**TeLeMaCo**

a collaborative repository for training and teaching materials

Elke Teich, Hannah Kermes, Jörg Knappen, José Manuel Martínez Martínez, Mihaela Vela, Marilisa Amoia
collaborative portal and repository for training and teaching materials
implemented as web service
different media types: text, presentations, videos
technical documentation, learning and teaching material
categorization allows for fine-grained structures search
semantic tagging of content
user-defined ranking criteria

https://fedora.clarin-d.uni-saarland.de/hub/
Digital repositories for language resources

- digital repositories for linguistic resources
  - The Language Archive (TLA)
  - the Virtual Language Observatory (VLO)
- WebLicht: Web-Based Linguistic Chaining Tool
- Federated Content Search (FCS) to query linguistic resources
- but how can we find relevant training and learning materials
Why not use google?

- google does not index everything which is on the web
- keyword search but no categorization ⇒ no fine-grained search
- ambiguity of terms
  - *Habeas corpus*, a legal mechanism to end detention of a suspect
  - *Corpus delicti*, a legal term meaning "body of the crime"
  - *Corpus callosum*, a structure in the brain
  - *Corpus cavernosum (disambiguation)*, a pair of structures in human genitals
1. **corpus linguistics:**
   definition (2), teaching/learning material (1), introductions (text) (3), institute (1), journal (1), article (1)

2. **corpus linguistics teaching:**
   article (text; 4) (slides; 1), website with links and references (1), course description (1), conference website (1), research group (1)

3. **corpus linguistics teaching material:**
   articles (text; 9) (slides; 1)

4. **corpus linguistics introduction:**
   text (3), teaching/course material (3), book (3), course description (1)
Metadata annotation

- bibliographic features: author, title, publication year, URL
- pedagogically relevant features: prerequisites, target audience, objectives, level
- topic-relevant features: description
- technical features: language, media, type of the resource
- accessibility information: license, access
- bibliography facilities: BibTeX, BibTeX type
Media types and types of resources

- media types: text, presentations, audio/video
- types of resources:
  - technical documentation: tutorials, how-to’s, documentations
  - learning material for self study: tutorials, exercises, videos of courses, etc.
  - teaching material: short modules (2-4 hours), full courses (2-4 SWS), material for training courses
  - show-cases: example projects using language resources and tools
Semantic tagging

- user feedback refers to these categories
- Was this resource relevant for you?
- What level has the resource in your opinion?
- Did you reach the objectives given?
- Were the prerequisites appropriate?
Describe a new resource

Title: 
Keywords: 
Author: 
Institution: 
Year of publication: 
BibTeX type: misc Language: 
Objective: 
Audience: 
Prerequisites: 
Level: 
Type: 
Media: 
URL: 
Licence: 
Access: open academic
Description:
The seminar gives an introduction to corpus linguistic methods and techniques. Corpus linguistics offers a broad spectrum of methodologies for empirically oriented questions in linguistics and translatology. In this course, we will get
Advanced search mechanism

- the annotated metadata allows for a fine-grained search
- the different categories allow to specify the query more precisely, e.g., in terms of:
  - difficulty level
  - target audience
  - prerequisites
  - type of resource
- an example:
  - topic: corpus linguistics; difficulty level: beginner; type of resource: tutorial for self study
  - topic: corpus linguistics; difficulty level: beginner; type of resources: teaching material
Advanced search results

Query results

- Corpus-based linguistics and translation

Your advanced search query

Title: 
Keywords: corpus linguistics
Author: 
Language: 
Objective: 
Audience: 
Prerequisites: 
Level: basic
Type: teaching material
Media: 
URL: 
Licence: 
Corpus-based linguistics and translotation

* Elke Teich * Stefania Degaetano-Ortlieb * Hannah Kermes * Ekaterina Lapshinova-Koltunski *

Keywords: corpus linguistics, translotation

http://fr46.uni-saarland.de/lasteich/SS2012_HS-CL/program.php

The seminar gives an introduction to corpus linguistic methods and techniques. Corpus linguistics offers a broad spectrum of methodologies for empirically oriented questions in linguistics and translotation. In this course, we will get to know a selection of these methods (corpus compilation, corpus annotation, corpus analysis). The application of these methods will be illustrated with selected research questions (e.g. collocation, register analysis, translation related text analysis). In the tutorial we learn how to apply and use different corpus tools (concordancers, annotation tools, corpus query) and how to use them for research.

Feedback

Sorry, there is no feedback available. Be the first one to provide feedback!

Give feedback
Resource details

Institution: Universität des Saarlandes
Year of publication: 2012
Language: english
Type: teaching material
Audience: linguists, translators
Level: basic
Prerequisites: linguistics
Media: pdf slides
Objective: basic knowledge about corpus compilation, corpus annotation, corpus analysis
Licence:
Access: academic
Creation date: Wednesday, 14 November 2012 11:29:01
Last modified: Wednesday, 14 November 2012 11:29:01
BibTeX type: @misc
BibTeX entry:

@misc(TeLeMaCo:167,
author = "Teich, Elke",
title = "Corpus-based linguistics and translationology",
year = "2012",
url = "http://fr46.uni-saarland.de/lsteich/SS2012\_HS-CL\_program.php"
)
Future work

- automatic generation of keywords from the given metadata
- upload mechanism for resources, which are not yet accessible on the internet
- a collaborative repository needs input
The Saarbrücken CLARIN-D Team:

Elke Teich
Hannah Kermes
Jörg Knappen
José Manuel Martínez Martínez
Mihaela Vela
Marilisa Amoia

https://fedora.clarin-d.uni-saarland.de/hub/