



**EuroVoc and thesauri from EU institutions and agencies
Interoperability and perspectives for collaborative thesaurus
management**

Publications Office of the EU
Christine LAABOUDI-SPOIDEN

ISKO-UK Conference, 9 July 2013, London

- Maintained by the Publications Office of the EU
- Version 4.4 : 6 800 concepts, 21 Domains, 120 Microthesauri
 - Published on 15 December 2012
- Multilingual
 - 24 languages
 - 23 EU official languages + Serbian
 - Links to other languages
 - Basque, Catalan, Macedonian
- Multidisciplinary thesaurus
 - Strong coverage
 - Parliamentary activities
 - European Union, EU legislation, EU activities, EU policies, EU Institutions
 - EU regions
- Exact equivalence between concepts
 - Symmetry
 - No coverage for national or regional concepts



■ A concept represented in every language

<http://eurovoc.europa.eu/6541>

politica migratoria dell'UE [4.4]

CONTRIBUTE

MAP

UF *politica di immigrazione comunitaria*
politica migratoria comunitaria
politica migratoria dell'Unione europea

28 QUESTIONI SOCIALI

MT 2811 migrazione

BT1 politica migratoria

BT2 migrazione

RT libera circolazione delle persone [1231]

LANGUAGE EQUIVALENTS

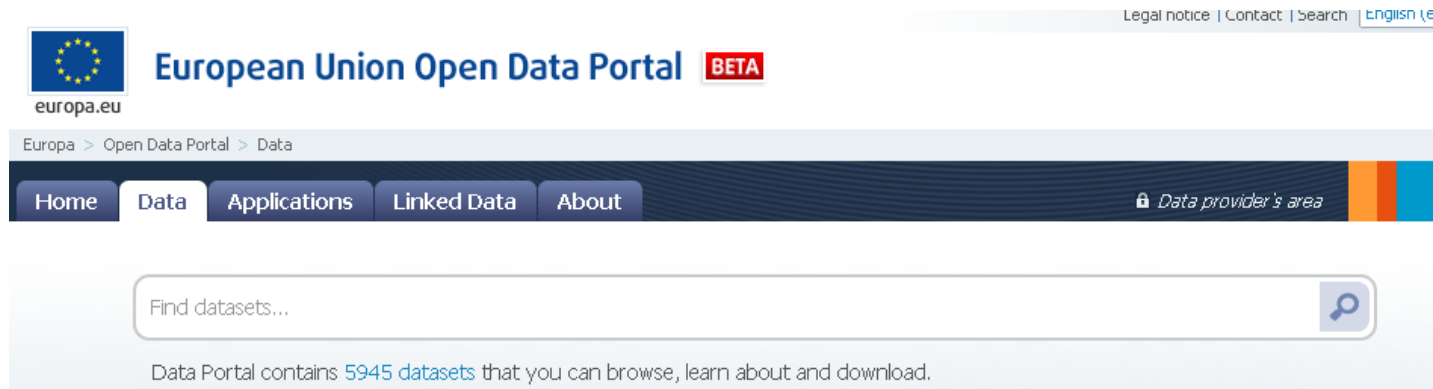
BG миграционна политика на ЕС [4.4]
ES política migratoria de la UE [4.4]
CS migrační politika EU [4.4]
DA EU's migrationspolitik [4.4]
DE EU-Migrationspolitik [4.4]
ET ELi migratsioonipoliitika [4.4]
EL μεταναστευτική πολιτική της ΕΕ [4.4]
EN EU migration policy [4.4]
FR politique migratoire de l'UE [4.4]
IT politica migratoria dell'UE [4.4]
LV ES migrācijas politika [4.4]
LT ES migracijos politika [4.4]
HU uniós migrációs politika [4.4]
MT politika dwar il-migrazzjoni tal-UE [4.4]
NL migratiebeleid van de EU [4.4]
PL polityka migracyjna UE [4.4]
PT política migratória da UE [4.4]
RO politica UE în domeniul migrației [4.4]
SK migračná politika EÚ [4.4]
SL migracijska politika EU [4.4]
FI EU:n siirtolaispolitiikka [4.4]
SV EU:s migrationspolitik [4.4]

HR migracijska politika EU-a [4.4]
SR EU migration policy (*under translation*)




■ EU Open Data Portal <http://odp/en/>

- Single point of access to metadata produced by the EU institutions and other bodies
- Free to use, reuse, link and redistribute for commercial or non-commercial purposes.
- EuroVoc is distributed in RDF or XML



Legal notice | Contact | Search | English (e)

 **European Union Open Data Portal** **BETA**
europa.eu

Europa > Open Data Portal > Data

Home Data Applications Linked Data About 🔒 Data provider's area

Find datasets... 🔍

Data Portal contains [5945 datasets](#) that you can browse, learn about and download.



EuroVoc, the EU's multilingual thesaurus

Publisher

Publications Office »

Description

EuroVoc is a multilingual, multidisciplinary thesaurus covering the activities of the EU. It contains terms in 22 EU languages (Bulgarian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish and Swedish), plus Croatian and Serbian.

Resources

- [DOWNLOAD](#) *Distribution – Alignment – EuroVoc and Agrovoc »*
- [DOWNLOAD](#) *Distribution – Alignment – EuroVoc and ECLAS »*
- [DOWNLOAD](#) *Distribution – Alignment – EuroVoc and Gemet »*
- [DOWNLOAD](#) *Distribution – EuroVoc SKOS/RDF »*
- [DOWNLOAD](#) *Distribution – EuroVoc XML »*
- [DOWNLOAD](#) *Distribution – EuroVoc ontology (OWL) »*
- [DOWNLOAD](#) *Documentation – About EuroVoc »*
- [DOWNLOAD](#) *Documentation – EuroVoc Help »*
- [DOWNLOAD](#) *Documentation – EuroVoc Ontology »*
- [DOWNLOAD](#) *Documentation – Publications and documents »*

Additional Information

URL
<http://eurovoc.europa.eu/>

Alternative Title
EuroVoc

Status
Completed

Keywords

EU activity
EU policy
controlled vocabulary
multilingual thesaurus

Licence: Europa Legal Notice



Thesaurus interoperability initiatives at the PO

- 2008: Study for a proposed methodology for thesaurus interoperability
 - Feasibility study
 - ITTIG - Institute of Legal Information Theory and Techniques (IT)
 - EuroVoc \leftrightarrow Eclas, Gemet, Unesco, ETT

- 2010 : Development of a Thesaurus Alignment Tool
 - TAE - Thesaurus Alignment Environment
 - Developed under a contract of the Publications Office
 - Interface for **automated mapping** and **manual validation**

EuroVoc - Eclas
Alignment parameters

concept groups	EuroVoc Eclas
algorithm	AROMA Alignment
thresholds	0.5 1
parameters	


ECLAS
 European Commission Libraries Catalogue
 European Commission > Central Library > ECLAS Thesaurus



Thesaurus Alignment Environment (TAE) (1/2)

Number of correspondences
between two thesauri

ALIGNMENT

Statistics

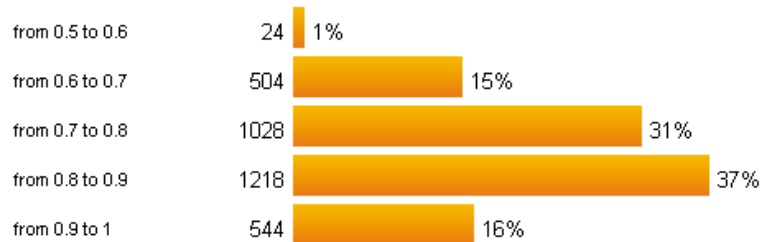
3318 correspondences are bound to EuroVoc - Eclac

[export statistics](#)

breakdown by review status



breakdown by score



breakdown by mapping property



Thesaurus Alignment Environment (TAE) (2/2)

Semantic correspondences
between concepts of two thesauri

MAPPING

Mappings for Eclas

Actions: [validate](#) [review](#) [reject](#) [delete](#)

select elements in all pages

	<input type="checkbox"/> Concept T2	Mapping type	Concept T1	Score	Status
1	<input type="checkbox"/> Political doctrines	exact match	political ideology	0.7	validated
2	<input type="checkbox"/> Political ideologies	exact match	political ideology	0.7	validated
3	<input type="checkbox"/> Political institutions	exact match	political institution	0.8	validated
4	<input type="checkbox"/> Socialist international	exact match	Socialist International	0.9	validated
5	<input type="checkbox"/> Legality	exact match	legality	0.8	validated
6	<input type="checkbox"/> Conservatism	exact match	conservatism	0.7	validated
7	<input type="checkbox"/> Democracy	exact match	democracy	0.7	validated
8	<input type="checkbox"/> People's Democracies	exact match	people's democracy	0.8	validated
9	<input type="checkbox"/> Dictatorship	exact match	dictatorship	0.7	validated
10	<input type="checkbox"/> Participatory democracy	exact match	participatory democracy	0.8	validated



Thesaurus Alignment - Automated generation

- Automated alignment
 - Algorithm: Aroma
- No pivot language
 - Alignment of concepts instead of terms
 - Alignments set up between **Uri's (concept)** and not terms
- Score
 - Assigned according to the number of matching lexical values

Actions: [validate](#) [review](#) [reject](#) [delete](#)

select elements in all pages

<input type="checkbox"/>	Concept T1	Mapping type	Concept T2	Score	Status
<input type="checkbox"/>	political ideology	exact match	Political doctrines	0.7	validated
<input type="checkbox"/>	political ideology	exact match	Political ideologies	0.7	validated
<input type="checkbox"/>	political institution	exact match	Political institutions	0.8	validated
<input type="checkbox"/>	Socialist International	exact match	Socialist international	0.9	validated

- Final Results
 - SKOS-Core (RDF)
 - EDDOAL



Thesaurus Alignment - Validation

Human validation

- Status: validated, to be reviewed, rejected

Actions: validate review reject delete 1

select elements in all pages

	<input type="checkbox"/>	Concept T1	Mapping type	Concept T2	Score	Status
1	<input type="checkbox"/>	political ideology	exact match	political doctrines	0.7	validated
2	<input type="checkbox"/>	political ideology	exact match	political ideologies	0.7	validated
3	<input type="checkbox"/>	political ideology	exact match	political ideologies	0.8	validated
4	<input type="checkbox"/>	Socialist	exact match	Socialist	0.9	validated
5	<input type="checkbox"/>	legality	exact match	legality	0.8	validated
6	<input type="checkbox"/>	conservative	exact match	conservative	0.7	validated
7	<input type="checkbox"/>	democracy	exact match	democracy	0.7	validated
8	<input type="checkbox"/>	people's	exact match	people's	0.8	validated
9	<input type="checkbox"/>	dictatorship	exact match	dictatorship	0.7	validated
10	<input type="checkbox"/>	participatory	exact match	participatory democracy	0.8	validated
11	<input type="checkbox"/>	protocol	exact match	Protocol	0.7	validated
12	<input type="checkbox"/>	social democracy	exact match	Social democrats	0.8	validated
13	<input type="checkbox"/>	veto	exact match	Veto	0.9	validated
14	<input type="checkbox"/>	rule of law	exact match	Rule of law	0.7	validated
15	<input type="checkbox"/>	head of State	exact match	Heads of State	0.8	validated
16	<input type="checkbox"/>	republic	exact match	Republic	0.7	validated



Thesaurus Alignment - Types of matches

▪ exactMatch

<input type="checkbox"/>	Concept T1	Mapping type	▼ Concept T2	Score	Status
<input type="checkbox"/>	acid rain	exact match	acid rain	1.0	validated

▪ broadMatch

<input type="checkbox"/>	Concept T1	Mapping type	Concept T2	Score	Status
<input type="checkbox"/>	palaeontology	broad match	earth sciences	1.0	validated

▪ narrowMatch

<input type="checkbox"/>	Concept T1	Mapping type	Concept T2	Score	Status
<input type="checkbox"/>	mammal	narrow match	wild mammal	1.0	validated
<input type="checkbox"/>	mammal	narrow match	marine mammal	1.0	validated

▪ relatedMatch

<input type="checkbox"/>	Concept T1	Mapping type	Concept T2	Score	Status
<input type="checkbox"/>	cinema	related match	film industry	0.9	validated

▪ closeMatch

<input type="checkbox"/>	Concept T1	Mapping type	Concept T2	Score	Status
<input type="checkbox"/>	pay	close match	wage	0.9	validated



Thesaurus alignment - Constraints (1/2)

■ Datasets stored in SKOS-Core

- Conversion in SKOS-Core (from SKOS-XL)
- Reload vocabularies with each new releases

■ One dedicated installation for alignment

- No gateway to our Thesaurus Maintenance System
 - SKOS-XL

■ 2 different installations for thesaurus maintenance in the PO

- EuroVoc
- ETT (Cedefop)
- Two similar data models



Thesaurus alignment - Constraints (2/2)

■ Published URI's → resources available as LOD

- ECLAS (Thesaurus of the European Commission's Library)
- AgroVoc (FAO thesaurus - UN)
- GEMET (General Multilingual Environmental Thesaurus - EEA)

■ Published but No URI's → resources not available as LOD

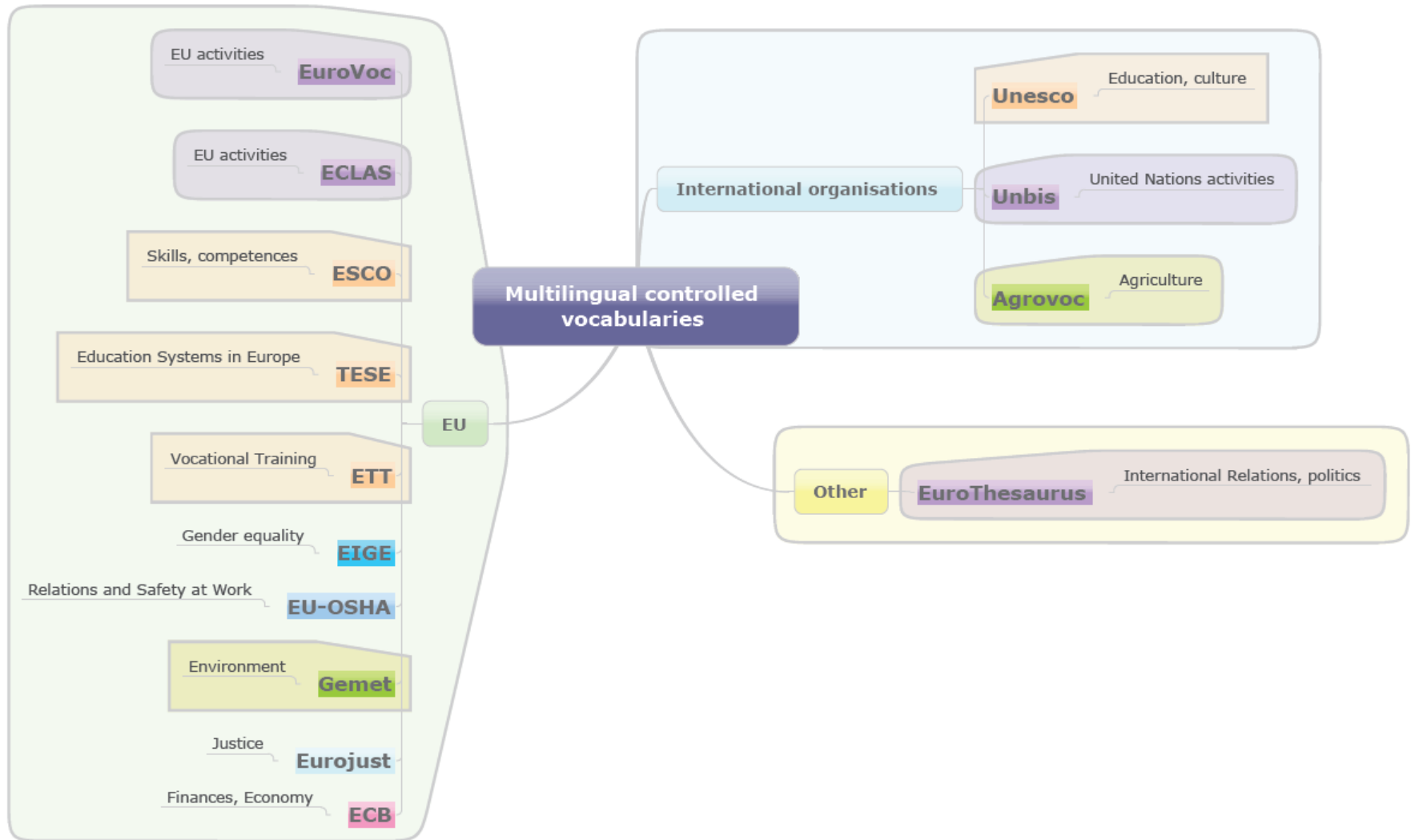
- ETT - European Training Thesaurus (Cedefop)
- TESE - Thesaurus of education systems in Europe

■ Not Published

- EIGE (Agency for gender equality)
- Eurojust (EU judicial cooperation unit)
- OSHA - European agency for health at work



A network of multilingual thesauri (1/2)



A network of multilingual Thesauri (2/2)

■ Published by

- EU institutions and agencies
- International organisations
- National or governmental bodies

■ Coverage

- Different levels of specialisation
 - Generic \leftrightarrow specialised
- Overlapping
 - Domains, concepts
 - Language coverage

■ Dissemination

- Individual publication
- Different formats (SKOS/RDF, HTML, PDF, Excel)
- Different presentations
 - Hierarchy
 - Alphabetical list of terms
- Online vs. Not published (PDF, XLS)



Thesaurus Alignment - Lessons learnt

■ We maintain

- Similar or close concepts
- Similar translations

	ExactMatch	Total Number of concepts	%
OSHA	688	3264	21
GEMET	1129	5208	22
ETT	1089	1564	70
ECLAS	3141	6041	52

■ We use

- Different maintenance systems
 - Maintenance, hosting or licensing fees
- Different formats and standards (data model)
- Different means of dissemination

■ Our concepts and metadata are not interlinked (LOD)

■ Duplicated efforts and resources in terms of thesaurus maintenance and dissemination



- Discussion forum between multilingual thesaurus managers
 - EU multilingual thesauri
 - EU institutions (ECLAS, EuroVoc)
 - EU agencies (Cedefop, Eurydice Tese, Gemet, ESCO, EU-OSHA)
 - International multilingual thesauri
 - UNBIS, FAO, UNESCO
 - National multilingual thesaurus
 - Eurothesaurus (European Thesaurus on International Relations and Area Studies)
 - SWP Berlin

- First meeting in June 2010
 - Publications Office, Luxembourg

- Future thesaurus collaboration foreseen at different levels



Collaborative Thesaurus Environment (1/3)

■ Role of the Publications Office

- Offer a Thesaurus Management Platform to EU institutions and bodies
- One installation of **Vocbench 2.0**

■ Advantages

- Open-Source software developed by FAO
- Full SKOS compliant
- No licensing fees, shared development fee between EU actors
- Development of new modules (plug-in) shared amongst the Vocbench community
 - Collaborative maintenance module
 - Linked data discovery engine

■ Share the maintenance of similar concepts and their translations

- Management of generic and EU concepts at EuroVoc level
- Management of specialised concepts by the specialised thesauri (EU agencies)



- Expand the coverage of EuroVoc with specialized concepts
 - Integration of specialised concepts in EuroVoc
 - Maintained by the specialised agencies
 - Development of new specialised microthesauri

- EU and generic concepts maintained in EuroVoc
 - Institutions (EU and international)
 - EU policies, EU treaties
 - Types of documents
 - Geography
 - EU regions

- Ongoing
 - ETT - Cedefop (Vocational training)
 - European Central Bank (Finance and economics)

- Planned
 - EIGE (Gender equality)
 - Eurojust (Judicial cooperation)



Collaborative Thesaurus Environment (3/3)

■ Collaborative management of common concepts

- One concept belongs to different conceptSchemes (thesauri)
- One concept = One owner
 - Maintained by the specific institutions or agencies
- Individual exports by thesaurus + Dissemination
 - Creation of definite namespace / URI per thesauri
 - Generate Linked data

■ Planned

- ECLAS (EC Central Library Thesaurus)
- ETT (Cedefop)

■ Practical applications

- In-depth indexing of EU documentary collection
- Double indexing and cross-information retrieval
 - EU-Bookshop, EUR-Lex, Eclas



Thesaurus collaboration - Practical applications (1/2)

EUR-Lex **BETA TEST**
Access to European Union law

About EUR-Lex | Site map | A-Z | FAQ

Quick search
Advanced

EUROPA » EU law and publications » EUR-Lex » Search results » EUR-Lex - 32013D0129 - EN

Official Journal | EU law and related documents | National law | Legislative procedures | More

Back to result list | Next >>

Document 32013D0129

About this document | Text | Procedure | Linked documents | All | Collapse all | Expand all

Title and reference

2013/129/EU: Council Decision of 7 March 2013 on subjecting 4-methylamphetamine to control measures

OJ L 72, 15/03/2013, p. 11-12 (BG, ES, CS, DA, DE, ET, EL, EN, FR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV)

Dates

- **Date of document:** 07/03/2013
- **Date of effect:** 16/03/2013; Entry into force Date pub. +1 See Art 3
- **End of validity date:** 31/12/9999

Classifications

- **EUROVOC descriptor:**
health control
health risk
penalty
EU control
drug addiction
stimulant
drug surveillance
- **Subject matter:**
Justice and home affairs
- **Directory code:**
15.30.00.00 Environment, consumers and health protection / Health protection
19.30.00.00 Area of freedom, security and justice / Police and judicial cooperation in criminal and customs matters

Chemical structure: 4-methylamphetamine
Molecular formula: C₁₀H₁₅N
Molecular weight: 149.2 g/mol



This site is part of Europa



Publication details



A strategy for green skills?

A study on skill needs and training has wider lessons for successful transition to a green economy



Corporate author(s): [Cedefop – European Centre for the Development of Vocational Training](#)

Themes: [Teaching](#), [Education policy](#)

Target audience: [Education](#)

Key words: [education policy](#), [vocational training](#), [Cedefop](#)



Thesaurus collaboration - Ongoing projects

■ Alignment EuroVoc - ETT (Cedefop)

- Identified the exactMatch to be shared
- Developed a new “vocational training” microthesaurus in EuroVoc
- In-depth indexing in EU Collections



European Union Open Data Portal BETA

Resources

DOWNLOAD

Distribution – Alignment – EuroVoc and Agrovoc »

DOWNLOAD

Distribution – Alignment – EuroVoc and ECLAS »

DOWNLOAD

Distribution – Alignment – EuroVoc and Gemet »

■ Alignment EuroVoc - Eclas

- Identified the exactMatch to be shared
- Double indexing in ECLAS

■ Alignment EuroVoc - Unbis (United Nations Thesaurus)

- Practical applications
 - Cross-retrieval in EU and UN documentary collections
 - Expand the language coverage of EuroVoc
 - Arabian, Russian, Chinese



■ Today

- EuroVoc Resources
 - Dereferencable as HTML representation

■ 2nd semester 2013

- Dereferencing and Content negotiations
 - Conceptschemes (EuroVoc, domains, microthesauri)
 - Concepts
 - Ontology
- Linked dataset description with VOID
- Dissemination of the alignments in the website
 - Association under the term details (HTML representation)



Thanks for your attention

For more information, please contact :

eurovoc@publications.europa.eu

