LREC 2012 ISOcat tutorial – exercises: 'answers'

My Workspace Name Version Administration st Registration stal Check The matic Views The matic Views Metadata Metadata Morphosyntax Semantic Content Represented Syntax Syntax Language Resource Ont Syntax Syntax Language Resource Ont Syntax Syntax<th>ISOcat</th><th></th><th></th><th></th><th></th><th></th><th>1.</th><th>🖓 Welcome Gu</th><th>-</th>	ISOcat						1.	🖓 Welcome Gu	-
Image: Source of the sources 4503 absent 1:0 private private simple Hopperma Image: Source of the sources Image: Source of the sources 1:0 private private open Hopperma Image: Source of the sources Image: Source of the sources 1:0 private private open Hopperma Image: Source of the sources Image: Source of the sources 1:0 private private open Hopperma Image: Source of the sources Image: Source of the sources 1:0 private private open Hopperma Image: Source of the sources Image: Sources	nter keywords here 📸 📸 🔀	4.1 🗎	🔯 🧟 🕼 🖘		Metadata				* 🕄
Image: Thematic Views 1/0 private private implie implie Image: Thematic Views 3764 access protocol 1/0 private private open Hopperma Image: Thematic Views 3764 access protocol 1/0 private private open Hopperma Image: Thematic Views Second 1/0 private private open Hopperma Image: Thematic Views Second 1/0 private private open Hopperma Image: Thematic Views Second 1/0 private private open Hopperma Image: Thematic Views Second 1/0 private private open Hopperma Image: Thematic Views Second 1/0 private private occlainer Van Utva Image: Thematic Views Second 1/0 private private occlainer Van Utva Van Utva Van Utva Second File Second File Second File File Second File File File File Second <td>🟦 My Workspace 🔶</td> <td>#</td> <td>Name</td> <td>Version</td> <td>Administration s</td> <td>t Registration stat</td> <td>Check</td> <td>Туре</td> <td>Owned by</td>	🟦 My Workspace 🔶	#	Name	Version	Administration s	t Registration stat	Check	Туре	Owned by
Image: Solution of the second seco		4503	absent	1:0	private	private	V	simple	Hopperma
Image: Syntax 5048 access tool 1:0 private private open Hopperma Image: Syntax Image: Syntax 1:0 private private Image: Syntax		3784	access protocol	1:0	private	private	\checkmark	open	Hopperma
Image: Semantic Content Repring 4442 achievement test 1:0 private private # simple Hopperma Image: Semantic Content Repring 4146 actor 1:0 private private ✓ container Vau Uytvate Image: Syntax 4166 actor 1:0 private private ✓ container Vau Uytvate Image: Syntax 3985 adaptation 1:0 private private ✓ container Vau Uytvate Image: Syntax 3985 adaptation 1:0 private private ✓ container Vau Uytvate Image: Syntax 3985 adaptation 1:0 private private ✓ container Vau Uytvate Image: Syntax		5048	access tool	1:0	private	private	V	open	Hopperma
B Semantic Content Repre B Syntax Container 110 private private display=backgroup 4146 actor - 1:0 actor - 1:0 display=backgroup 412 display=backgroup actor - 1:0 display=backgroup actor		4442	achievement test	1:0	private	private	~	simple	Hopperma
B Syntax 4506 actual psychological state 1:0 private private visite Reichel, U B Language Resource Ont 3985 adaptation 1:0 private private simple Wright, Su B Language Codes III III m III IIII IIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	<u> </u>	4146	actor	1:0	private	private	1	container	Van Uytvar
Image: Star Star Stress Image: Star Star Star Star Star Star Star Star		4506	actual psychological state	1:0	private	private	√	closed	Reichel, U
B Lexicography III B Language Codes B Terminology B Terminology B Multilingual Information B Lexical Resources B Lexical Semantics B Lexical Semantics B Sign language B Audio Conver Van Uytvanck, Dieter Scope public B Asis isocat group B Basi isocat group <	🗉 🧾 Language Resource Ont	3985	adaptation	1:0	private	private	1	simple	Wright, Su
B Language Codes B Terminology B Terminology B Multilingual Information B Lexical Resources B Lexical Semantics B Lexical Semantics B Sign language B Sign language B Audio B Scope public Scope B Stass Core B Stass Core or bublic B Scope D LAdministration Information Section B Scope D LAdministration Record	🗉 📃 Lexicography								
B Multilingual Information B Lexical Resources B Lexical Resources B Lexical Semantics B Translation Type container Owner Van Uytvanck, Dieter B Scope Public 1.1 Administration Record	🖽 🔄 Language Codes								
B Lexical Resources Key 4146 B Lexical Semantics PID http://www.isocat.org/datat/DC-4146 B Sign language Owner container B Sign language Scope public B Audio Scope public B Scipa Societ group Athinistration Record 1. Administration Information Section		4.2 🖨	0 🖪 📝 🖌 👧		actor - 1:0	D			
B Lexical Semantics PID H Fit Fit H Sign language Owner Van Uvtvanck, Dieter Scope									
B Translation Type Container Owner Van Uytvanck, Dieter B Judio Scope B Audio Scope B Audio Audio CARIN-NL/VL I.1 Administration Record									
B Sign language Owner Van Uytvanck, Dieter B Audio Scope public B Athens Core public Internation Information Section B Scissional group Internation Record Internation Record					rg/datcat/DC-414	<u>16</u>			
Image: Constraint of the second se									
Image: CLARIN-NL/VL Image: CLARIN-NL/VL Image: CLARIN-NL/VL Image: CLARIN-NL/VL					er				
BAS isocat group 1. Administration Information Section B CLARIN-NL/VL 1.1 Administration Record		Jeope	publi						
CLARIN-NL/VL I.1 Administration Record				1. Adminis	tration Inform	ation Section			
		1 1 Adn	ninistration Record						
				r					
							_		
			let the second second		-	1	1-		
	1					-			
Name Version Administration st Registration stati Cherk Type Owned by Scope	4146 actor	1:0	private private	×	container	Van Uytvanck, Di	public		
Name Version Administration st Registration stati Cherk Type Owned by Scope									

- 1. Welcome menu
- 2. Workspace explorer
- 3. Search box and options
- 4. Main area
 - 4.1. DCS viewer
 - 4.2. DC viewer
- 5. Basket area

		Velcome Guest		🕑 Help *
	8	Log in	Г	
-	8	Shibb Log in	to	ISOcat
	0	Recover	R.	
	8	Register		

- If needed, register (4) with the ISOcat (testbed).
- Log in (²) to ISOcat (testbed) using your own credentials.
- Notice that after logging in you can change the automatic generated password into something easier to remember.
 - \circ You find the Password (\mathscr{P}) menu item in the Welcome menu ().

Hands-on: ISOcat basics

enter keywords here 🛛 🏙 🗳 😒	📥 🗔 🕞
My Workspace Private CGN-playground core WS description DCIF SYIFTAX-standards	Beta version The core of ISOcat would greatly app <u>homepage</u> for vari

- Select 'My Workspace'.
- In the right menu bar there a three buttons to create new groups (
), new selections (
) and new data categories (
).
 Create a new group (
).

a 🚳 👘			BWO Type	logical DB *		
Name		BWO Typological DB			v	
escription		This group works on the	database schem	a for a typological data	abase on basic word or	der.
orum		Contact the ISOcat syste	em administration	to create a forum for	this group.	
	•	Contact the ISOcat syste	em administration	to create a forum for	this group. Organization	
orum Iembers	•					Status ^

- Give your new group a unique name. This name has to be globally unique as when the group makes selections public it will appear in the workspace of every ISOcat user.
- Describe the purpose of the group.
- Add users to the group (+). These users will receive an invitation email and will only have access to the selections and categories shared with the group once they have accepted the invitation.
- Save the new or changed group (🐴).

Hands-on: creating Data Category selections



- Select 'My Workspace'.
- Create a new Data Category Selection (¹/₂).
- Give the selection a name.

	- 6 6 .	BWO db			
#	Name	Version Administration st Registration stati Check Type	Owned by	Scope	A

- The selection is directly loaded into the basket area.
- Use the search (📾) and/or navigation facilities to find matching data categories for the database columns.
 - For the generic columns, e.g., language name and code, one can use the options (*) to restrict the search to the metadata domain:



- You can also restrict the search to name and identifier to find data categories named similar to the database columns
- If you found a proper category it can be added (+) to the basket.
- The following is a possible selection of Data Categories:

		Law second			Contract of the second second second second				Law management	
#		Name	Version	Administration st	Registration state	Check	Туре	Owned by	Scope	
	2482	language ID	1:0	private	private	1	constrained	Athens Core	public	
	2484	language name	1:0	private	private	1	open	Athens Core	public	
•	2534	original source	1:0	private	private	1	open	Athens Core	public	
	3759	source	1:0	private	private	1	open	Nevskava, Irina	public	

• Why might you prefer these categories?

Column	Data Category	Discussion
language name	/language name/ (<u>DC-2484</u>)	The owner of this data category is Athens Core, which is actually a group of people working in the metadata thematic domain. This means the category is reviewed by a number of people and likely to be standardized somewhere in the future. However, it might still undergo chances.
language code	/language ID/ (<u>DC-2482</u>)	Same as for <i>/language name/</i> . Notice that <i>/language ID/</i> doesn't list all valid ISO 639-3 codes (this conceptual domain is too big to be conveniently managed in ISOcat), but it does define a constraint to limit the value domain to 3 lower case letters.
Source	/source/ (<u>DC-3759</u>) /original source/ (<u>DC-2534</u>)	The MDF /source/ is targeted towards people as the source of information. We do have a separate column for informant. If this column is only used for bibliographical references the /original source/ owned by Athens Core.
informant	/source/ (<u>DC-3759</u>)	See the source row above.

• Save () the changes to your Data Category Selection.

source	aa 🐉 🛱	3 00 0	🔒 🔯 🧟 🧟		BWO db	
🗏 🏫 My Workspi	ace	#	Name	Version	Administration st	Registratio
🖻 🚨 Private		24	82 language ID	1:0	private	private
BWC		24	84 language name	1:0	private	private
	-playground WS description	25	34 original source	1:0	private	private
		37	59 source	1:0	private	private

- Load the DCS in the DCS viewer.
- Change the scope (\Re) of the DCS.

Change the scope of this Data Category S	election ×
Groups with read and write access to your/their Da this Data Category Selection	ata Categories in
BWO Typological DB	^
CGN-playground	
CLARIN Web Services	
CLARIN-NL/VL	E
GilAndSusan	
RELISH	
STTS	
TDS Curator	-
T VU-DNC	
Set scope to <i>public</i> (everyone has read access))
Make all your DCs in this DCS also public	
🔀 Cancel	< Change

- Mark the checkbox for the group you've created before.
- Save the changes (). Now all group members, who have accepted the invitation, will have access to this DCS.

Hands-on: creating Data Category specifications



- Select 'My Workspace'.
- Create a new Data Category (
).

subjectObjectVerb - 1:0 *		_ <i>8</i> ×
🖻 🖌		
Administrative Information Sec	tion	
Identification Identifier Version Type Justification	subjectObjectVerb 110 simple Needed for a typological database on basic word order.	
Origin		E
Administration status Registration status	private	
Dates		
Effective date Until date	yyyy-MM-dd	
Additional information		
	E. F.	_
🔀 Cancel		Next
Description Section		
Conceptual Domains and Linguis	stic Sections	

- Enter a camel case identifier.
- Choose the proper type.
- Enter a justification, just in case you want to standardize the Data Category later.

×					
ministrative Information Sect	ion				
scription Section					
Data Element Name	Name	•	Source	- "	
	= sov		BWO typological databas		
				Ŧ	
Language Sections					
English *			Lar	nguage	٠
Language section					T
Language section	ne	Stat	us	<u>∽</u> ♣	-
Name sections • Nar	ne bject object verb		us nitted name	^ (
Name sections • Nar				~ ት	
Name sections • Nar				 中 中 	
Name sections • Nar sub Definitions • Def	oject object verb	adn	Note	*	-
Name sections • Nar sub Definitions • Def	inition	Source	Note	× *	N

- The abbreviation we use in our database we can enter as a Data Element Name.
- The English language section is mandatory and should contain at least one name and exactly one definition.
- Save () the new Data Category.

Save a DC	×
Jiso cat	
Change description	
Created SQV word order value,	
This Data Category will now be created. Please enter a description of the changes you have made. Notice, that this change description will be a standard noti of the DCIE export of the DCIE category.	4 III +
X Cancel	

- Enter a meaningful change description and save (
 .
 .
- Agree to add the new Data Category to the basket which contains the DCS we're filling.
 Don't forget to save (¹) the changes.
- Same can be done for the other word order values and for the basic word order closed data category.

								_	Ð
ministrative Information Section	n								
scription Section									
ceptual Domains and Linguistic	Sections								
Conceptual Domain									
Data type	• strin	g 🔻							
	0 V-1							_ ↓	
Value domain	Valu	e						~ ~	
									-
									-
	Data Ca	tegory Selector					×		
			Query isy, p				8 2		
Linguistic Sections	1.	Name	Query jsx_p	Administratio	on st Registration s				
Linguistic Sections	#	Name	Version		on st Registration s	tati Cher	k Typ *		_
Linguistic Sections	4624	subject Complement		private	private	tati Cher		Language	
	4624 4187	subject Complement	Version 1:0	private private	private private	tati Cher	sim	Language '	
	4624 4187 5141	subject Complement subject Control subject object verb	Version 1:0 1:0	private	private	tati Cher	k Tvc * sim sim	Language	
Language *	4624 4187 5141 4188	subject Complement subject Control subject object verb subject Raising	* Version 1:0 1:0 1:0	private private private	private private private	tati Cher	k Tyc ^ sim sim sim		•
Language *	4624 4187 5141 4188 368	subject Complement subject Control subject object verb subject Raising subjunctive	Version 1:0 1:0 1:0 1:0	private private private private	private private private private	tati Cher	sim sim sim sim	Language *	
Language *	4624 4187 5141 4188 368 1843	subject Complement subject Control subject object verb subject Raising	Version 1:0 1:0 1:0 1:0 1:0 1:0	private private private private private	private private private private private	tati Cher V V V A	sim sim sim sim sim		•
Language *	4624 4187 5141 4188 368 1843 3493	subject Complement subject Control subject object verb subject Raising subjunctive subjunctive	* Version 1:0 1:0 1:0 1:0 1:0 1:0 1:0	private private private private private private	private private private private private private	tati Cher V V V A	k Tyc ^ sim sim sim sim sim		
Linguistic section	4624 4187 5141 4188 368 1843 3493 3494	subject Complement subject Control subject object verb subject Raising subjunctive subjunctive SubjunctiveMood	* Version 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0	private private private private private private	private private private private private private private	tati Cher V V V A Cher V V V V V V V V V V V V V	k Tyr ^ sim sim sim sim sim sim		
Language *	4624 4187 5141 4188 368 1843 3493 3494 1392	subject Complement subject Control subject object verb subject Raising subjunctive subjunctive SubjunctiveMood SubjanctiveMood	* Version 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0	private private private private private private private private	private private private private private private private	tati Cher	k Tyr ^ sim sim sim sim sim sim sim		
Language *	4624 4187 5141 4188 368 1843 3493 3494 1392	subject Complement subject Control subject object verb subject Raising subjunctive subjunctive Subjunctive Subjunctive subjunctive subjunctive subjunctive subjative case	Version 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0 1:0	private private private private private private private private private	private private private private private private private private	tati Cher V V V A Cher V V V V V V V V V V V V V	k Tvc ^ sim sim sim sim sim sim sim		

Add (⁺/₊ and *√*) the simple Data Categories, like /subject object verb/, to the value domain of the closed complex Data Category /basic word order/.

Take home: beyond ISOcat - data category references

A database schema is defined using the Data Definition Language (DDL) part of SQL. Using @dcr:datcat annotations inside comments the DDL statements can be annotated:

```
-- @dcr:datcat 'sov' http://www.isocat.org/datcat/DC-...
-- @dcr:datcat 'svo' http://www.isocat.org/datcat/DC-...
-- @dcr:datcat 'osv' http://www.isocat.org/datcat/DC-...
-- @dcr:datcat 'ovs' http://www.isocat.org/datcat/DC-...
-- @dcr:datcat 'vos' http://www.isocat.org/datcat/DC-...
-- @dcr:datcat 'vso' http://www.isocat.org/datcat/DC-...
CREATE TYPE bwo AS ENUM (
    'sov',
    'svo'.
    'osv',
    'ovs',
    'vos',
    'vso'
);
-- @dcr:datcat code http://www.isocat.org/datcat/DC-2482
-- @dcr:datcat name http://www.isocat.org/datcat/DC-2484
-- @dcr:datcat bwo http://www.isocat.org/datcat/DC-...
-- @dcr:datcat source http://www.isocat.org/datcat/DC-2534
-- @dcr:datcat informant http://www.isocat.org/datcat/DC-3759
CREATE TABLE "typ-word-order" (
   code character varying(3) NOT NULL,
    name character varying NOT NULL,
   bwo bwo,
    source character varving,
    informant character varying
```

);

Take home: beyond ISOcat – ontological relationships

verb-object-subject (vos) and verb-subject-object (vso) can be seen as subclasses of predicate-initial:

/verb-object-subject/rel:subClassOf/predicate-initial/

/verb-subject-object/ rel:subClassOf /predicate-initial/

In the ontology of the Typological Database System (TDS) project these kind of theoretical perspectives were both included but kept separate in their own parts of the subsumption hierarchy. In the TDS different theories describing the same linguistic phenomena are linked using the special tds:equatesWith predicateⁱ. Depending on your preference the same can be stored in the Relation Registry RELcat:

/verb-object-subject/tds:equatesWith/predicate-initial/

```
/verb-subject-object/ tds:equatesWith/predicate-initial/
```

However to make it possible for generic semantic search algorithms to make use of this relationship tds:equatesWith has to be embedded in the taxonomy of relation types:

1. rel:related

- 1.1. rel:almostSameAs
 - 1.1.1. tds:equatesWith

¹ A. Dimitriadis, A. Saulwick, M. Windhouwer. Semantic relations in ontology mediated linguistic data integration. In Proceedings of the E-MELD Workshop on Morphosyntactic Annotation and Terminology: Linguistic Ontologies and Data Categories for Linguistic Resources (E-MELD 2005), Cambridge, Massachusetts, July 1-3, 2005. <u>http://www.emeld.org/workshop/2005/papers/saulwick-paper.html</u>