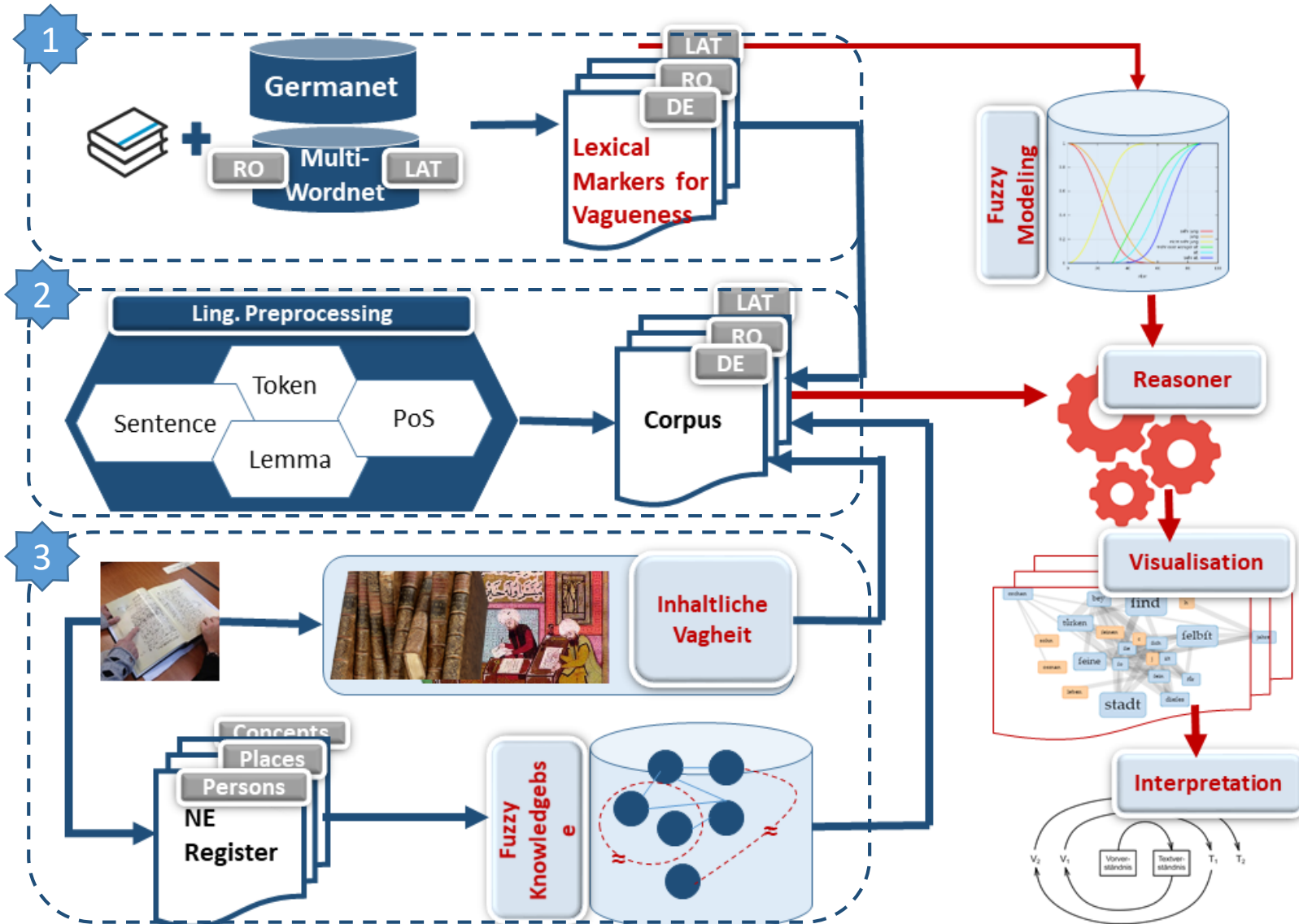


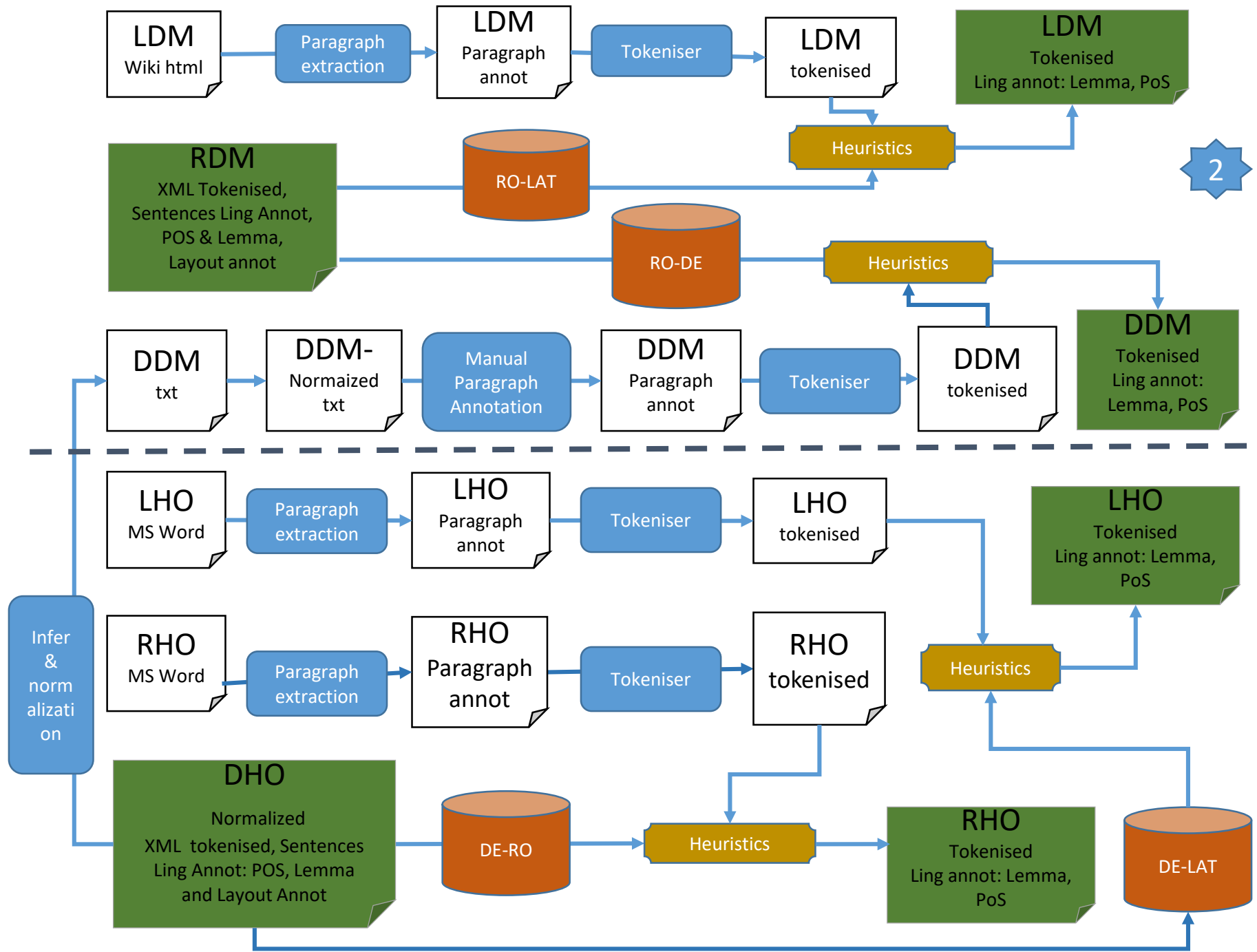
HerCoRe

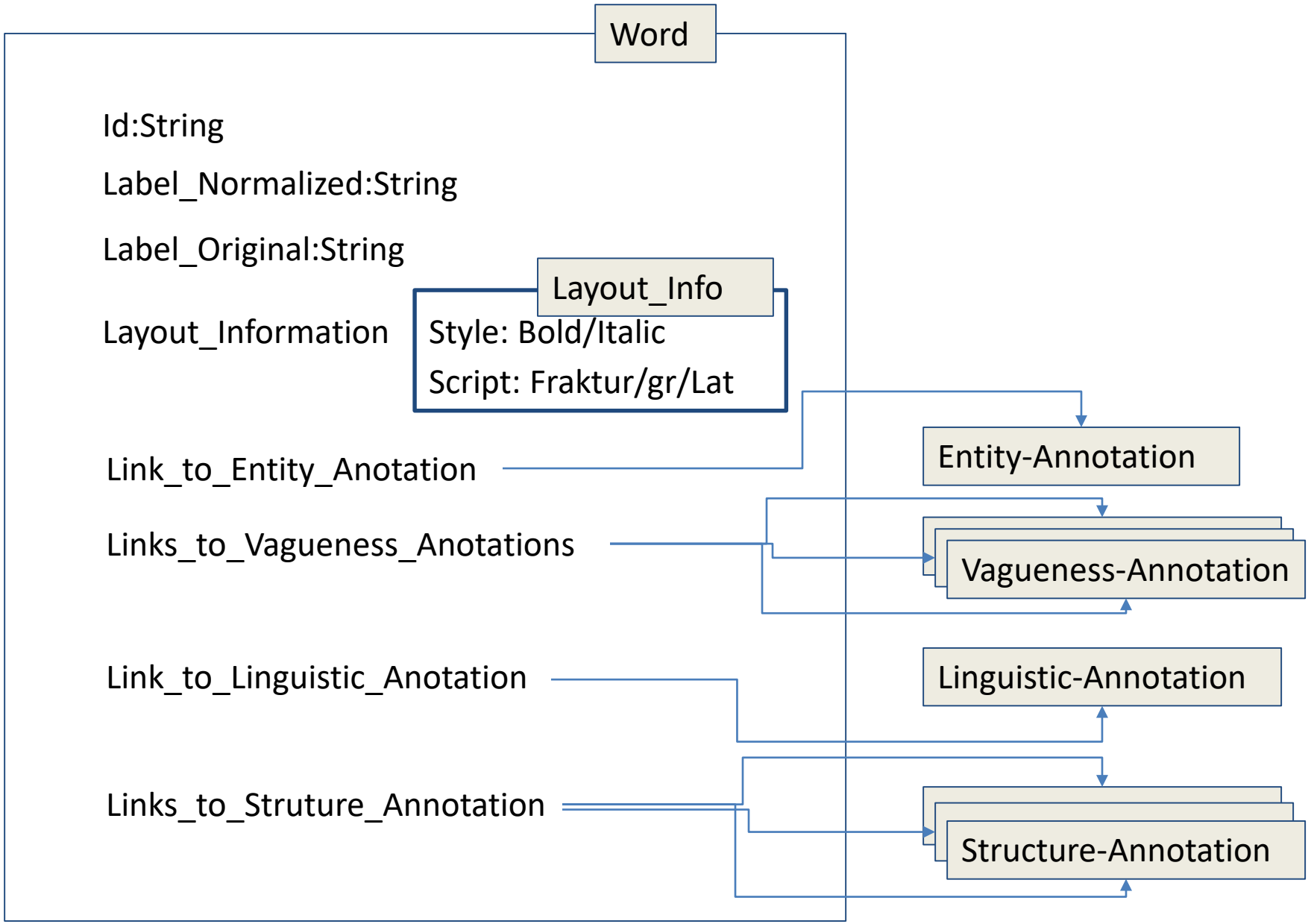
System Architecture
Workflow for the preprocessing step
Architecture of the annotation tool
Cristina Vertan, University of Hamburg

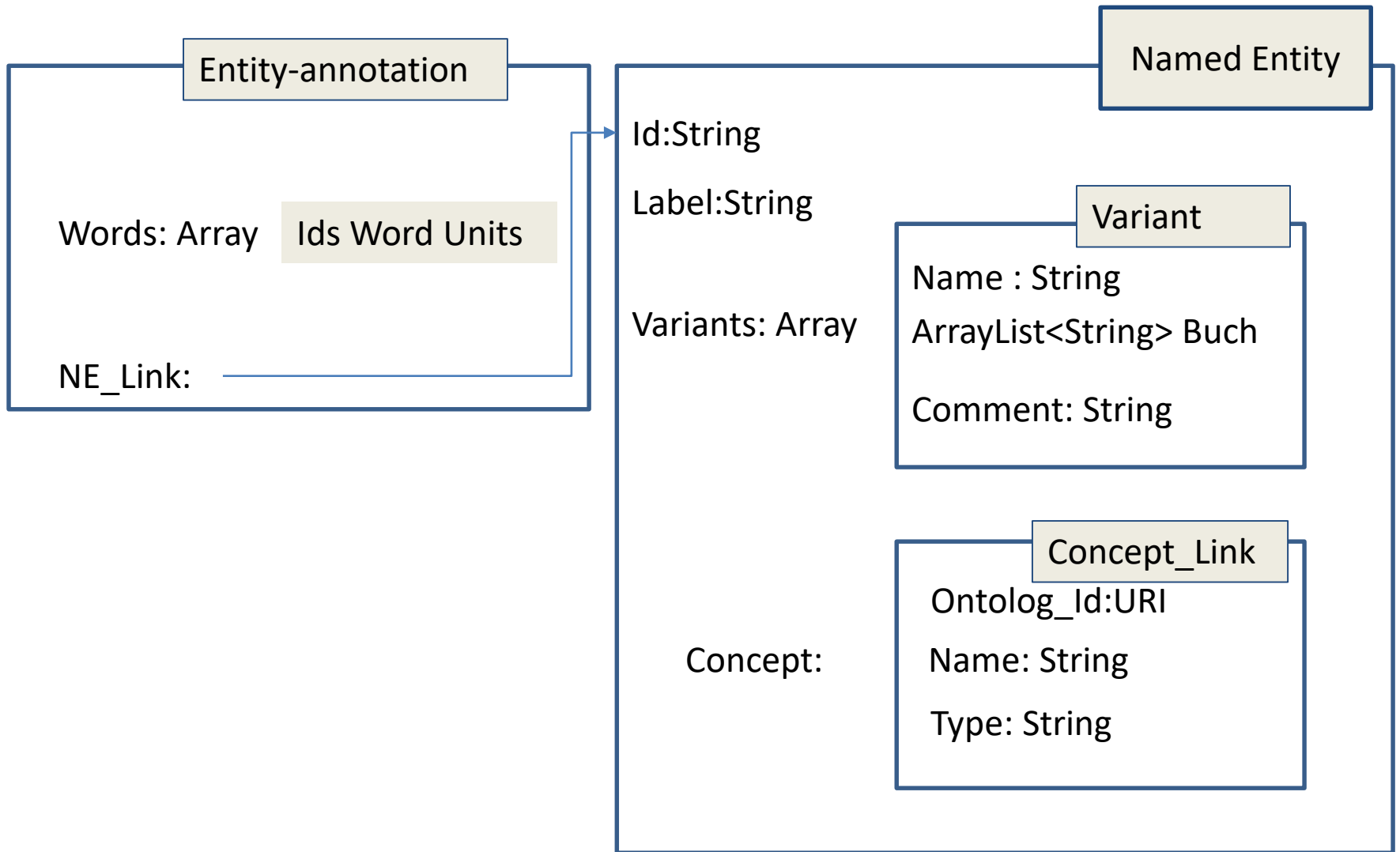
System Architecture



P
R
E
-
P
R
O
C
E
S
S
I
N
G







Vagueness-annotation

Id:String

Type: String (Quotation /Linguistic / Edition /Geo/Genre)

Subtype: String (dependent on each type)

Words: Array Word_Id

Confidence: String (low/medium/high)

Linguistic-annotation

Id:String

Lemma: String

PoS: String

Morpho_Features: String

Words: Array Word_Id

Structure-annotation

Id:String

Level: String (Chap/Paragraph/Sentence)

Type: String (Author/Editor)

Words: Array Word_Id

S1
Level: Sentence
Type: Author
Words: W1,W2,W3,W4,W5,W6,W7,W8,W9,W10,W11

W1 Die Die n frak	W2 Moldau Moldau b frak	W3 hat hat n frak	W4 nicht nicht n frak	W5 zu zu n frak	W6 allen allen n frak	W7 Zeiten Zeiten n frak	W8 eben eben n frak	W9 dieselben dieselben n frak	W10 Grenzen Gränzen n frak	W11 gehabt gehabt n frak
--------------------------------------	--	--------------------------------------	--	------------------------------------	--	--	--	--	---	---

Ling Ann :L2
Lemma:Moldau
PoS :N
Morpho:
Words: W2

Ent Ann :E1
Words: W2
NE20

NE20
Label: Moldova
Var: Moldau (DDM)
URI: https://..
Name: placeName
Type: country

Vagueness Ann :V1
Type: ling
Subtype : time expression
Confid: low
Words: W5,W6,W7

Ling Ann :L3
Lemma:haben
PoS :Verb
Morpho: 3.sg. Perf.
Words: W3,W11

