное перемещение	ek. ģ.	1975			14
997	laikapstākļu pārvietošanās; laikapstākļu pārnese	перенос погоды	meteor.		38	120	
998	mitruma pārvietošanās atmosfērā	перенос влаги в атмосфере	meteor.		38	120	

Sheet1 Sheet2 Sheet3

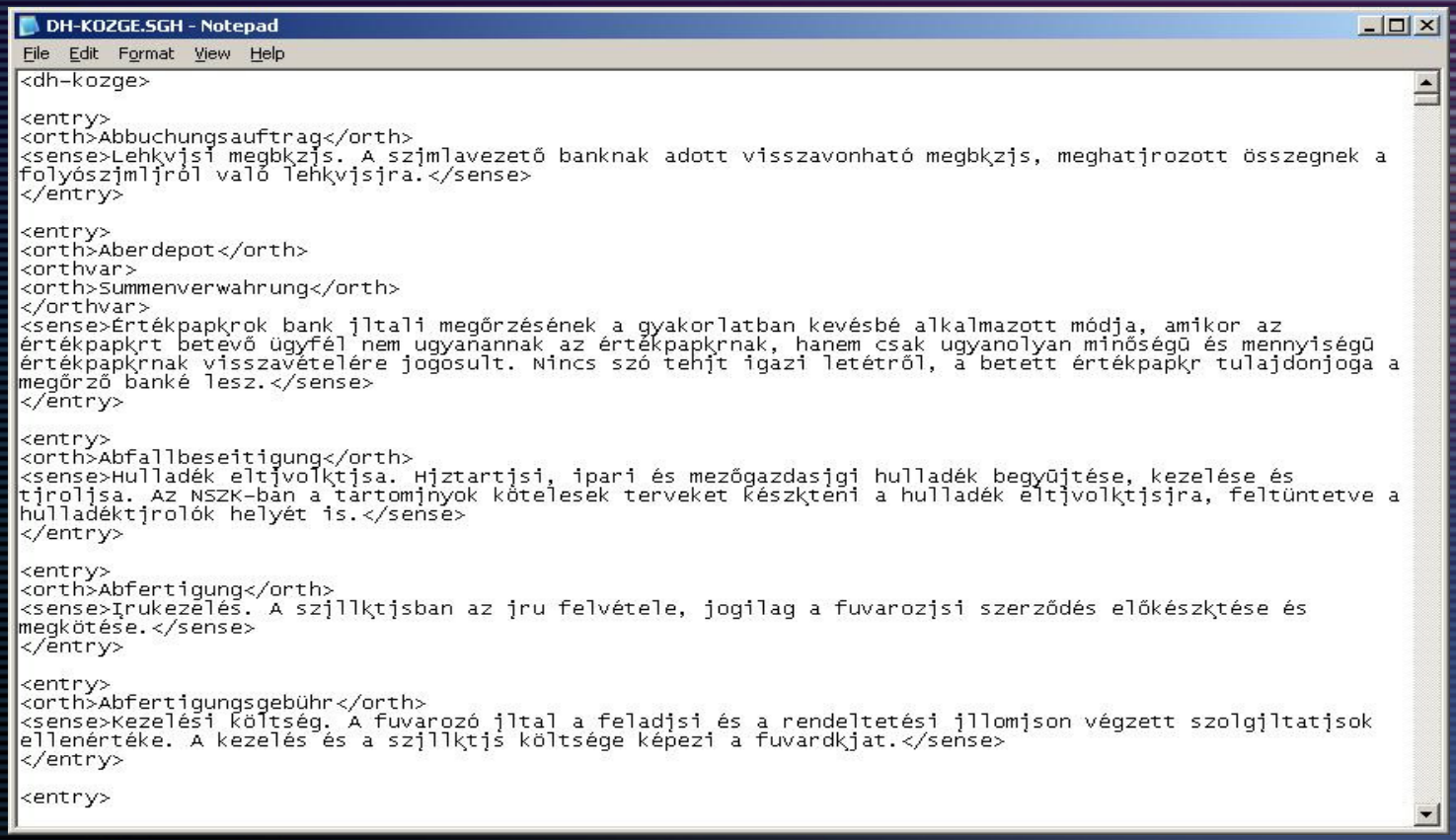
Ready NUM

eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

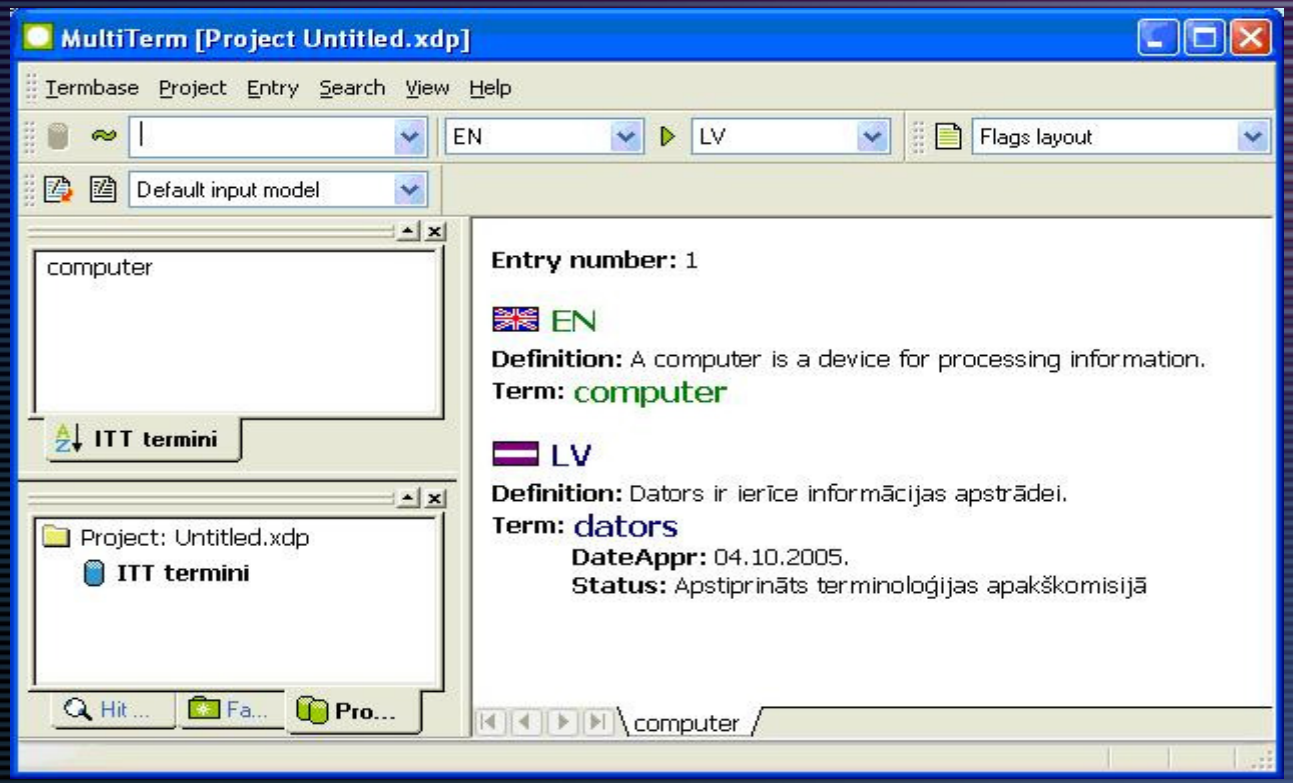
GE-HU Financial Terminology in Tagged text format



eContent

EuroTermBank

EN-LV IT Termbase in Trados MultiTerm



eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.

EuroTermBank

Standards – technical base for cooperation

- To ensure completeness and uniformity of data
- To ensure exchangeability of data
- To ensure wide applicability of data

eContent

EuroTermBank

Terminology Standards for Resource Description and Data Modeling

- TeDIF - The Terminology Documentation Interchange Format
- ISO 12200: Computer applications in terminology – Machine-readable terminology interchange format (MARTIF)
- ISO 12620: Computer applications in terminology – Data categories
- ISO 16642: Computer applications in terminology – Terminological markup framework (TMF)
- TBX: LISA (Localization Industry Standards Association) developed industry standard compliant with TMF, specifying data categories from ISO 12620 and adapting XML style compatible with MARTIF

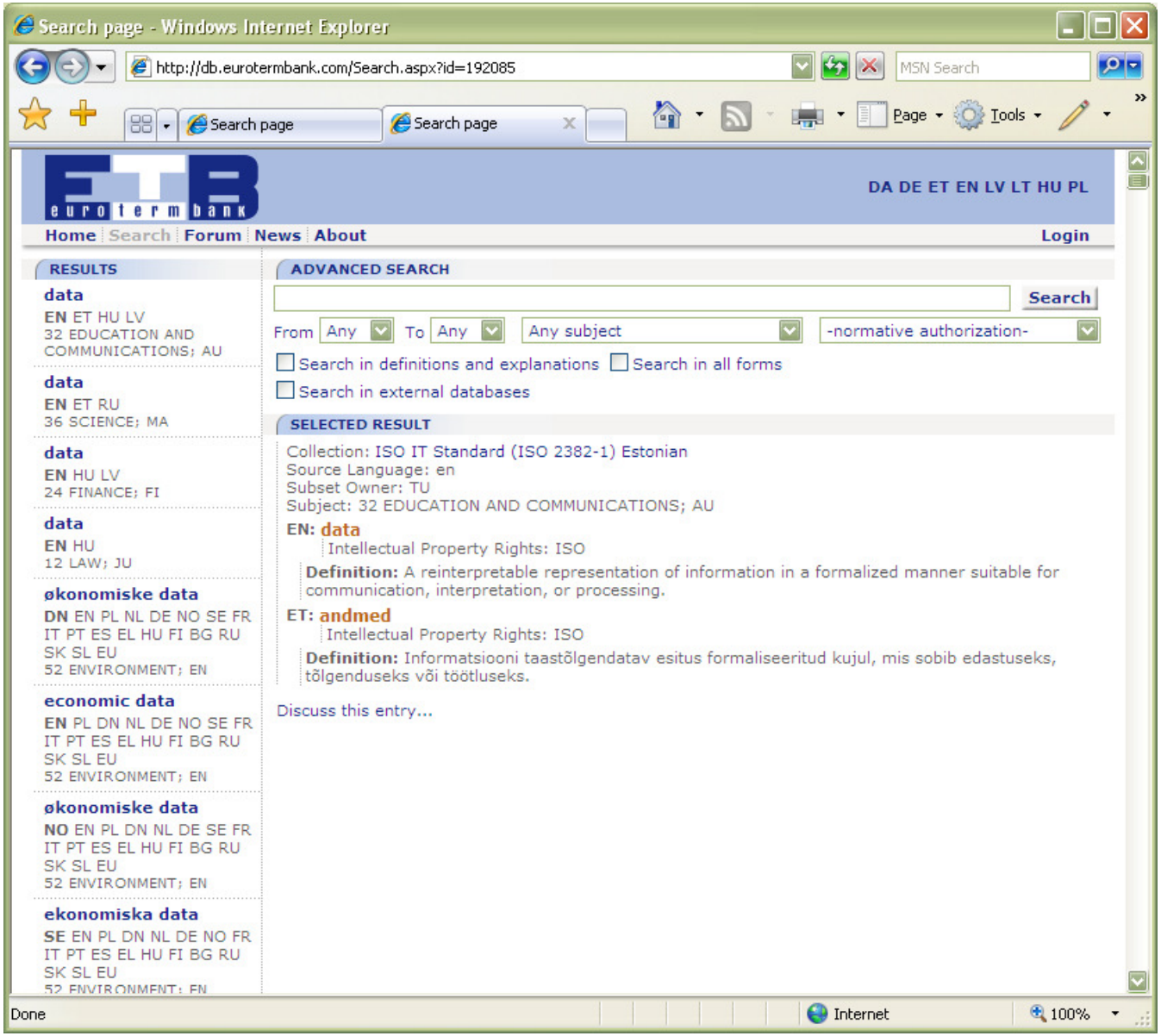
eContent

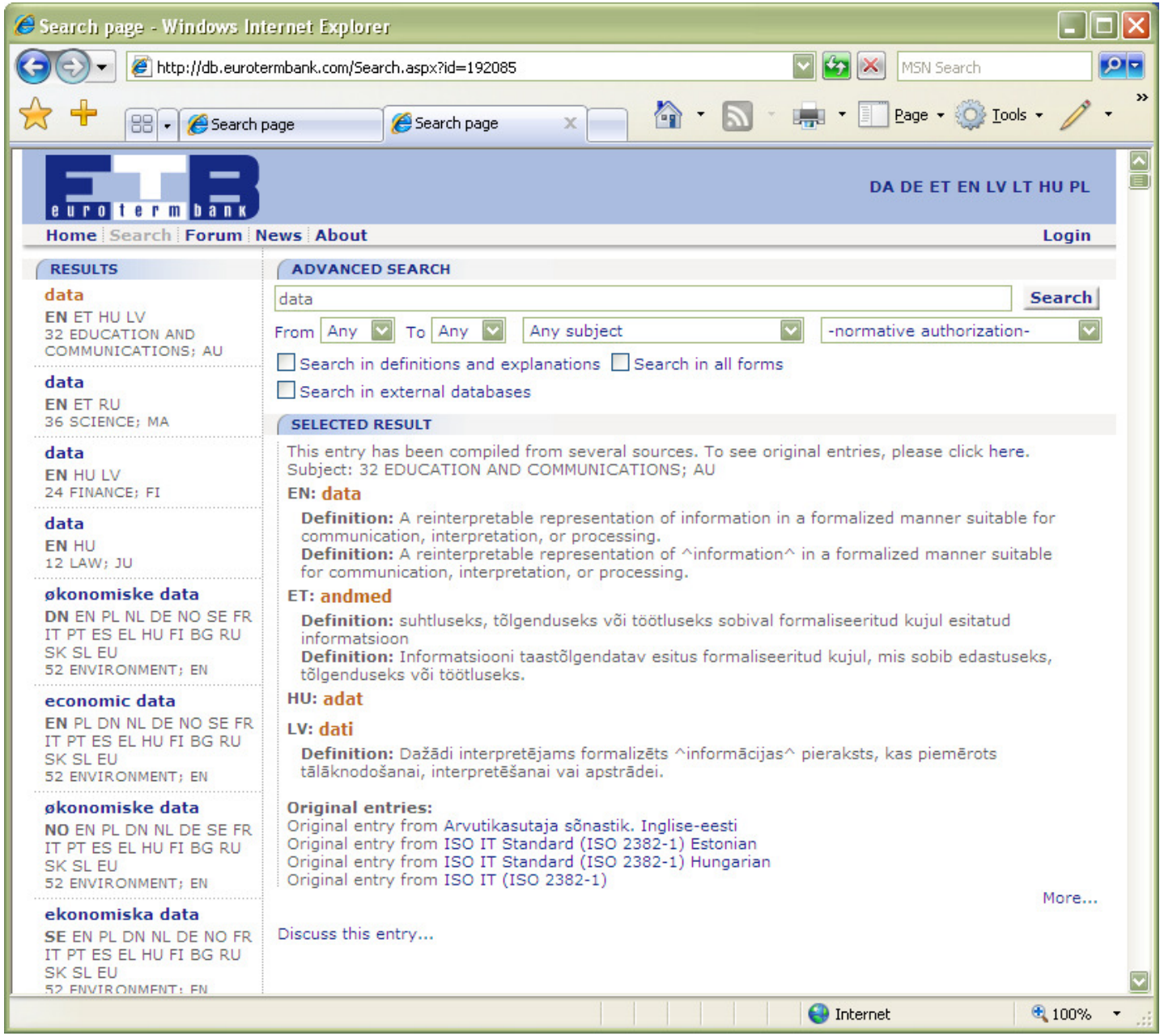
EuroTermBank

Current Status and Further Developments

- Current methodology and best practices in new EU member countries summarized
- Methodology recommendations prepared
- 468 terminology resources identified and described
- >700 000 entries collected, >400 000 in database, 200 000 digitized from paper format
- *Beta* version of online database launched, publicly available from December 2006
- Developments on entry consolidation from several sources to form unified concept-oriented representation

eContent


Search page - Windows Internet Explorer
http://db.eurotermbank.com/Search.aspx?id=192085
MSN Search
Search page
Search page
Page
Tools
DA DE ET EN LV LT HU PL
Home Search Forum News About
Login
RESULTS
data
EN ET HU LV
32 EDUCATION AND COMMUNICATIONS; AU
data
EN ET RU
36 SCIENCE; MA
data
EN HU LV
24 FINANCE; FI
data
EN HU
12 LAW; JU
økonomiske data
DN EN PL NL DE NO SE FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
economic data
EN PL DN NL DE NO SE FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
økonomiske data
NO EN PL DN NL DE SE FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
ekonomiska data
SE EN PL DN NL DE NO FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
ADVANCED SEARCH
Search
From Any To Any Any subject -normative authorization-
Search in definitions and explanations
Search in all forms
Search in external databases
SELECTED RESULT
Collection: ISO IT Standard (ISO 2382-1) Estonian
Source Language: en
Subset Owner: TU
Subject: 32 EDUCATION AND COMMUNICATIONS; AU
EN: data
Intellectual Property Rights: ISO
Definition: A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing.
ET: andmed
Intellectual Property Rights: ISO
Definition: Informatsiooni taastõlgendatav esitus formaliseeritud kujul, mis sobib edastuseks, tõlgenduseks või töötluseks.
Discuss this entry...
Done
Internet
100%

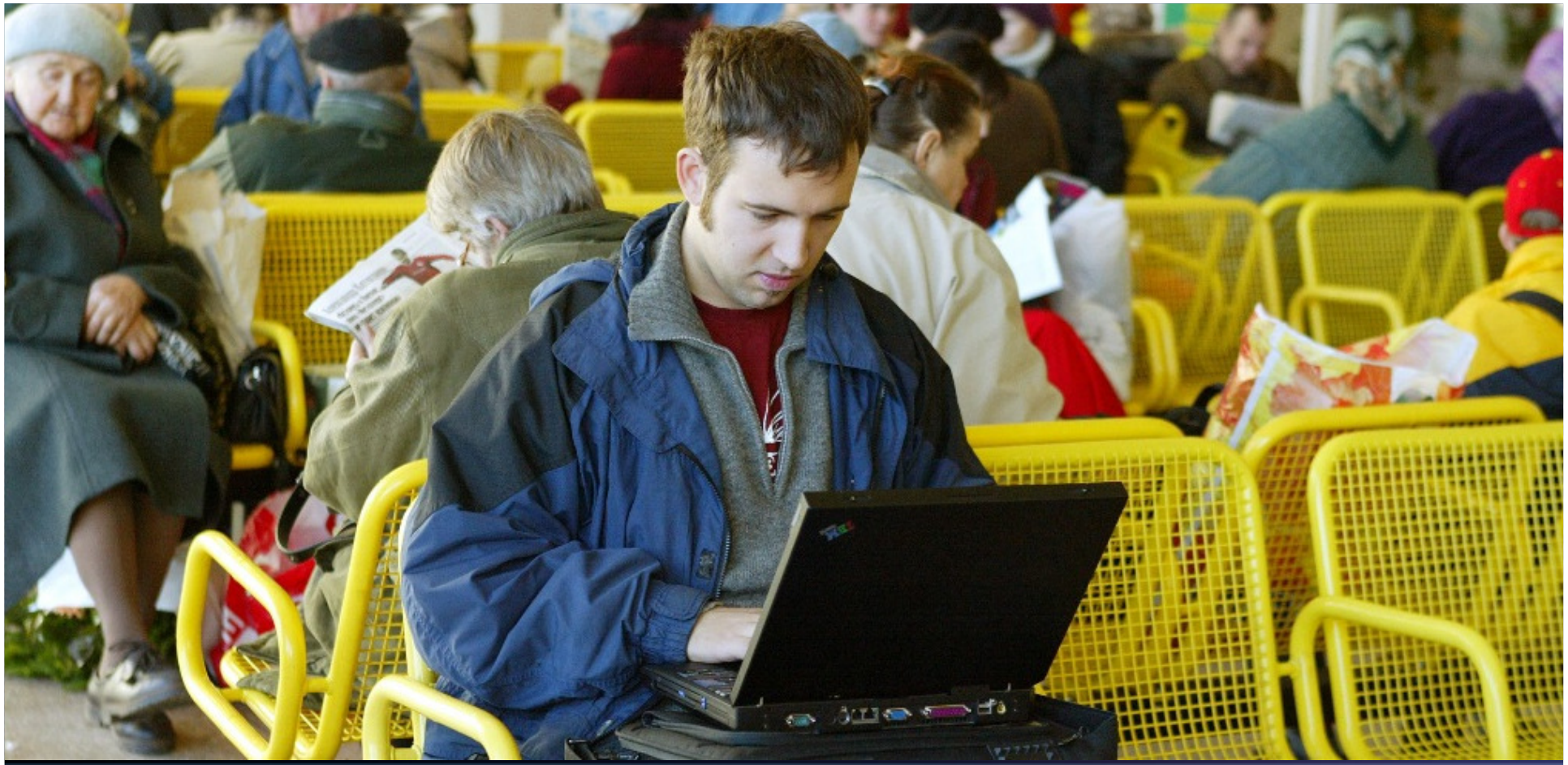

Search page - Windows Internet Explorer
http://db.eurotermbank.com/Search.aspx?id=192085
MSN Search
Search page
Search page
Page
Tools
DA DE ET EN LV LT HU PL
Home Search Forum News About
Login
RESULTS
data
EN ET HU LV
32 EDUCATION AND COMMUNICATIONS; AU
data
EN ET RU
36 SCIENCE; MA
data
EN HU LV
24 FINANCE; FI
data
EN HU
12 LAW; JU
økonomiske data
DN EN PL NL DE NO SE FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
economic data
EN PL DN NL DE NO SE FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
økonomiske data
NO EN PL DN NL DE SE FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
ekonomiska data
SE EN PL DN NL DE NO FR IT PT ES EL HU FI BG RU SK SL EU
52 ENVIRONMENT; EN
ADVANCED SEARCH
data
Search
From Any To Any Any subject -normative authorization-
Search in definitions and explanations
Search in all forms
Search in external databases
SELECTED RESULT
This entry has been compiled from several sources. To see original entries, please click here.
Subject: 32 EDUCATION AND COMMUNICATIONS; AU
EN: data
Definition: A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing.
Definition: A reinterpretable representation of ^information^ in a formalized manner suitable for communication, interpretation, or processing.
ET: andmed
Definition: suhtluseks, tõlgenduseks või töötluseks sobival formaliseeritud kujul esitatud informatsioon
Definition: Informatsiooni taastõlgendatav esitus formaliseeritud kujul, mis sobib edastuseks, tõlgenduseks või töötluseks.
HU: adat
LV: dati
Definition: Dažādi interpretējams formalizēts ^informācijas^ pieraksts, kas piemērots tālāknodošanai, interpretēšanai vai apstrādei.
Original entries:
Original entry from Arvutikasutaja sõnastik. Inglise-eesti
Original entry from ISO IT Standard (ISO 2382-1) Estonian
Original entry from ISO IT Standard (ISO 2382-1) Hungarian
Original entry from ISO IT (ISO 2382-1)
More...
Discuss this entry...
Internet
100%

EuroTermBank

Key Success Factors

- Assessment and accumulation of the best practice
- Reliance on ISO and industry standards
- Standardized procedures and agreements
- Distributed approach
- Openness and flexibility

eContent



Thank you
and
Welcome to

www.eurotermbank.com

andrejs@tilde.lv

eContent

Terminology Summit 2006, Brussels
13-14 November, 2006.