


# Anotação de dependências *enhanced* - CONLLuEditor

[Sitemap](#) [Open source](#) [Orange in the communities](#) [News](#) [Orange on GitHub](#) [Français](#)

 **Orange Open Source**  
free and open source software

[Mobile software](#) [Home software](#) [Personal software](#) [Professional software](#)

[Orange Open Source](#) > [Languages](#) > [JavaScript](#) > [ConllEditor: editor for syntactic annotations in text](#)

## ConllEditor: editor for syntactic annotations in text

ConllEditor is a tool which facilitates the editing of syntactic relations and morphological features of files in CoNLL-U format ([universaldependencies.org](http://universaldependencies.org)). It uses a Java-based server and a HTML/CSS/JavaScript based front-end. The editor loads the CoNLL-U file and saves every change to disk (and performs a git commit if the file is under git version control).

CoNLL-U is a Stanford format used to annotate texts in syntax (notably dependency trees). The annotated texts can be used for machine learning tasks in Natural Language Processing. The tool permits to edit easily the syntactic tree structures in a web browser.

The editor provides the following functionalities:








- editing words (forms, lemmas, upos, xpos, features, enhanced dependencies)
- editing dependency relations
- join/split words (to correct tokenization errors)
- join/split sentences (to correct tokenization errors)
- undo/redo (partially)
- git support







This software is under the 3-Clause BSD License. ConllEditor is now available on [Orange-OpenSource GitHub](#).


### Recent Posts

- GPSGhoster 1 February 2023
- With Sylla, Telco Cloud is on the path of open source 8 December 2022
- Autocoder based Anomaly detection 23 March 2022
- Pyngali 23 March 2022
- its-client 15 March 2022

filename: /Users/ariadna/cafe.conllu changes: 0

  read sentence 1     features misc flat graph  right to left save changes help simple search 

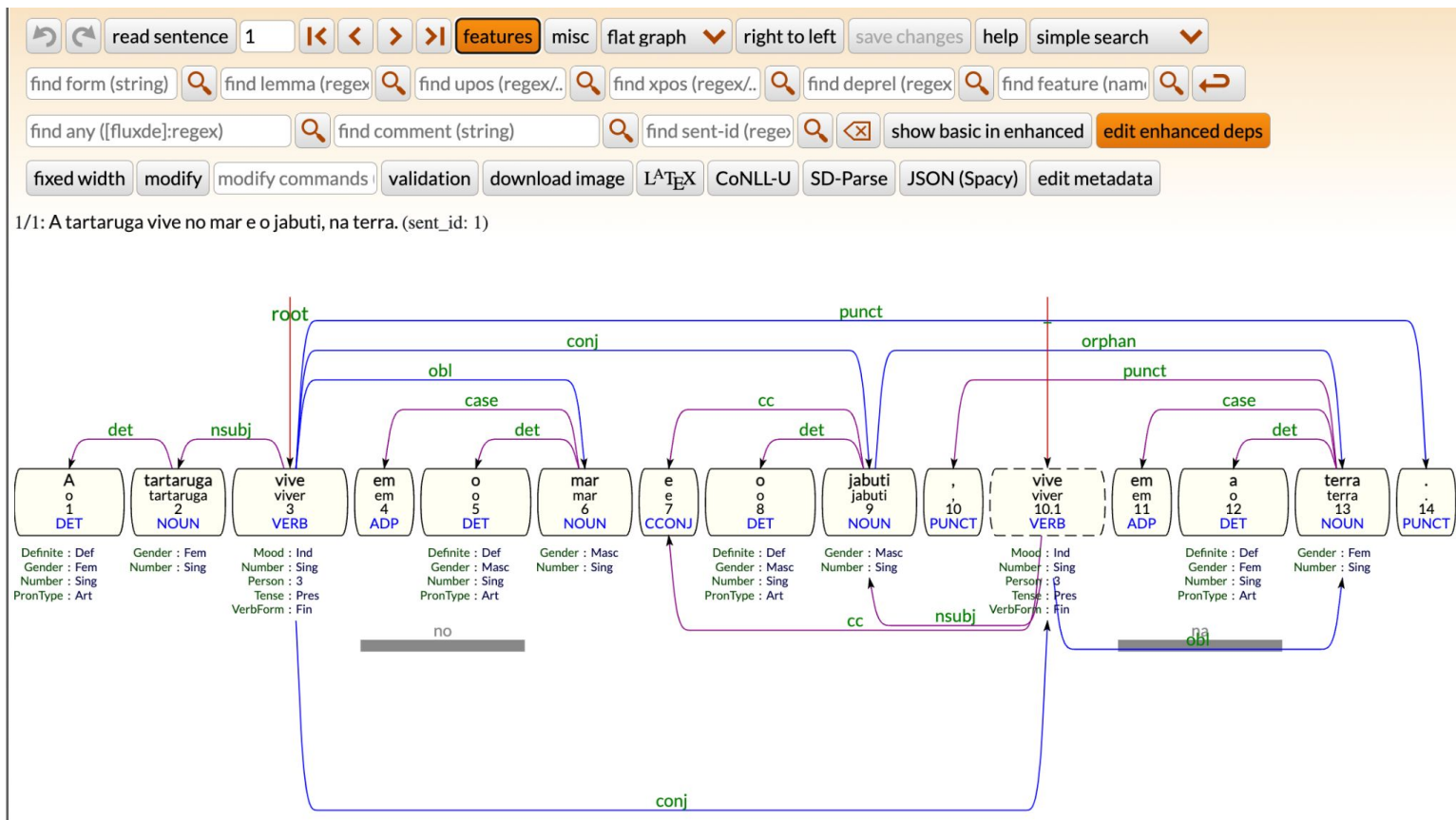
find form (string)  find lemma (regex)  find upos (regex/..  find xpos (regex/..  find deprel (regex)  find feature (name)  

find any ([fluxde]:regex)  find comment (string)  find sent-id (regex)   show basic in enhanced edit enhanced deps


fixed width modify modify commands validation download image L<sup>A</sup>T<sub>E</sub>X CoNLL-U SD-Parse JSON (Spacy) edit metadata



# Dependências *enhanced* - inserção de deprels



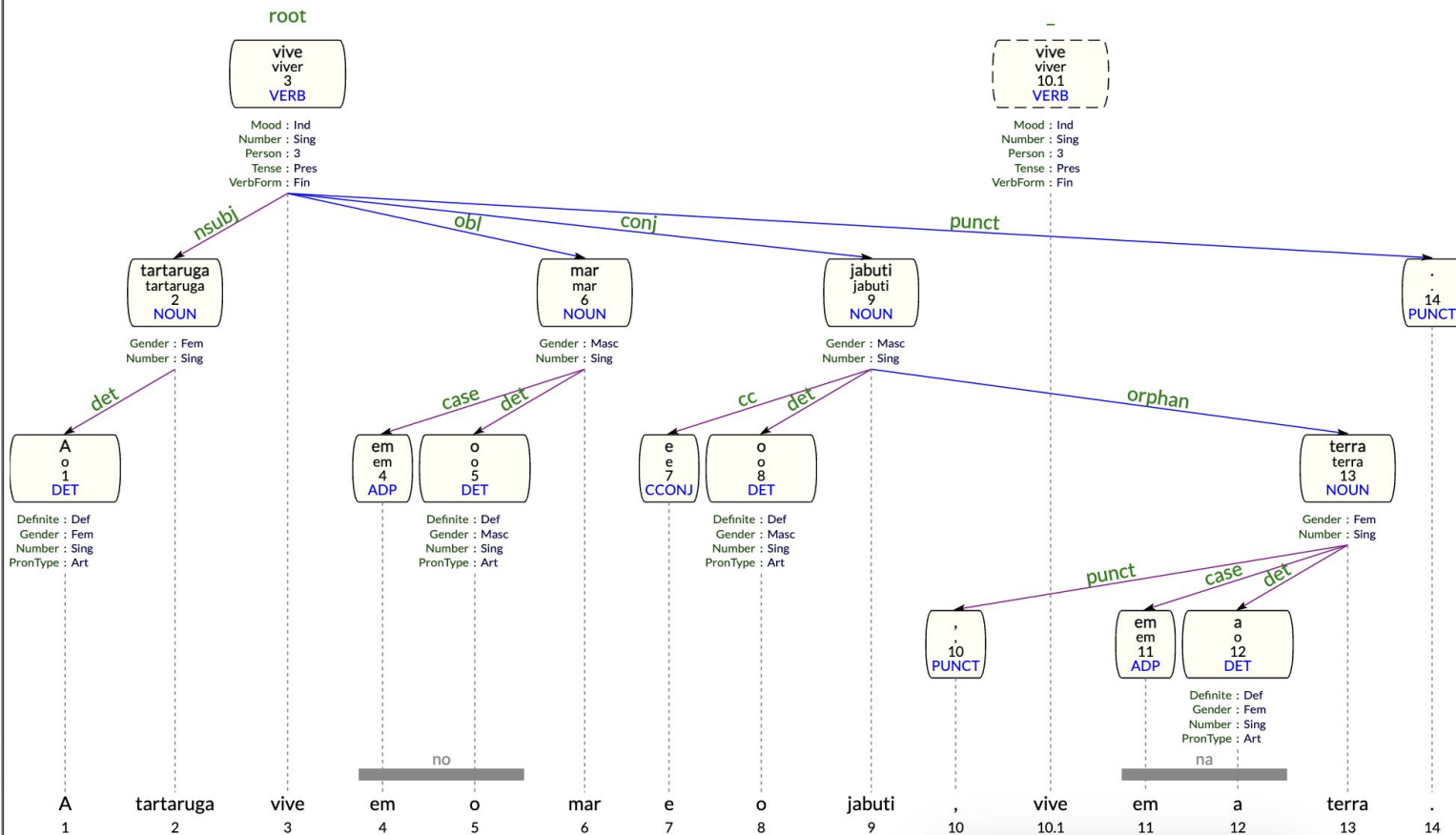
# Dependências *enhanced* - CoNLL-U



ID	FORM	LEMMA	UPOS	XPOS	FEATS	HEAD	DEPREL	DEPS	MISC
1	A	o	DET	_	Definite=Def Gender=Fem Numbe	2	det	_	_
2	tartaruga	tartaruga	NOUN	_	Gender=Fem Number=Sing	3	nsubj	_	_
3	vive	viver	VERB	_	Mood=Ind Number=Sing Person=3	0	root	_	_
4-5	no	-	-	-	-	-	-	-	-
4	em	em	ADP	_	_	6	case	_	_
5	o	o	DET	_	Definite=Def Gender=Masc Numb	6	det	_	_
6	mar	mar	NOUN	_	Gender=Masc Number=Sing	3	obl	_	_
7	e	e	CCONJ	_	_	9	cc	10.1:cc	_
8	o	o	DET	_	Definite=Def Gender=Masc Numb	9	det	_	_
9	jabuti	jabuti	NOUN	_	Gender=Masc Number=Sing	3	conj	10.1:nsubj	SpaceAfter=No
10	,	,	PUNCT	_	_	13	punct	_	_
10.1	vive	viver	VERB	_	Mood=Ind Number=Sing Person=3	0	_	3:conj	_
11-12	na	-	-	-	-	-	-	-	-
11	em	em	ADP	_	_	13	case	_	_
12	a	o	DET	_	Definite=Def Gender=Fem Numbe	13	det	_	_
13	terra	terra	NOUN	_	Gender=Fem Number=Sing	9	orphan	10.1:obl	SpaceAfter=No
14	.	.	PUNCT	_	_	3	punct	_	_

# Dependências enhanced - árvore

1/1: A tartaruga vive no mar e o jabuti, na terra. (sent\_id: 1)



## SD-Parser

A tartaruga vive em o mar e o jabuti , em a terra .

det(tartaruga, A)

nsubj(vive, tartaruga)

case(mar, em)

det(mar, o)

obl(vive, mar)

cc(jabuti, e)

det(jabuti, o)

conj(vive, jabuti)

punct(terra, ,)

case(terra, em)

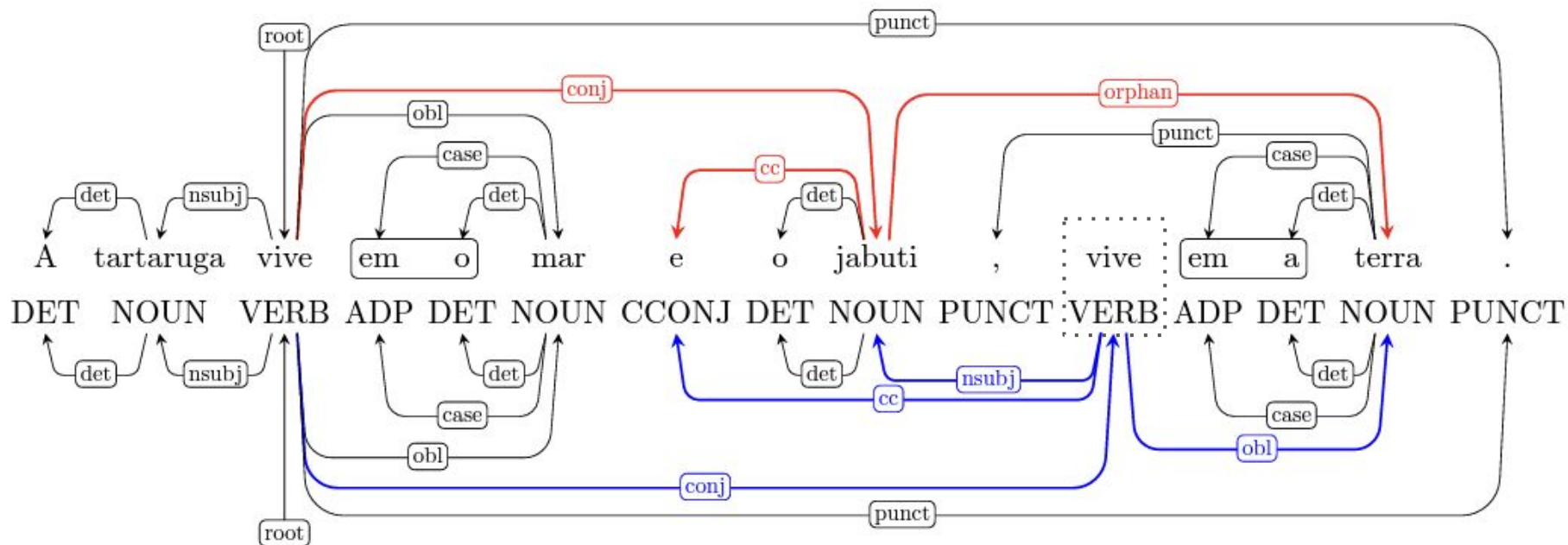
det(terra, a)

orphan(jabuti, terra)

punct(vive, .)



# Visualização por meio do LATEX



\usepackage{tikz-dependency}



# Visualização de dependências *enhanced* - **greW**

: # newpar

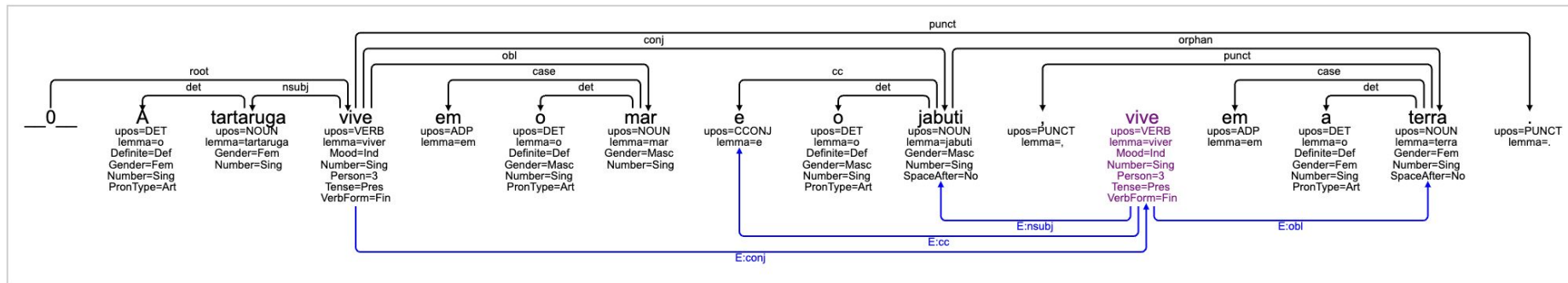
sent\_id: 1

text: A tartaruga vive no mar e o jabuti, na terra.

generator: UDPipe 2, <https://lindat.mff.cuni.cz/services/udpipe>

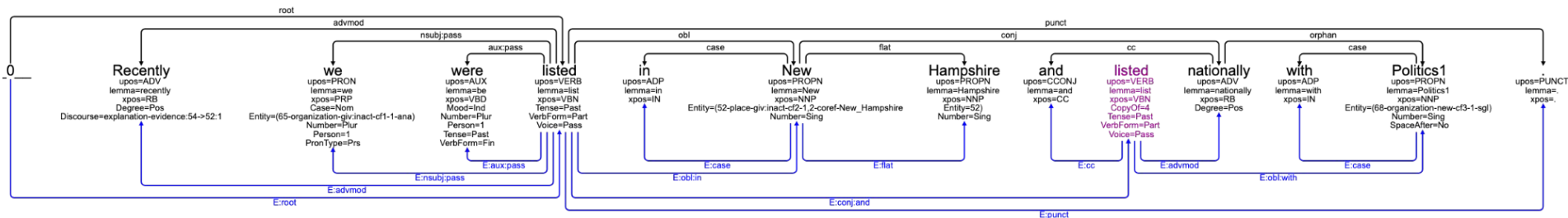
udpipe\_model: portuguese-bosque-ud-2.10-220711

udpipe\_model\_licence: CC BY-NC-SA



<https://web.grew.fr/>

# UD\_English-GUM



# UD\_English-GUM

# sent\_id = GUM\_interview\_hill-30

# s\_prominence = 3

# s\_type = decl

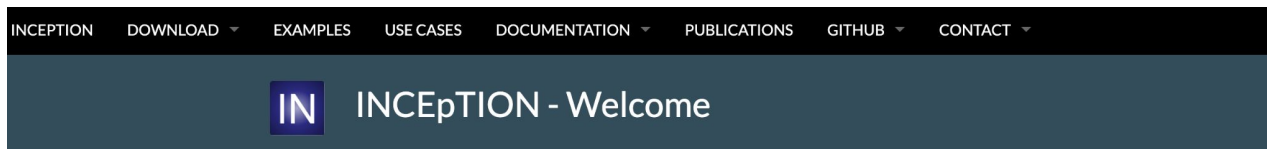
# speaker = ChristopherHill

# transition = zero

# text = Recently we were listed in New Hampshire and nationally with Politics1.

1	Recently	recently	ADV	RB	Degree=Pos	4	advmod	4:advmod	Discourse=explanation-evidence:54->52:1
2	we	we	PRON	PRP	Case=Nom   Number=Plur   Person=1   PronType=Prs	4	nsubj:pass	4:nsubj:pass	Entity=(65-organization-giv:inact-cf1-1-ana)
3	were	be	AUX	VBD	Mood=Ind   Number=Plur   Person=1   Tense=Past   VerbForm=Fin	4	aux:pass	4:aux:pass	_
4	listed	list	VERB	VBN	Tense=Past   VerbForm=Part   Voice=Pass	0	root	0:root	_
5	in	in	ADP	IN	_	6	case	6:case	_
6	New	New	PROPN	NNP	Number=Sing	4	obl	4:obl:in	Entity=(52-place-giv:inact-cf2-1,2-coref-New_Hampshire
7	Hampshire	Hampshire	PROPN	NNP	Number=Sing	6	flat	6:flat	Entity=52)
8	and	and	CCONJ	CC	_	9	cc	8.1:cc	_
8.1	listed	list	VERB	VBN	Tense=Past   VerbForm=Part   Voice=Pass	_	_	4:conj:and	CopyOf=4
9	nationally	nationally	ADV	RB	Degree=Pos	6	conj	8.1:advmod	_
10	with	with	ADP	IN	_	11	case	11:case	_
11	Politics1	Politics1	PROPN	NNP	Number=Sing	9	orphan	8.1:obl:with	Entity=(68-organization-new-cf3-1-sgl)   SpaceAfter=No
12	.	.	PUNCT	.	_	4	punct	4:punct	_

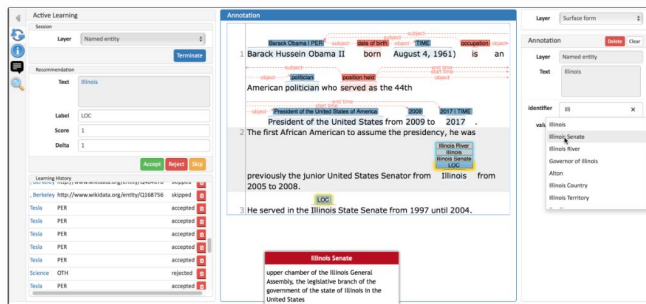
# Anotação de dependências *enhanced* - INCEpTION



A semantic annotation platform offering intelligent assistance and knowledge management

The annotation of specific semantic phenomena often require compiling task-specific corpora and creating or extending task-specific knowledge bases. Presently, researchers require a broad range of skills and tools to address such semantic annotation tasks.

In the recently funded INCEpTION project, UKP Lab at TU Darmstadt aims towards building an annotation platform that incorporates all the related tasks into a joint web-based platform.



How to cite

If you use INCEpTION in scientific work or want to learn more about it, please have a look at



Download  
**INCEpTION 28.2**  
(Released on 2023-06-06)



Try  
**INCEpTION**  
online



Check out our latest publication  
**From Zero to Hero: Human-In-The-Loop Entity  
Linking in Low Resource Domains**  
at [ACL 2020](#)



Details



Documents



Users



Workload



Layers



Annotation



Curation



Knowledge Bases



Recommenders



Tagsets



Document Repositories



CAS Doctor

## uds

## Layers ?

Import

Create

Chunk  
Coreference  
Dependency  
Lemma  
Morphological features  
Named entity  
Orthography Correction  
Part of speech  
Referentes  
SemArg  
SemPred  
Surface form

## Layer Details

Export

Save

Cancel

## Properties ?

Name

Dependency

Description

☒ Enabled

## Technical Properties ?

Internal Name

de.tudarmstadt.ukp.dkpro.core.a  
pi.syntax.type.dependency.Depe  
ndency

Type

Relation

Attach to layer

Token :: pos -> Part of  
speech

## Behaviors ?

☐ Read-only☒ Show arcs

## Features ?

↑ ↓ Create

Flavor : [String]

Relation : [String]

## Feature Details

Save

Cancel

Options

- ☒ Enabled
- ☐ Required
- ☒ Visible
- ☐ Remember
- ☒ Curatable
- ☐ Hide when no constraints apply
- ☐ Multiple rows

Tagset

Dependency flavors

Editor type

Auto (depending on tagset s

Key bindings

basic

enhanced



Details



Documents



Users



Workload



Layers



Annotation



Curation



Knowledge Bases



Recommenders



Tagsets



Document Repositories



CAS Doctor

## uds

## Tagsets

Import

Create

Coreference mentions  
Coreference relations  
Dependency flavors  
Named Entity tags  
Operation  
UD Universal Dependencies (v2)  
UD Universal POS tags (v2)

## Tagset Details

Export

Save

Cancel



## Name

UD Universal Dependencies (v2)

## Language

mul

## Description

The idea of universal dependencies is to propose a set of universal grammatical relations which can be used with relative fidelity to capture any dependency relation between words

☐ Annotators may add new tags

## Tags

Create

acl  
advcl  
advmod  
amod  
appos  
aux  
case  
cc  
ccomp  
clf  
compound  
conj  
cop  
csubj  
dep  
det  
discourse  
dislocated

# Anotação de dependências *enhanced* - INCEpTION

INCEpTION Projects Dashboard

Help Administration Adriana Log out 26 mi

Annotations

Group by label

Sort by score

Recommendations first

enhanced | advcl

→ descansar por 20 mir

enhanced | conj

→ Sove a massa e deixe

enhanced | nsubj

→ descansar por 20 mir

→ dobrar de volume.

enhanced | obj

→ deixe descansar por

uds receita.conllu

1-1 / 1 sentences [doc 2 / 5]

1 Sove a massa e deixe descansar por 20 minutos ou até dobrar de volume.

Diagram illustrating dependency annotations for the sentence: "Sove a massa e deixe descansar por 20 minutos ou até dobrar de volume."

Basic dependencies (labeled "basic |"): root, obj, conj, det, cc, cc, xcomp, case, nsubj, advcl, obl, case, nummod, conj, cc, mark, obl, case, punct.

Enhanced dependencies (labeled "enhanced |"): punct, nsubj, conj, nsubj, obj, cc, xcomp, advcl, case, nsubj, advcl, obl, case, nummod, conj, cc, mark, obl, case, punct.

Layer: Referentes

Annotation: Delete Reverse Clear

Layer: Dependency

From: descansar

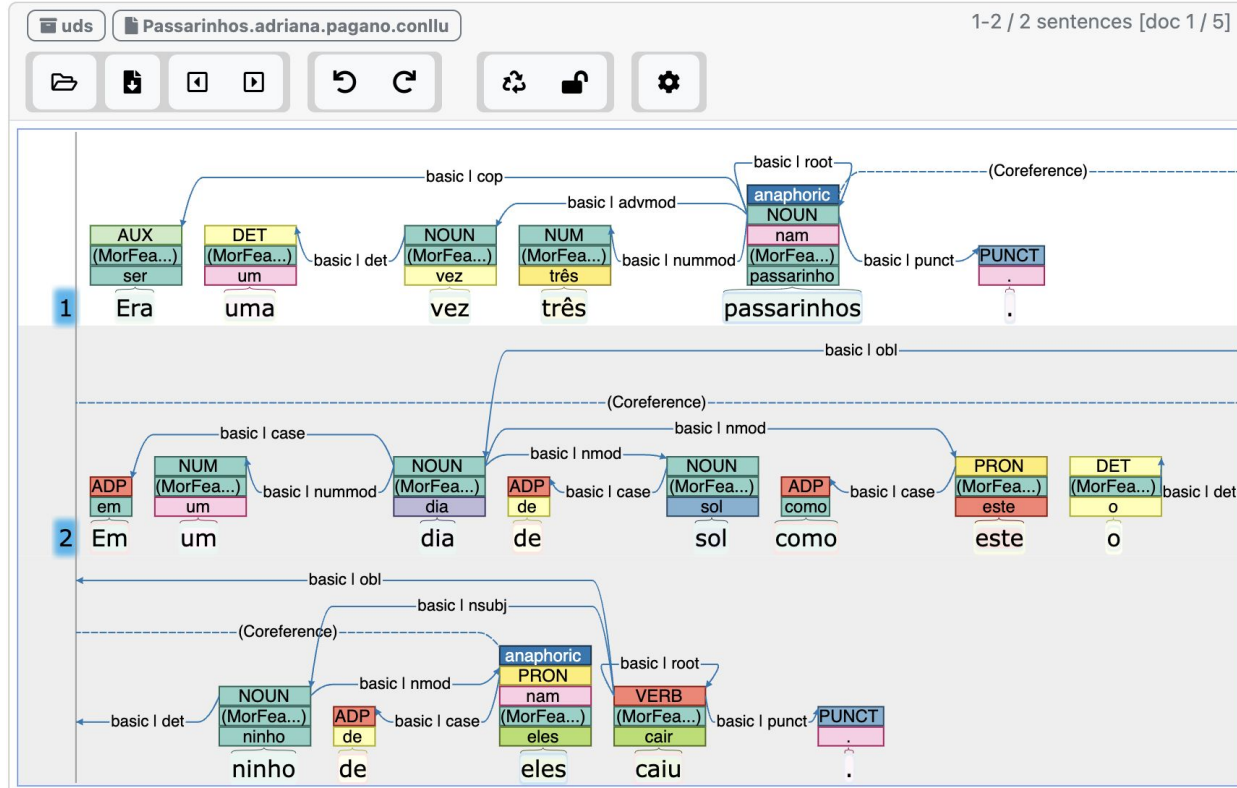
To: massa

Flavor: basic enhanced

Relation: nsubj



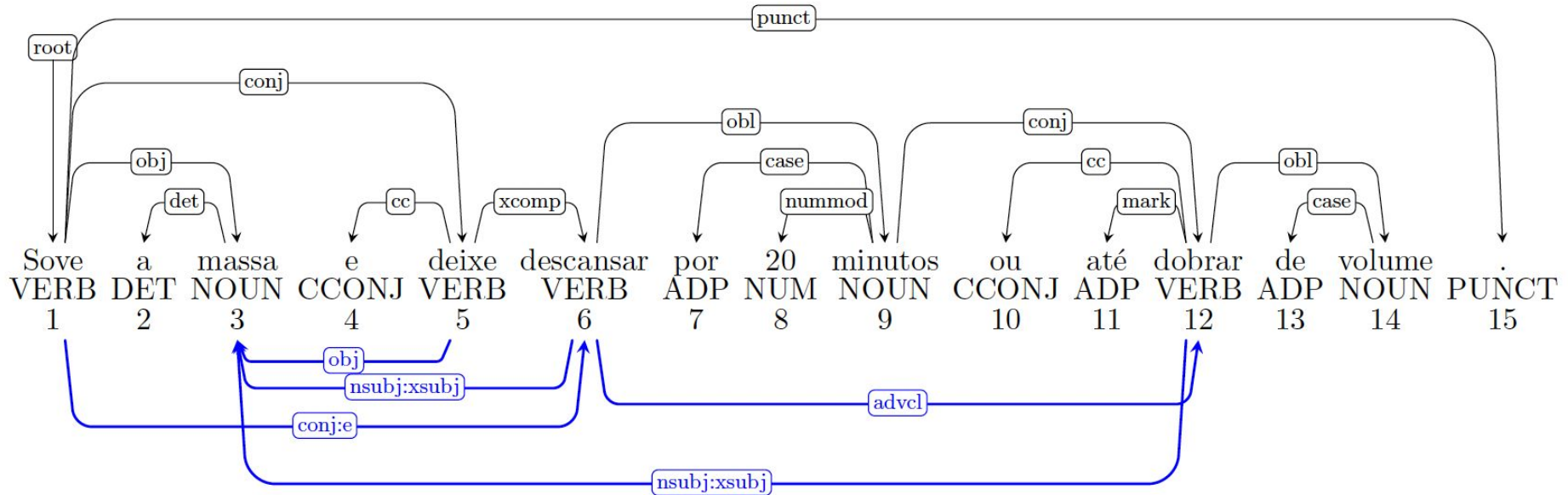
# Anotação de dependências + correferência



# Exportação de dados

1-1	0-4	Sove	*	*	*	*	*	*	Imp	*	*	Sing	2	*	*	*	*	*	Mood=Imp\ N umber=Sing\  Person=2	*	*	*	VERB	sovar	root	basic	1-1	
1-2	5-6	a	*	*	*	*	*	Fem	*	*	*	Sing	*	*	Art	*	*	*	Definite=Def\  Gender=Fem\  Number=Sing\  PronType=Ar t	*	*	*	DET	o	det	basic	1-3	
1-3	7-12	massa	*	*	*	*	*	Fem	*	*	*	Sing	*	*	*	*	*	*	Gender=Fem\  Number=Sing	*	*	*	NOUN	massa	obj ns ubj ns ubj ob j	basic  enhan ced e nhanc ed en hance d	1-1 1- 6 1- 12 1-5	
1-4	13-14	e	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	CCONJ	e	cc	basic	1-5		
1-5	15-20	deixe	*	*	*	*	*	*	Imp	*	*	Sing	2	*	*	*	*	*	Mood=Imp\ N umber=Sing\  Person=2	*	*	*	VERB	deixar	conj	basic	1-1	
1-6	21-30	descan sar	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	VerbForm=Inf	Inf	*	*	VERB	descan sar	xcomp  conj	basic  enhan ced	1-5 1- 1	
1-7	31-34	por	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	ADP	por	case	basic	1-9		
1-8	35-37	20	*	*	*	*	*	*	*	*	Card	*	*	*	*	*	*	*	NumType=Car d	*	*	*	NUM	20	numm od	basic	1-9	
1-9	38-45	minut os	*	*	*	*	*	Masc	*	*	*	Plur	*	*	*	*	*	*	Gender=Masc\  Number=Plur	*	*	*	NOUN	minut o	obl	basic	1-6	
1-10	46-48	ou	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	CCONJ	ou	cc	basic	1-12		
1-11	49-52	até	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	ADP	até	mark	basic	1-12		
1-12	53-59	dobrar	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	VerbForm=Inf	Inf	*	*	VERB	dobrar	conj a dvcl	basic  enhan ced	1-9 1- 6	
1-13	60-62	de	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	ADP	de	case	basic	1-14		
1-14	63-69	volum e	*	*	*	*	*	Masc	*	*	*	Sing	*	*	*	*	*	*	Gender=Masc\  Number=Sing	*	*	*	NOUN	volum e	obl	basic	1-12	
1-15	69-70	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	PUNCT	.	punct	basic	1-1		

# Visualização no LaTeX



# Visualização /Anotação de dependências

Ferramenta	Visualização	Anotação
Arborator Grew	não	não
Annotatrix	não	não
CoNLL-U Viewer	não	não
Grew web	sim	não
ConLLuEditor	sim	sim
INCEpTION	sim (parcial) erro ao importar CoNLLU com token adicionado por causa de ellipse	sim
Annodoc <a href="https://spyysalo.github.io/annodoc/#what-is-annodoc">https://spyysalo.github.io/annodoc/#what-is-annodoc</a>	sim	sim

# Anotação de dependências *enhanced* - elipse

At the moment, working with zero-width spans is the only option for ellipsis annotation. Zero-width spans are not attached to tokens. The UI only allows creating zero-width spans at the beginning or end of a token or even within a token. You cannot create an annotation on the white space between tokens. But that doesn't imply that the span is attached to the token. That said, a zero-width span does not behave like a token either. Currently, a token is an invisible annotation type and a few other annotations like POS, Lemma and a few others internally hook up to the token. The UD importer/exporter knows how to work with these internal Token, POS or Lemma types. But if I remember correctly, the UD importer/exporter currently does not support ellipsis. You can build custom annotation layers which allow

zero-width annotations and also define relations over them and such, but you can then only

export the data as UIMA CAS XMI or WebAnno TSV and a conversion to/from CoNLL-U would need to happen externally e.g. using a Python script with the help of the DKPro Cassis library for XMI files.

# Anotação de dependências *enhanced* - elipse

## *Zero-width spans*

To create a zero-length annotation, hold `Shift` and click on the position where you wish to create the annotation. To avoid accidental creations of zero-length annotations, a simple single-click triggers no action by default. The **lock to token** behavior cancels the ability to create zero-length annotations.



A zero-width span between two tokens that are directly adjacent, e.g. the full stop at the end of a sentence and the token before it ( `end.` ) is always considered to be **at the end of the first token** rather than at the beginning of the next token. So an annotation between `d` and `.` in this example would be rendered at the right side of `end` rather than at the left side of `.`.