

# زمانه وانیی کۆمپیوتەری

بابەتی 2: سەرچاوه زمانه وانیه کان

سینا ئەحمەدی

[sinaahmadi.github.io/KurdishCL](https://sinaahmadi.github.io/KurdishCL)



1 دهسپیک

2 پیدراوهی زمان

3 پیدراوهکان بهپیی فۆرماتهکان

4 بهکارهینانی سه رچاوه زمانیهکان

5 بهشی پراکتیک



# دهسپیک





© Sina Ahmadi



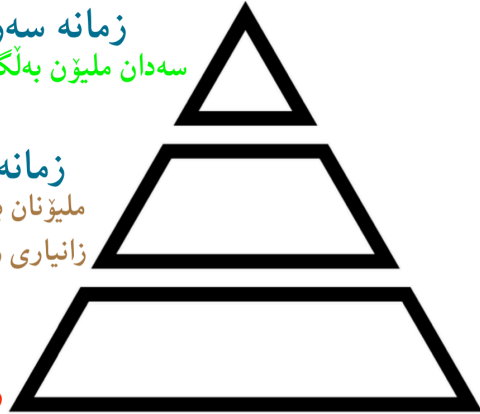
# سەرچاوه زمانه وانیه کان بو زمانه وانیه کو مپیوتهری

- پیدراوهی زمانه وانیه ← یه کهم ههنگاو بو پرۆسهس کردنی زمان

زمانه سەرچاوه دهوله مهنده کان  
سه دان ملیون به لگهی سهرهیل و پیدراوه و زانیاری  
وهک ئینگلیسی

زمانه مامناوهندیه کان  
ملیونان به لگهی سهرهیل و پیدراوه و  
زانیاری زۆربهی زمانه ئه ورووپیه کان

زمانه ههژاره کان  
بی یان کهم/سەرچاوه  
زۆربهی زمانه کانی دنیا



# سەرچاوه زمانه وانیه کان بۆ زمانه وانیه کۆمپیوتهری

- پیدراوهی زمانه وانیه ← یه کهم ههنگاو بۆ پرۆسهس کردنی زمان
- بهشیکی زۆری زمانه کانی دنیا سەرچاوهی زمان و زمانه وانیه کان بۆ کهمه [1]

زمانه سەرچاوه دهوله مهنده کان  
سه دان ملیۆن به لگهی سه رهیل و پیدراوه و زانیاری  
وهک ئینگلیسی

زمانه مامناوه ندییه کان  
ملیۆنان به لگهی سه رهیل و پیدراوه و  
زانیاری زۆربهی زمانه ئه ورووپیه کان

زمانه هه ژاره کان  
بی یان کهم/سه رهچاوه  
زۆربهی زمانه کانی دنیا



# سەرچاوه زمانه وانیه کان بو زمانه وانیه کو مپیوتهری

- پیدراوهی زمانه وانیه ← یه کهم ههنگاو بو پرۆسهس کردنی زمان
- بهشیکی زۆری زمانه کانیه دنیا سەرچاوهی زمان و زمانه وانیه کان بو کهمه [1]
- زمانی کوردی زمانیهی کهم سەرچاوهی کهمه (less-resourced)

زمانه سەرچاوه دهوله مهنده کان  
سه دان ملیۆن به لگه ی سهرهیل و پیدراوه و زانیاری  
وهک ئینگلیسی

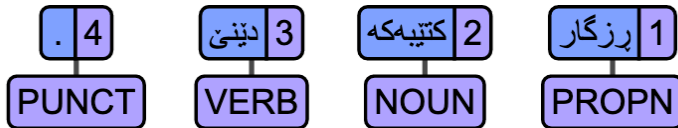
زمانه مامناوه ندییه کان  
ملیۆنان به لگه ی سهرهیل و پیدراوه و  
زانیاری زۆربه ی زمانه ئه ورووپیه کان

زمانه هه ژاره کان  
بی یان کهم/سەرچاوه  
زۆربه ی زمانه کانیه دنیا



# به کارهینانی سه رچاوه کانی زمان

- ناسینه وهی به شه کانی وشه (part-of-speech tagging)





# به کارهییانی سه رچاوه کانی زمان

- ناسینه وهی به شه کانی وشه (part-of-speech tagging)
- راست کردنه وهی هه لهی رینووس (spelling error correction)

نوام چامسکی و زنجیره پلهی زمان و ریزمانه فورماله کان

زنجیره پلهی

هه زنجیره ره قه

Ignore

Ignore All

Add to Dictionary

Spelling...



# به کارهییانی سه رچاوه کانی زمان

- ناسینه وهی به شه کانی وشه (part-of-speech tagging)
- راست کردنه وهی هه لهی رینووس (spelling error correction)
- لیک دانه وهی واتای وشه کان: «داریکی به دهسته وه بوو.» (word-sense disambiguation)



دار



- ناسینه وهی به شه کانی وشه (part-of-speech tagging)
- راست کردنه وهی هه لهی ریڤنوس (spelling error correction)
- لیک دانه وهی واتای وشه کان: «داریکی به دهسته وه بوو.» (word-sense disambiguation)
- ناوانی رۆله ماناداره کان (semantic role labeling)



# به کارهییانی سه رچاوه کانی زمان

- ناسینه وهی به شه کانی وشه (part-of-speech tagging)
- راست کردنه وهی هه لهی رینووس (spelling error correction)
- لیک دانه وهی واتای وشه کان: «داریکی به دهسته وه بوو.» (word-sense disambiguation)
- ناوانی رۆله ماناداره کان (semantic role labeling)
- په یوهندی دانی چه مکه کان له رسته دا (entity linking)



- ناسينه وهی به شه کانی وشه (part-of-speech tagging)
- راست کردنه وهی هه لهی رینووس (spelling error correction)
- ليک دانه وهی واتای وشه کان: «داريکی به دهسته وه بوو.» (word-sense disambiguation)
- ناوانی رۆله ماناداره کان (semantic role labeling)
- په يوهندی دانی چه مکه کان له رسته دا (entity linking)
- ويړای پيشکه وتنه کان له بواری ژيريی دهستکرددا، سه رچاوه کانی زمان هيشتاش رۆليکی گرینگيان له تهکنۆلۆژيای زماندا ههيه.



## سەرچاوهی زمانه وانیه/زمانی

سەرچاوهی زمانه وانیه بریتیه له پیکهاتهی بهرهمه زمانه وانیه و زمانیهکان که بو چی کردن، باستر کردن یان هه لسه نگاندهی تهکنولۆژیای زمان و تووژینهوهی سهر زمان به کار دههیندرین.

1 پیدراوهی زمانه وانیه: نهو سه چاوه یانهی زمانیک له ریگای پیدراوهوه ده ناسین، وهک فرههنگ، لیستی وشه یان ته نانهت دهنگی شیوه زاریک



## سەرچاوهی زمانه وانیه/زمانی

سەرچاوهی زمانه وانیه بریتیه له پیکهاتهی بهرهمه زمانه وانیه و زمانیهکان که بو چی کردن، باستر کردن یان هه لسه نگاندهی ته کنولۆژیای زمان و تووژینه وهی سهر زمان به کار دههیندرین.

- 1 پیدراوهی زمانه وانیه: نهو سه چاوه یانهی زمانیک له ریگای پیدراوه وه ده ناسین، وهک فرههنگ، لیستی وشه یان ته نانهت دهنگی شیوه زاریک
- 2 نهو ئامیرانهی که بو به کارهینان، نیشان دان یان دهست تی وهردانی پیدراوهکان پیویستن

## سەرچاوهی زمانه وانیه/زمانی

سەرچاوهی زمانه وانیه بریتیه له پیکهاتهی بهرهمه زمانه وانیه و زمانیهکان که بو چی کردن، باستر کردن یان هه لسه نگاندهی ته کنولۆژیای زمان و تووژینه وهی سهر زمان به کار دههیندرین.

- 1 پیدراوهی زمانه وانیه: نهو سه چاوه یانهی زمانیک له ریگای پیدراوه وه ده ناسین، وهک فرههنگ، لیستی وشه یان ته نانهت دهنگی شیوه زاریک
- 2 نهو ئامیرانهی که بو به کارهینان، نیشان دان یان دهست تی وهردانی پیدراوهکان پیوستن
- 3 میتاداتا و وشه سازیهکان که بو مامه له کردن له گه ل پاراستنی سه رچاوهکان به کار دین، به تایه تی بهرده وامی (sustainability)، بهرده ستیوون و ئه رشیفکردنی درێژخایه ن





# پیداوہی زمان



هه‌ندیك له گرنگترین جوړه‌کانی پیداوه زمانه‌وانییه‌کان ئەمانه‌ن:  
① وه‌سفی زمان وه‌ک ری‌زمان (grammar) و مودی‌ل کردنی فۆرمال



هه‌ندیك له گرنگترین جوړه‌کانی پیداوه زمانه‌وانییه‌کان ئەمانه‌ن:

- 1 وه‌سفی زمان وه‌ک ری‌زمان (grammar) و مودی‌ل کردنی فۆرمال
- 2 کۆرپس (corpora / corpus به‌کۆ): پیکهاته‌یه‌ک له ده‌ق یان ده‌نگ به‌ شیوه‌ی دیجیتال



هه‌ندیك له گرنگترین جوړه‌کانی پیداوه زمانه‌وانییه‌کان ئەمانه‌ن:

- 1 وه‌سفی زمان وه‌ک ری‌زمان (grammar) و مودی‌ل کردنی فۆرمال
- 2 کۆرپس (corpora / corpus به‌کۆ): پیکهاته‌یه‌ک له ده‌ق یان ده‌نگ به‌ شیوه‌ی دیجیتال
- 3 سه‌رچاوه‌ی واتاناسی-وشه‌ناسی (lexical-semantic): ری‌ک خستی سه‌رچاوه‌ی زمان به‌ پی‌ی بو‌چوون یان تیورییه‌ک له واتاناسی یان وشه‌ناسیدا



هه‌ندیك له گرنگترین جوړه‌کانی پیداوه زمانه‌وانییه‌کان ئەمانه‌ن:

- 1 وه‌سفی زمان وه‌ک ری‌زمان (grammar) و مودی‌ل کردنی فۆرمال
- 2 کۆرپس (corpora / corpus به‌کۆ): پیکهاته‌یه‌ک له ده‌ق یان ده‌نگ به‌ شیوه‌ی دیجیتال
- 3 سه‌رچاوه‌ی واتاناسی-و‌شه‌ناسی (lexical-semantic): ری‌ک خستی سه‌رچاوه‌ی زمان به‌ پی‌ی بۆ‌چوون یان تیورییه‌ک له واتاناسی یان و‌شه‌ناسیدا
- 4 ئۆنتۆلۆژی (ontology)، زاراوه‌ناسی (terminology) و فه‌ره‌ه‌نگۆک (glossary)



هەندیک له گرنگترین جوّرهکانی پیداوه زمانهوانیهکان ئەمانه:

- 1 وهسفی زمان وهک ریزمان (grammar) و مودیل کردنی فۆرمال
- 2 کۆرپس (corpora / corpus به کۆ): پیکهاتهیهک له دهق یان دهنگ به شیوهی دیجیتال
- 3 سه‌رچاوهی واتاناسی-وشه‌ناسی (lexical-semantic): ریک خستنی سه‌رچاوهی زمان به پیی بوچوون یان تیورییهک له واتاناسی یان وشه‌ناسیدا
- 4 ئۆنتۆلۆژی (ontology)، زاراوه‌ناسی (terminology) و فهره‌نگۆک (glossary)
- 5 مۆدیلی زمان (language model) که له سه‌ر کۆرپسه‌گه‌وره‌کاندا فیرده‌کری

هه‌ندیك له گرنه‌تیرین جوهره‌کانی پیدراوه زمانه‌وانیه‌کان ئه‌مانه‌ن:

- 1 وه‌سفی زمان وه‌ک ریژمان (grammar) و مودیل کردنی فۆرمال
- 2 کۆرپس (corpora / corpus به‌ کۆ): پیکهاته‌یه‌ک له ده‌ق یان ده‌نگ به‌ شیوه‌ی دیجیتال
- 3 سه‌رچاوه‌ی واتاناسی-و‌شه‌ناسی (lexical-semantic): ریک خستی سه‌رچاوه‌ی زمان به‌ پیی بو‌چوون یان تیورییه‌ک له واتاناسی یان و‌شه‌ناسیدا
- 4 ئۆنتۆلۆژی (ontology)، زاراوه‌ناسی (terminology) و فه‌ره‌ه‌نگۆک (glossary)
- 5 مۆدیلی زمان (language model) که له سه‌ر کۆرپسه‌ گه‌وره‌کاندا فی‌رده‌کری
- 6 گرافی زانیاری (knowledge graph) و بنکه‌ی زانیاری (knowledge base)



## • ریزمانی زمان وهک ریزمانی کوردیی تاکستۆن [2]

### AFFIRMATIVE PRESENT

دهمهوئى <b>dámawe</b>	دهمانهوئى <b>dámânawe</b>
دهتهوئى <b>dátawe</b>	دهتانهوئى <b>dátânawe</b>
دهیهوئى <b>dáyawe</b>	دهیانهوئى <b>dáyânawe</b>

### NEGATIVE PRESENT

نهمهوئى <b>námawe</b>	نهمانهوئى <b>námânawe</b>
نتهوئى <b>nátawe</b>	نتهتانهوئى <b>nátânawe</b>
نیهوئى <b>náyawe</b>	نیهتانهوئى <b>náyânawe</b>

### PRESENT SUBJUNCTIVE

بمهوئى <b>bímawe</b>	بمانهوئى <b>bímânawe</b>
بتهوئى <b>bítawe</b>	بتانهوئى <b>bítânawe</b>
بییهوئى <b>bíyawe</b>	بییهتانهوئى <b>bíyânawe</b>





- ریزمانی زمان وهک ریزمانی کوردیی تاکستۆن [2]
- فۆرمالیسم و یاساکان

$$\left[ \begin{array}{l} /XY(\mathbf{in})/_{V} \\ x (= \text{چاۆگ}) \\ \text{کردن} (kirdin) \end{array} \right] \leftrightarrow \left[ \begin{array}{l} /XY/_{V} \\ x \text{ رهگی رابردووی} \\ \text{کرد} (kird) \end{array} \right] \leftrightarrow \left[ \begin{array}{l} /XZ/_{V} \\ x \text{ رهگی ئیستای} \\ \text{که} (ke) \end{array} \right]$$

- ریزمانی زمان وهک ریزمانی کوردیی تاکستۆن [2]
- فورمالیسم و یاساکان

$$\left[ \begin{array}{l} /XY(\text{in})/_{V} \\ x (= \text{چاوغ}) \\ \text{کردن} (kirdin) \end{array} \right] \leftrightarrow \left[ \begin{array}{l} /XY/_{V} \\ x \text{ رهگی رابردووی} \\ \text{کرد} (kird) \end{array} \right] \leftrightarrow \left[ \begin{array}{l} /XZ/_{V} \\ x \text{ رهگی ئیستای} \\ \text{که} (ke) \end{array} \right]$$

که تییدا  $X$  رهگی کرداره که یه و  $Y$  و  $Z$  یش ههردووکیان پاشگری دارپژانن.

- له سه‌بنه‌مایه‌کی واتاناسی یان وشه‌ناسیدا به شیوه‌یه‌کی وه‌سفکه‌ر و پیکهاته‌دار ری‌ک خراون.



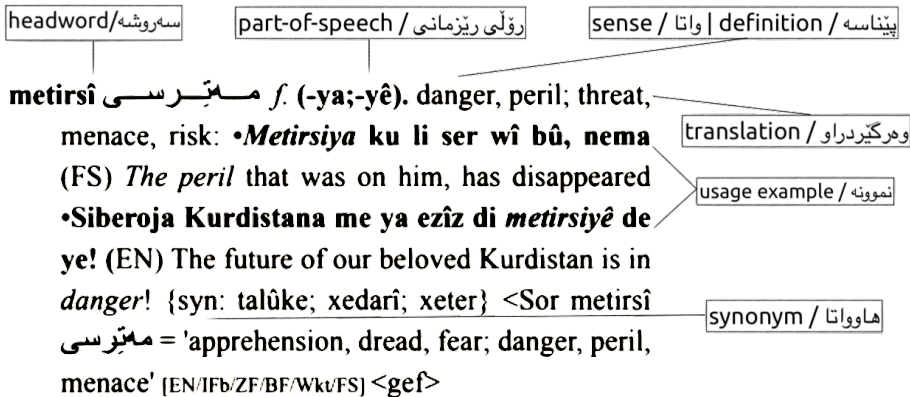
# پیداووی زمان: سه‌رچاوه‌ی واتاناسی-وشه‌ناسی

- له سه‌ر بنه‌مایه‌کی واتاناسی یان وشه‌ناسیدا به شیوه‌یه‌کی وه‌سفکه‌ر و پیکهاته‌دار ری‌ک خراون.
- هه‌ندی‌ک له گرینگترینی ئەو سه‌رچاوه‌یانه ئەمانه‌ن:



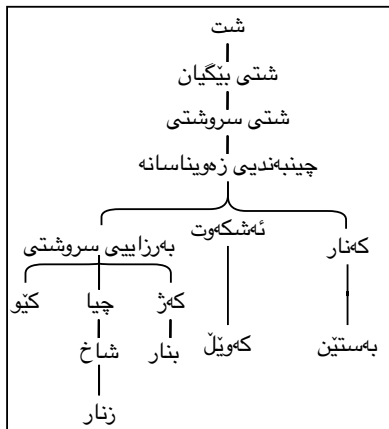
# پیداوهی زمان: سه‌رچاوهی واتاناسی-وشه‌ناسی

- له سه‌ر بنه‌مایه‌کی واتاناسی یان وشه‌ناسیدا به شیوه‌یه‌کی وه‌سفکه‌ر و پیکهاته‌دار رپیک خراون.
- هه‌ندی‌ک له گرینگترینی ئه‌و سه‌رچاوه‌یانه ئه‌مانه‌ن:
- فره‌هنگی وشه وه‌ک هه‌نبانه بو‌رینه یان Ferhenga Birûskî [3]



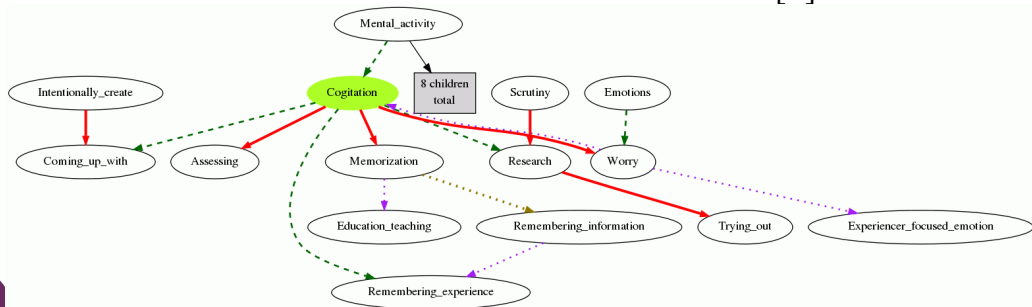
# پیداوهی زمان: سه‌رچاوهی واتاناسی-وشه‌ناسی

- له سه‌ر بنه‌مایه‌کی واتاناسی یان وشه‌ناسیدا به شیوه‌یه‌کی وه‌سفکه‌ر و پیکهاته‌دار ری‌ک خراون.
- هه‌ندی‌ک له گرینگترینی ئه‌و سه‌رچاوه‌یانه ئه‌مانه‌ن:
- فره‌ه‌نگی وشه وه‌ک هه‌نبانه بو‌رینه یان [3] Ferhenga Birûskî
- وو‌ردنی‌ت (WordNet) [4] (به‌سته‌ر) وه‌ک کوردنی‌ت [5]



# پیداوهی زمان: سه رچاوهی واتاناسی-وشه ناسی

- له سه ر بنه مایه کی واتاناسی یان وشه ناسیدا به شیوهیه کی وه سفکه ر و پیکهاته دار ریگ خراون.
- هه ندیک له گرینگترینی ئه و سه رچاوه یانه ئه مانه ن:
- فرههنگی وشه وهک هه نبانه بۆرینه یان [3] Ferhenga Birûskî
- ووۆردنیی (WordNet) [4] (به سته ر) وهک کوردنیی [5]
- [6] FrameNet



- زاراوه یان زار (terminology) به وشه یهک ده کوتری که تایبته به بواریکه وهک زاراوه ی پزیشکی یان زاراوه ی ته کنۆلۆژیا



- زاراوه یان زار (terminology) به وشه یهک ده کوترئ که تایبته به بواریکه وهک زاراوه ی پزیشکی یان زاراوه ی ته کئولۆژیا
- سه رجه م زاره کان زۆر جار پئیان ده کوترئ فهرهه نگوک (glossary)

# سەرچاوه زمانه وانیهکان: زاراوه سازی

- زاراوه یان زار (terminology) به وشهیهک دهکوتری که تایبته به بواریکه وهک زاراوهی پزیشکی یان زاراوهی تهکنۆلۆژیا
- سهرجهم زارهکان زۆر جار پییان دهکوتری فهرههنگۆک (glossary)
- هاوشیوهی فهرههنگ له سهروهه پیک هاتوهه و ههندیك جار زانیاریی زمانه وانیشی لهگه لدایه. تیرمینۆلۆژیی یهکیهتی ئه ورووپا (بهستهه)

The screenshot shows a web interface for a terminology database. At the top, it displays the ID '1341869' and the language 'en'. The entry is for 'morpheme' in the 'social sciences' category. It includes a term reference to ISO/R 1087:1969 and a Wikipedia link. The definition is 'smallest semantically meaningful unit in a language'. It also shows a language level note and creation/modification dates. At the bottom, there is a French entry for 'morphème'.

1341869 1

social sciences [SCIENCE, humanities] COM

en morpheme \*\*\*

Term reference: - ISO/R 1087:1969 Vocabulary of terminology (withdrawn)  
- Wikipedia > Morpheme, <http://en.wikipedia.org/wiki/Morpheme> [10.1.2012]

Definition: smallest semantically meaningful unit in a language

Definition reference: Wikipedia > Morpheme, <http://en.wikipedia.org/wiki/Morpheme> [10.1.2012]

Language level note: See also:  
lexeme [ IATE:1615017 ]

Creation date: 24.9.2003 1:00 Legacy database: Eurodicautom  
Created by: (COM) Legacy ID: BTL-AFN90-200355  
Modification date: 19.5.2014 12:36  
Modified by: (CdT)

fr morphème \*\*



- زاراوه یان زار (terminology) به وشه یهک دهکوتری که تایبته به بواریکه وهک زاراوه ی پزیشکی یان زاراوه ی تهکنۆلۆژیا
- سه رجه م زارهکان زۆر جار پییان دهکوتری فرههنگۆک (glossary)
- هاوشیوه ی فرههنگ له سهروشه پیک هاتوو و هه ندیک جار زانیاری زمانه وانیشی له گه لدایه. تیرمینۆلۆژی یه کیه تیی ئه ورووپا (به سته ر)
- ئامانجی زاراوه ریک خستن و پۆلاندنی زانیاریه بو ساز کردنی پۆلینه ناسی (taxonomy)



- زاراوه یان زار (terminology) به وشه‌یه‌ک ده‌کوتری که تایبته به بواریکه وهک زاراوه‌ی پزیشکی یان زاراوه‌ی ته‌کنۆلۆژیا
- سه‌رجه‌م زاره‌کان زۆر جار پییان ده‌کوتری فه‌ره‌ه‌نگۆک (glossary)
- هاوشیوه‌ی فه‌ره‌ه‌نگ له سه‌روه‌ه‌ پیک هاتووه و هه‌ندی‌ک جار زانیاریی زمانه‌وانی‌شی له‌گه‌ل‌دایه.
- تیرمینۆلۆژیی یه‌کیه‌تیی ئه‌ورووپا (به‌سته‌ر)
- ئامانجی زاراوه‌ریک خستن و پۆلاندنی زانیارییه‌ بو‌ سازکردنی پۆلینه‌ناسی (taxonomy)
- له‌ ته‌کنۆلۆژیای زمان زۆر به‌سوودن بو‌ زۆرکردنه‌وه‌ی ئاستی ناسینه‌وه‌ی وشه‌کان



- زاراوه یان زار (terminology) به وشه یهک دهکوتری که تایبته به بواریکه وهک زاراوه ی پزیشکی یان زاراوه ی تهکنۆلۆژیا
- سه رجه م زارهکان زۆر جار پییان دهکوتری فرههنگۆک (glossary)
- هاوشیوه ی فرههنگ له سهروشه پیک هاتوه و هه ندیک جار زانیاری زمانه وانیشی له گه لدایه.
- تیرمینۆلۆژی یه کیه تیی ئه ورووپا (به سته ر)
- ئامانجی زاراوه ریک خستن و پۆلاندنی زانیاریه بو ساز کردنی پۆلینه ناسی (taxonomy)
- له تهکنۆلۆژیای زمان زۆر به سوودن بو زۆر کردنه وه ی ئاستی ناسینه وه ی وشهکان
- هه ندیک بنکه ی زاری کوردیی به سوود



- زاراوه یان زار (terminology) به وشه‌یه‌ک ده‌کوتری که تایبته به بواریکه وهک زاراوه‌ی پزیشکی یان زاراوه‌ی ته‌کنۆلۆژیا
- سه‌رجه‌م زاره‌کان زۆر جار پییان ده‌کوتری فه‌ره‌ه‌نگۆک (glossary)
- هاوشیوه‌ی فه‌ره‌ه‌نگ له سه‌روه‌ه‌ پیک هاتووه و هه‌ندی‌ک جار زانیاریی زمانه‌وانیسی له‌گه‌ل‌دایه.
- تیرمینۆلۆژیی یه‌کیه‌تیی ئه‌ورووپا (به‌سته‌ر)
- ئامانجی زاراوه‌ریک خستن و پۆلاندنی زانیارییه‌ بو سازکردنی پۆلینه‌ناسی (taxonomy)
- له ته‌کنۆلۆژیای زمان زۆر به‌سوودن بو زۆرکردنه‌وه‌ی ئاستی ناسینه‌وه‌ی وشه‌کان
- هه‌ندی‌ک بنکه‌ی زاری کوردیی به‌سوود
- وشه‌نامه‌ی هوژین: <https://hojan.org/dic/>



- زاراوه یان زار (terminology) به وشهیهک دهکوتری که تایبته به بواریکه وهک زاراوهی پزیشکی یان زاراوهی تهکنۆلۆژیا
- سهرجهه زارهکان زۆر جار پێیان دهکوتری فهرههنگۆک (glossary)
- هاوشیوهی فهرههنگ له سهروشه پیک هاتوووه و ههندیك جار زانیاریی زمانه وانیشی لهگهڵدایه. تیرمینۆلۆژیی یهکیهتی ئه ورووپا (بهستهه)
- ئامانجی زاراوه ریک خستن و پۆلاندنی زانیارییه بۆ ساز کردنی پۆلینه ناسی (taxonomy)
- له تهکنۆلۆژیای زمان زۆر به سوودن بۆ زۆرکردنه وهی ئاستی ناسینه وهی وشهکان ههندیك بنکهی زاری کوردیی به سوود
- وشه نامه ی هۆژین: <https://hojan.org/dic/>
- یاپراخ: <https://dict.linux.krd/>



## کۆرپس

کۆمهلیک پیدراوهی زمانه وانیه که به شیوهی دهق یان دهنگ کۆده کریته وه و وه کوو خالی دهستییک بۆ سهلماندنی گریمانه کان سه بارهت به زمانیک به کار دههیندری.

- زۆربهی کۆرپسه کان لانی کهم چهند ملیۆن وشه یان تیدایه.





## کۆرپس

کۆمهلیک پیدراوهی زمانه وانیهکان که به شیوهی دهق یان دهنگ کۆده کریتته وه و وهکوو خالی دهستییک بۆ سهلماندنی گریمانیهکان سهبارهت به زمانیک به کار دههیندری.

- زۆربهی کۆرپسهکان لانی کهم چهند ملیۆن وشه یان تیدایه.
- کۆرپس ته نیا چیژییک له زمانه کهیه، نهک زمانه که بۆ خوی.

## کۆرپس

کۆمهلیک پیدراوهی زمانه وانیه که به شیوهی دهق یان دهنگ کۆده کریتته وه و وه کوو خالی دهستییک بو سه لماندنی گریمانیه کان سه بارهت به زمانیک به کار دههیندری.

- زۆربهی کۆرپسه کان لانی کهم چه ند ملیۆن وشه یان تیدایه.
- کۆرپس ته نیا چیژییک له زمانه کهیه، نهک زمانه که بو خوی.
- به رای چۆمسکی، کۆرپس قهت ناتوانی هه موو دیاردهکانی زمانیک بینیتته سه لماندن [7].

## کۆرپس

کۆمهلیک پیدراوهی زمانه وانیهکان که به شیوهی دهق یان دهنگ کۆده کریتته وه و وهکوو خالی دهستییک بۆ سهلماندنی گریمانهکان سهبارهت به زمانیک به کار دههیندری.

- زۆربهی کۆرپسهکان لانی کهم چهند ملیۆن وشه یان تیدایه.
- کۆرپس ته نیا چیژییک له زمانه کهیه، نهک زمانه که بۆ خوی.
- به رای چۆمسکی، کۆرپس قهت ناتوانی هه موو دیاردهکانی زمانیک بیئیتته سهلماندن [7].
- بۆ ههر بابتهتییک، وهک زانست یان رامیاری، و ههر پیناسهیهکی زمان، وهک بنزاراوه یان هه پرهتی قسه کردن، دهتوانین کۆرپسمان هه بی.



## کۆرپس

کۆمهلیک پیدراوهی زمانه وانیهکان که به شیوهی دهق یان دهنگ کۆده کریتته وه و وهکوو خالی دهستییک بۆ سهلماندنی گریمانهکان سهبارهت به زمانیک به کار دههیندری.

- زۆربهی کۆرپسهکان لانی کهم چهند ملیۆن وشه یان تیدایه.
- کۆرپس ته نیا چیژیک له زمانه کهیه، نهک زمانه که بۆ خوی.
- به رای چۆمسکی، کۆرپس قهت ناتوانی هه موو دیاردهکانی زمانیک بیئیتته سهلماندن [7].
- بۆ ههر بابتهتیک، وهک زانست یان رامیاری، و ههر پیناسهیهکی زمان، وهک بنزاراوه یان ههرهتی قسه کردن، دهتوانین کۆرپسمان هه بی.
- کۆرپس ده بی نمونه، هاوسهنگ، نوینه ر و سروشتی بی [8].



## کۆرپس

کۆمهلیک پیدراوهی زمانه وانیهکان که به شیوهی دهق یان دهنگ کۆده کریتته وه و وهکوو خالی دهستیپک بۆ سهلماندنی گریمانیهکان سهبارهت به زمانیک به کار دههیندری.

- زۆربهی کۆرپسهکان لانی کهم چهند ملیۆن وشه یان تیدایه.
- کۆرپس ته نیا چیژیک له زمانه کهیه، نهک زمانه که بۆ خوێ.
- به رای چۆمسکی، کۆرپس قهت ناتوانی هه موو دیاردهکانی زمانیک بیئیتته سهلماندن [7].
- بۆ ههر بابتهتیک، وهک زانست یان رامیاری، و ههر پیناسهیهکی زمان، وهک بنزاراوه یان ههرهتی قسه کردن، دهتوانین کۆرپسمان هه بی.
- کۆرپس ده بی نمونه، هاوسهنگ، نوینه ر و سروشتی بی [8].
- مژاری سه رهکیی «زمانه وانیه کۆرپس» (corpus linguistics) تیگه یشتن له کۆرپسه.



## کۆرپس

کۆمهلیک پیدراوهی زمانه وانیه که به شیوهی دهق یان دهنگ کۆده کریتته وه و وه کوو خالی دهستییک بۆ سهلماندنی گریمانیه کان سه بارهت به زمانیک به کار دههیندری.

- زۆربهی کۆرپسه کان لانی کهم چه ند ملیۆن وشه یان تیدایه.
- کۆرپس ته نیا چیژیک له زمانه کهیه، نهک زمانه که بۆ خوی.
- به رای چۆمسکی، کۆرپس قهت ناتوانی هه موو دیاردهکانی زمانیک بیئیه سهلماندن [7].
- بۆ هه ر بابه تییک، وهک زانست یان رامیاری، و هه ر پیناسه یه کی زمان، وهک بنزاراوه یان هه ره تی قسه کردن، ده توانین کۆرپسمان هه بی.
- کۆرپس ده بی نمونه، هاوسه نگ، نوینه ر و سروشتی بی [8].
- مژاری سه ره کیی «زمانه وانیه کۆرپس» (corpus linguistics) تیگه یشتن له کۆرپسه.
- کۆرپسه کان به به ربلاوی له پرۆسه س کردنی زمانیشدا به کار دههیندرین.



# سەرچاوه زمانه وانیه کان: کۆرپس (دریژه)

نیوی کۆرپسه که	سەرچاوه	بنزاراوه	قهباره (وشه #)
کۆرپسی پیوان [9]	هه و آل	سۆرانی	18M
		کورمانجی	4M
کۆرپسی دهقی فۆلکلۆر [10]	گۆرانی، بهند و بهیت	سۆرانی	49K
کۆرپسی ئاسۆسافت [11]	هه و آل	سۆرانی	80M
کۆرپسی زازا-گۆرانی [12]	هه و آل	زازاکی	10M
		گۆرانی	2M
کۆرپسی پهروه ده [13]	کتیبی پهروه ده	سۆرانی	0.6M
کۆرپسی مه نچستر [14]	کۆکردنه وهی مهیدانی	سۆرانی	
		کورمانجی	

لیستی هه موو کۆرپسه کان بۆ کوردی لی ره دایه:

<https://github.com/sinaahmadi/awesome-kurdish>





- مۆدیل کردنی ئاماریی زمان که تایبه تمه ندیه کانیه زمانیک «فیڕ ده بی» رۆلی مۆرفۆلۆژی و سینتاکس، په یوه ندیه نیوان وشه کان و هتد



- مۆدیل کردنی ئاماری زمان که تایبه تمه نندیهکانی زمانیک «فیڕ دهبی» رۆلی مۆرفۆلۆژی و سینتاکس، په یوه نیدی نیوان وشهکان و هتد
- به سوود له مۆدیل کردنی زۆر دیاردهی زمان و بهربلاوه له ئه ورۆی تهکنۆلۆژیای زماندا

- مۆدیل کردنی ئاماریی زمان که تایبه تمه نندییهکانی زمانیک «فیڕ دهبی» رۆلی مۆرفۆلۆژی و سینتاکس، په یوه نیدی نیوان وشهکان و هتد
- به سوود له مۆدیل کردنی زۆر دیاردهی زمان و بهربلاوه له ئه ورۆی تهکنۆلۆژیای زماندا ئه و وشانهی له باری واتاوه له وشه «ماف» نزیکایه تیایان ههیه ↓



## ئۆنتۆلۆژی (ontology)

به پیکهاتهی ئه و یاسا و زانیارییهکانی دهکوتری که بو ناساندنی چه مکهکانی بواریک و په یوهندی نیوانیان به شیوهی فورمال به کارده هیندرین.

- ئۆنتۆلۆژییه کان به توندی فهرمین و به زمانیکی مهنتیقی دیاری کراون



## ئۆنتۆلۆژی (ontology)

به پیکهاتهی ئه و یاسا و زانیارییهکانی دهکوتری که بو ناساندنی چه مکهکانی بواریک و په یوهندی نیوانیان به شیوهی فۆرمال به کارده هیندریین.

- ئۆنتۆلۆژییهکان به توندی فهرمین و به زمانیکی مهنتیقی دیاری کراون
- «مانگ» یهک وشهیه له فهرههنگیکدا، بهلام دهتوانی به چهند شیوه پیناسه بکری: مانگی ههتاوی و مانگی هیجری

## ئۆنتۆلۆژی (ontology)

به پیکهاتهی ئه و یاسا و زانیارییهکانی دهکوترئ که بۆ ناساندنی چه مکهکانی بواریک و په یوهندی نیوانیان به شیوهی فۆرمال به کارده هیندریین.

- ئۆنتۆلۆژییهکان به توندی فهرمین و به زمانیکی مهنتیقی دیاری کراون
- «مانگ» یهک وشهیه له فهرههنگیکدا، بهلام دهتوانی به چهند شیوه پیناسه بکری:
- مانگی ههتاوی و مانگی هیجری
- سەرچاوه زمانه وانیهکان به پیی ئاستی فۆرمال بوونیان له ئۆنتۆلۆژی دهچن.

## ئۆنتۆلۆژی (ontology)

به پیکهاتهی ئه و یاسا و زانیارییهکانی دهکوترئ که بۆ ناساندنی چه مکهکانی بواریک و په یوهندی نیوانیان به شیوهی فۆرمال به کارده هیندریین.

- ئۆنتۆلۆژییهکان به توندی فهرمین و به زمانیکی مهنتیقی دیاری کراون
- «مانگ» یهک وشهیه له فهرههنگیکدا، بهلام دهتوانی به چهند شیوه پیناسه بکری:
- مانگی ههتاوی و مانگی هیجری
- سەرچاوه زمانه وانیهکان به پیی ئاستی فۆرمال بوونیان له ئۆنتۆلۆژی دهچن.
- ئۆنتۆلۆژی له بواری مهنتیق و کۆمپیوتهردا جیاوازی ههیه. چاو بکه له [15]



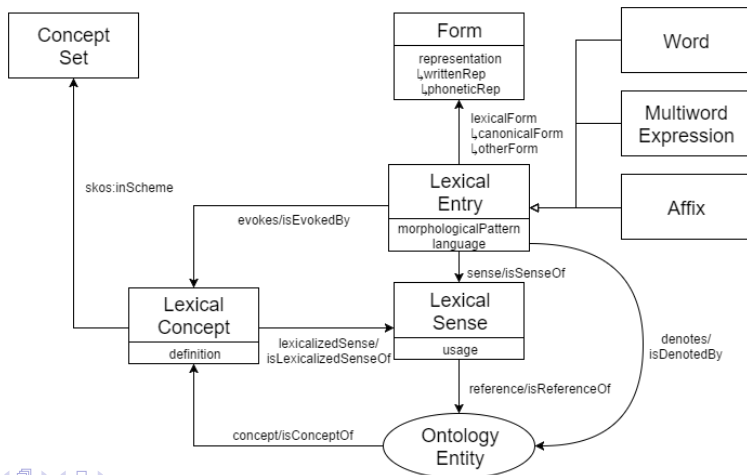
## ئۆنتۆلۆژی (ontology)

به پیکهاتهی ئه و یاسا و زانیاریانی ده کوترئ که بو ناساندنی چه مکه کانی بواریک و په یوه ندیی نیوانیان به شیوه ی فۆرمال به کارده هیندرین.

- ئۆنتۆلۆژییه کان به توندی فه رمین و به زمانیکی مه نتیقی دیاری کراون
- «مانگ» یه ک وشه یه له فه ره هه نگیکدا، به لام ده توانی به چه ند شیوه پیناسه بکری:
- مانگی هه تاوی و مانگی هه جری
- سه رچاوه زمانه وانیه کان به پیی ئاستی فۆرمال بوونیان له ئۆنتۆلۆژی ده چن.
- ئۆنتۆلۆژی له بواری مه نتیق و کۆمپیوته ردا جیاوازی هه یه. چا و بکه له [15]
- ئەندازیاریی ئۆنتۆلۆژی باهتی سه ره کیی بواری «وییی سیمانتیک» (semantic web) ه.

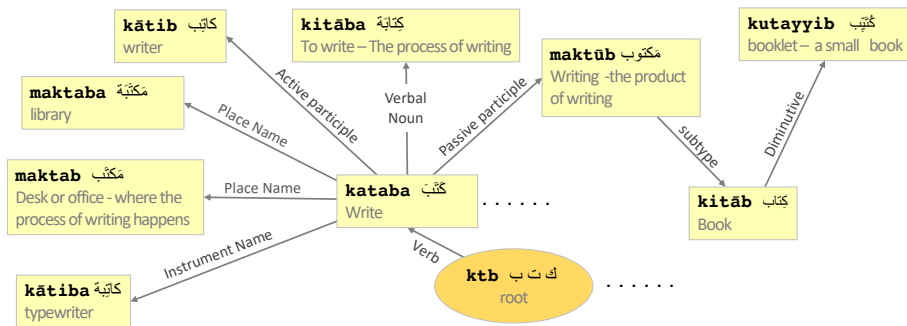
# سەرچاوه زمانه وانیه کان: ئونتۆلۆژی (دریژه)

- وشه سازی و مۆدیلیکی پیدراوه بۆ سه رچاوه کانی وشه، وه ک فهره ننگه کان [16]



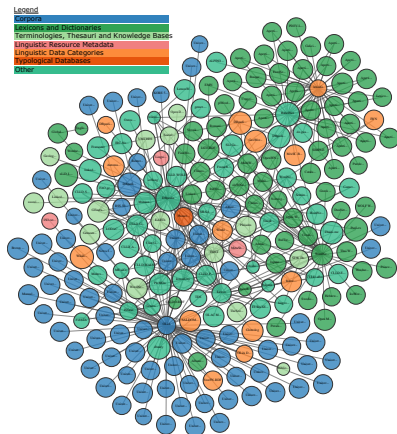
# سەرچاوه زمانه وانیهکان: ئۆنتۆلۆژی (دریژه)

- وشه‌سازی و مۆدییلیکی پیدراوه بۆ سه‌رچاوه‌کانی وشه، وه‌ک فه‌ره‌ه‌نگه‌کان [16]
- مۆدییلیکی پیدراوه بۆ نیشان دانی په‌یوه‌ندی واتایی و بنه‌ماله‌یی نیوان وشه‌کان [17]



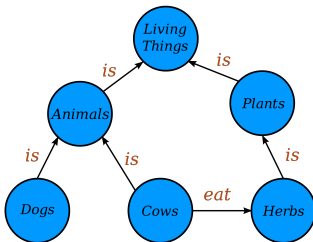
# سەرچاوه زمانه وانیه کان: ئۆنتۆلۆژی (دریژه)

- وشه سازی و مۆدیلیکی پیدراوه بۆ سەرچاوه کانی وشه، وهک فهرههنگه کان [16]
- مۆدیلیکی پیدراوه بۆ نیشان دانی په یوهندی واتایی و بنه مالهیی نیوان وشه کان [17]
- ههوری سەرچاوه په یوهندیاره کان (بهسته)




# سەرچاوه زمانه وانیهکان: گرافی زانیاری و بنکهی زانیاری


- نواندنی زانیاری و پیدراوهکان به شیوای گراف (graph)



# سەرچاوه زمانه وانیهکان: گرافی زانیاری و بنکهی زانیاری

- نواندنی زانیاری و پیدراوهکان به شیوازی گراف (graph)
- پشت بهستن به بناغهی ویی سیمانتیک و مۆدیل کردنی پیدراوهکان

**Barack Obama** 

44th U.S. President 

[barackobama.com](http://barackobama.com)

Barack Hussein Obama II is an American politician who served as the 44th president of the United States from 2009 to 2017. A member of the Democratic Party, he was the first African-American president of the United States. [Wikipedia](#)

**Born:** August 4, 1961 (age 61 years), Kapi'olani Medical Center for Women & Children, Honolulu, Hawaii, United States

**Full name:** Barack Hussein Obama II [Trending](#)

**Height:** 1.87 m

**Presidential term:** January 20, 2009 – January 20, 2017

**Education:** Harvard Law School (1988–1991), MORE

**Children:** Malia Ann Obama, Sasha Obama



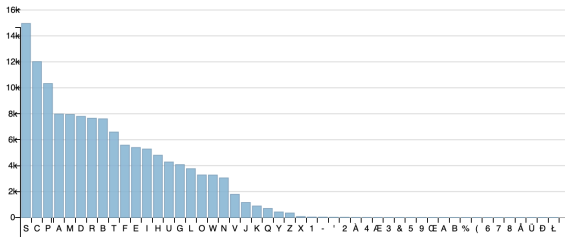
# سەرچاوه زمانه وانیه کان: گرافی زانیاری و بنکهی زانیاری

- نواندنی زانیاری و پیدراوه کان به شیوازی گراف (graph)
- پشت بهستن به بناغهی ویی سیمانیک و مودیل کردنی پیدراوه کان
- لیک کارکردنی یه کجار زوری زانیاری له سهر وییدا ← چاره سه ریکی باش بو ئیستای «پیدراوه په یوه نیدیاری» هکان (linked data)



# سەرچاوه زمانه وانیهکان: گرافی زانیاری و بنکهی زانیاری

- نواندنی زانیاری و پیدراوهکان به شیوازی گراف (graph)
- پشت بهستن به بناغهی ویبی سیمانتیک و مۆدیل کردنی پیدراوهکان
- لیك کارکردنی یه کجار زۆری زانیاری له سهر وییدا ← چاره سه ریکی باش بو ئیستای «پیدراوه په یوه نیدار»هکان (linked data)
- به سوودیشه بو زانیاریی زمانه وانیهکان
- ئەژماری وشه ئینگلیسییهکان به پیی یه کهم پیتیان (به ستهر)





# سەرچاوه زمانه وانیهکان: گرافی زانیاری و بنکهی زانیاری

- نواندنی زانیاری و پیدراوهکان به شیوازی گراف (graph)
- پشت بهستن به بناغهی ویبی سیمانتیک و مۆدیل کردنی پیدراوهکان
- لیک کارکردنی یه کجار زۆری زانیاری له سهر وییدا ← چاره سه ریکی باش بو ئیستای «پیدراوه په یوه نیدیدار» هکان (linked data)
- به سوودیشه بو زانیاریی زمانه وانیه
- ئەژماری وشه ئینگلیسییهکان به پیی یه کهم پیتیان (به ستهر)
- کۆمهلی گرافهکانی زانیاری دهبی به بنکهی زانیاری (knowledge base)



# سەرچاوه زمانه وانیهکان: گرافی زانیاری و بنکهی زانیاری

- نواندنی زانیاری و پیدراوهکان به شیوازی گراف (graph)
- پشت بهستن به بناغهی ویبی سیمانتیک و مۆدیل کردنی پیدراوهکان
- لیک کارکردنی یه کجار زۆری زانیاری له سهر وییدا ← چاره سه ریکی باش بو ئیستای «پیدراوه په یوه نیدیادار» هکان (linked data)
- به سوودیشه بو زانیاریی زمانه وانیه
- ئەژماری وشه ئینگلیسییهکان به پیی یه کهم پیتیان (به ستهر)
- کۆمهلی گرافهکانی زانیاری دهبی به بنکهی زانیاری (knowledge base)
- لیستی پشکنه رهکانی ئاسمان به یه ک کرته (به ستهر)



# سەرچاوه زمانه وانیهکان: گرافی زانیاری و بنکهی زانیاری

- نواندنی زانیاری و پیدراوهکان به شیوازی گراف (graph)
- پشت بهستن به بناغهی ویی سیمانتیک و مۆدیل کردنی پیدراوهکان
- لیک کارکردنی یه کجار زۆری زانیاری له سهر وییدا ← چاره سه ریکی باش بو ئیستای «پیدراوه په یوه نیدی دار» هکان (linked data)
- به سوودیشه بو زانیاریی زمانه وانیهکان
- ئەژماری وشه ئینگلیسییهکان به پیی یه کهم پیتیان (بهسته ر)
- کۆمه لی گرافهکانی زانیاری ده بی به بنکهی زانیاری (knowledge base)
- لیستی پشکنه رهکانی ئاسمان به یه ک کرته (بهسته ر)
- بنکه به سوودهکان:



# سه‌رچاوه زمانه‌وانییه‌کان: گرافی زانیاری و بنکه‌ی زانیاری

- نواندنی زانیاری و پیدراوه‌کان به شیوازی گراف (graph)
- پشت به‌ستن به بناغهی ویی سیمانتیک و مودیل کردنی پیدراوه‌کان
- لی‌ک کارکردنی یه‌کجار زوری زانیاری له سه‌ر وییدا ← چاره‌سه‌ریکی باش بو ئیستای «پیدراوه پهیوه‌ندی‌دار»ه‌کان (linked data)
- به‌سوودیشه بو زانیاریی زمانه‌وانی
- ئەژماری وشه ئینگلیسییه‌کان به پیی یه‌که‌م پیتیان (به‌سته‌ر)
- کۆمه‌لی گرافه‌کانی زانیاری ده‌بی به بنکه‌ی زانیاری (knowledge base)
- لیستی پشکنه‌ره‌کانی ئاسمان به یه‌ک کرته (به‌سته‌ر)
- بنکه به‌سووده‌کان:
- Wikidata (سه‌یری کوردی بکه: به‌سته‌ر)



# سه‌رچاوه زمانه‌وانییه‌کان: گرافی زانیاری و بنکه‌ی زانیاری

- نواندنی زانیاری و پی‌دراوه‌کان به شیوازی گراف (graph)
- پشت به‌ستن به بناغه‌ی وی‌بی سیمانتیک و مۆدیل کردنی پی‌دراوه‌کان
- لیک کارکردنی یه‌کجار زۆری زانیاری له سه‌ر وی‌یدا ← چاره‌سه‌ریکی باش بو ئیستای «پی‌دراوه‌ی په‌یوه‌ندی‌دار»ه‌کان (linked data)
- به‌سوودیشه بو زانیاریی زمانه‌وانی
- ئەژماری وشه ئینگلیسییه‌کان به پی‌ی یه‌که‌م پیتیان (به‌سته‌ر)
- کۆمه‌لی گرافه‌کانی زانیاری ده‌بی به بنکه‌ی زانیاری (knowledge base)
- لیستی پشکنه‌ره‌کانی ئاسمان به یه‌ک کرته (به‌سته‌ر)
- بنکه به‌سووده‌کان:
- Wikidata (سه‌یری کوردی بکه: به‌سته‌ر)
- DBpedia (سه‌یری کوردی بکه: به‌سته‌ر)



# سه‌رچاوه زمانه‌وانییه‌کان: گرافی زانیاری و بنکه‌ی زانیاری

- نواندنی زانیاری و پیدراوه‌کان به شیوازی گراف (graph)
- پشت به‌ستن به بناغهی ویی سیمانتیک و مودیل کردنی پیدراوه‌کان
- لیک کارکردنی یه‌کجار زوری زانیاری له سهر وییدا ← چاره‌سه‌ریکی باش بو ئیستای «پیدراوه پهیوه‌ندی‌دار»ه‌کان (linked data)
- به‌سوودیشه بو زانیاریی زمانه‌وانی
- ئەژماری وشه ئینگلیسییه‌کان به پیی یه‌که‌م پیتیان (به‌سته‌ر)
- کۆمه‌لی گرافه‌کانی زانیاری ده‌بی به بنکه‌ی زانیاری (knowledge base)
- لیستی پشکنه‌ره‌کانی ئاسمان به یه‌ک کرته (به‌سته‌ر)
- بنکه به‌سووده‌کان:
- Wikidata (سه‌یری کوردی بکه: به‌سته‌ر)
- DBpedia (سه‌یری کوردی بکه: به‌سته‌ر)
- ConceptNet (سه‌یری کوردی بکه: به‌سته‌ر)



# پیداوه‌کان به‌پیی فۆرماتەکان



SEROKÊ IRAQÊ ŞPÊWAZÎYA ENDAMÊ SEROKATÎYA ENCÛMENA  
 BILINDA ISLAMÎYA IRAQÊ KIR [10:30] 2010/Oct  
 /25 PNA Serokê komara Iraqê Celal Talebanî,  
 şpewaziya endamê serokatîya Encûmena  
 Bilinda Islamiya Iraqê Adil Abdulmehdî û  
 şandeya pê re kir. Herdu aliyan daxwaz  
 kirin, ku divê hukûmeta nû ya Iraqê zû pêk  
 bihê da ku pirs û şarêyên hevvelatîyan  
 çareser bibin û kar û barên gel bi rê ve  
 biçin. Di evê dîdarê de ya ku doh êvarê hat  
 sazdan, herdu aliyan bir û nerîn di bareya  
 pêkanîna hukûmeta nû ya Iraqê de bi hev  
 guherandin û behsa encamên diyaloga navbera  
 aliyên siyasî kirin.

• ناسراوترین پیناسه‌ی فۆرمات: .txt





SEROKÊ IRAQÊ ŞPÊWAZÎYA ENDAMÊ SEROKATÎYA ENCÛMENA  
 BILINDA ISLAMÎYA IRAQÊ KIR [10:30] 2010/Oct  
 /25 PNA Serokê komara Iraqê Celal Talebanî,  
 şpewaziya endamê serokatîya Encûmena  
 Bilinda Islamiya Iraqê Adil Abdulmehdî û  
 şandeya pê re kir. Herdu aliyan daxwaz  
 kirin, ku divê hukûmeta nû ya Iraqê zû pêk  
 bihê da ku pirs û şarêyên hevvelatîyan  
 çareser bibin û kar û barên gel bi rê ve  
 biçin. Di evê dîdarê de ya ku doh êvarê hat  
 sazdan, herdu aliyan bir û nerîn di bareya  
 pêkanîna hukûmeta nû ya Iraqê de bi hev  
 guherandin û behsa encamên diyaloga navbera  
 aliyên siyasî kirin.

- ناسراوترین پیناسه‌ی فۆرمات: .txt
- ✓ ساکار، خیرا و ئاسان



SEROKÊ IRAQÊ ŞPÊWAZÎYA ENDAMÊ SEROKATÎYA ENCÛMENA  
 BILINDA ISLAMÎYA IRAQÊ KIR [10:30] 2010/Oct  
 /25 PNA Serokê komara Iraqê Celal Talebanî,  
 şpewaziya endamê serokatîya Encûmena  
 Bilinda Islamiya Iraqê Adil Ebdulmehdî û  
 şandeya pê re kir. Herdu aliyan daxwaz  
 kirin, ku divê hukûmeta nû ya Iraqê zû pêk  
 bihê da ku pirs û şarêyên hevvelatîyan  
 çareser bibin û kar û barên gel bi rê ve  
 biçin. Di evê didarê de ya ku doh êvarê hat  
 sazdan, herdu aliyan bir û nerîn di bareya  
 pêkanîna hukûmeta nû ya Iraqê de bi hev  
 guherandin û behsa encamên diyaloga navbera  
 aliyên siyasî kirin.

- ناسراوترین پیناسه‌ی فۆرمات: .txt
- ✓ ساکار، خیرا و ئاسان
- × پیدراوه‌ی به‌بێ دارپێژران یان وه‌سف



```

<entry>
  <form>
    <orth>bend</orth>
    <pron>bænd</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
    <gen>f</gen>
  </gramGrp>
  <sense n="1">
    <cit type="translation" xml:lang="en">
      <quote>bond</quote>
    </cit>
    <cit type="example">
      <quote>divê em êdî li benda sibehê
        ranewestin.</quote>
      <cit type="translation" xml:lang="en">
        <quote>we shouldn't stand around
          waiting for tomorrow.</quote>
      </cit>
    </cit>
  </sense>
</entry>

```

## زمانی مارکاپی دريژه پيڊەر

### Extensible Markup Language (.xml)



```

<entry>
  <form>
    <orth>bend</orth>
    <pron>bænd</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
    <gen>f</gen>
  </gramGrp>
  <sense n="1">
    <cit type="translation" xml:lang="en">
      <quote>bond</quote>
    </cit>
    <cit type="example">
      <quote>divê em êdî li benda sibehê
        ranewestin.</quote>
      <cit type="translation" xml:lang="en">
        <quote>we shouldn't stand around
          waiting for tomorrow.</quote>
      </cit>
    </cit>
  </sense>
</entry>

```

- زمانی مارکاپی دريژه پيڊەر
- Extensible Markup Language (.xml)
- زمانی پروگرامسازيی مارکاپ هه نديک
- هيماي تايهت بو پيکهينانی به لگه به کاردينی



```

<entry>
  <form>
    <orth>bend</orth>
    <pron>bænd</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
    <gen>f</gen>
  </gramGrp>
  <sense n="1">
    <cit type="translation" xml:lang="en">
      <quote>bond</quote>
    </cit>
    <cit type="example">
      <quote>divê em êdî li benda sibehê
        ranewestin.</quote>
      <cit type="translation" xml:lang="en">
        <quote>we shouldn't stand around
          waiting for tomorrow.</quote>
      </cit>
    </cit>
  </sense>
</entry>

```

- زمانی مارکاپی دريژه پييدهر
- Extensible Markup Language (.xml)
- زمانی پروگرامسازيی مارکاپ ههنديک
- هيمای تاييهت بو پيکهينانی بهلگه بهکاردينی
- ✓ خووهسفکهرو دارپژراو



```

<entry>
  <form>
    <orth>bend</orth>
    <pron>bænd</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
    <gen>f</gen>
  </gramGrp>
  <sense n="1">
    <cit type="translation" xml:lang="en">
      <quote>bend</quote>
    </cit>
    <cit type="example">
      <quote>divê em êdî li benda sibehê
        ranewestin.</quote>
      <cit type="translation" xml:lang="en">
        <quote>we shouldn't stand around
          waiting for tomorrow.</quote>
      </cit>
    </cit>
  </sense>
</entry>

```

زمانی مارکاپی دريژه پييدهر  
Extensible Markup Language (.xml)

زمانی پروگرامسازيی مارکاپ ههنديک

هيماي تاييهت بو پيکهينانی بهلگه بهکاردينی

✓ خووهسفکههه و دارپژراو

✓ باش بو گواستنهوه و دهست وهردان له پي دراوه

```

<entry>
  <form>
    <orth>bend</orth>
    <pron>bænd</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
    <gen>f</gen>
  </gramGrp>
  <sense n="1">
    <cit type="translation" xml:lang="en">
      <quote>bond</quote>
    </cit>
    <cit type="example">
      <quote>divê em êdî li benda sibehê
        ranewestin.</quote>
    <cit type="translation" xml:lang="en">
      <quote>we shouldn't stand around
        waiting for tomorrow.</quote>
    </cit>
  </sense>
</entry>

```

زمانی مارکاپی دريژه پييدهر  
Extensible Markup Language (.xml)

زمانی پروگرامسازيی مارکاپ ههنديک

هيماي تاييهت بو پيکهينانی بهلگه بهکاردينی

✓ خووهسفکهههه و دارپژراو

✓ باش بو گواستههوه و دهسته و هردان له پي دراوه

× ههنديک جار ناخييرا و دريژ







```
{
  "firstName": "John",
  "lastName": "Smith",
  "isAlive": true,
  "age": 27,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": "10021-3100"
  },
  "phoneNumbers": [
    {
      "type": "home",
      "number": "212 555-1234"
    }
  ]
}
```

JavaScript Object Notation (.json) ●



```
{
  "firstName": "John",
  "lastName": "Smith",
  "isAlive": true,
  "age": 27,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": "10021-3100"
  },
  "phoneNumbers": [
    {
      "type": "home",
      "number": "212 555-1234"
    }
  ]
}
```

JavaScript Object Notation (.json) ●

✓ خۆوهسفکههه و دارپژراو ●



```

{
  "firstName": "John",
  "lastName": "Smith",
  "isAlive": true,
  "age": 27,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": "10021-3100"
  },
  "phoneNumbers": [
    {
      "type": "home",
      "number": "212 555-1234"
    }
  ]
}

```

## JavaScript Object Notation (.json)

✓ خۆوهسفکههه و دارپژراو

✓ باش بو گواستهوه و

دهست وهردان له پیدراوه



```

{
  "firstName": "John",
  "lastName": "Smith",
  "isAlive": true,
  "age": 27,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": "10021-3100"
  },
  "phoneNumbers": [
    {
      "type": "home",
      "number": "212 555-1234"
    }
  ]
}

```

## JavaScript Object Notation (.json)

✓ خۆوهسفکههه و دارپژراو

✓ باش بو گواستهوه و

دهست وهردان له پیدراوه

× کول و کورتتر له XML



```

{
  "firstName": "John",
  "lastName": "Smith",
  "isAlive": true,
  "age": 27,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": "10021-3100"
  },
  "phoneNumbers": [
    {
      "type": "home",
      "number": "212 555-1234"
    }
  ]
}

```

## JavaScript Object Notation (.json)

- ✓ خۆوهسفکەر و دارپژراو
- ✓ باش بۆ گواستنهوه و دهست وهردان له پیدراوه
- × کۆل و کورتتر له XML
- JSON و XML بربرهه
- هه لگواستنهوهی پیدراوهن له سهه و ییدا.



## بنکه دراوه (database)

سیستهمی نهرمه کالاً بو پیناسه کردن، چی کردن، پاراستن و دهست وهردان له پیدراوه به کار دههیندری [18].

- یه کجار خیرا و جی متمانه یه



## بنکه دراوه (database)

سیستهمی نهرمه کالاً بۆ پیناسه کردن، چۆ کردن، پاراستن و دهست وهردان له پیدراوه به کار دههیندری [18].

- یه کجار خیرا و جی متمانه یه
- ✓ زۆر بهربلاو له ئه ورۆی ویدا



## بنکه دراوه (database)

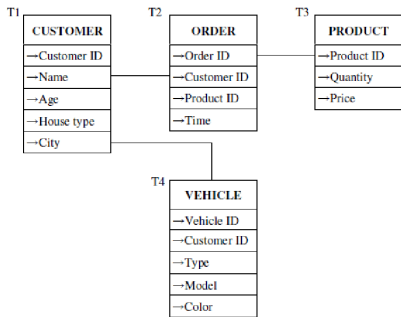
سیستهمی نهرمه کالاً بو پیناسه کردن، چی کردن، پاراستن و دهست وهردان له پیدراوه به کار دههیندری [18].

- یه کجار خیرا و جی متمانه یه
- ✓ زور بهربلاو له ئه ورۆی ویدا
- × خووه سفکه ر نین و پیکهاته کان به پیی مهنتقی په یوه ندییه وه له خشته دا ریک ده خرین



## بنکه دراوه (database)

سیستهمی نهرمه کالاً بو پیناسه کردن، چئی کردن، پاراستن و دهست وهردان له پیدراوه به کار دههیندری [18].



- یه کجار خیرا و جی متمانه یه
- ✓ زور بهربلاو له ئه ورۆی ویدا
- × خووه سفکه رنین و پیکهاته کان به پیی
- مهنتقی په یوه ندییه وه له خشته دا ریک ده خرین
- به نیو بانگترین بنکه دراوه کان  
په یوه ندیدارن (relational)  
PostgreSQL، MySQL  
و Oracle Database

## بنکه دراوه (database)

سیستهمی نهرمه کالاً بو پیناسه کردن، چئی کردن، پاراستن و دهست وهردان له پیدراوه به کار دههیندری [18].

- یه کجار خیرا و جی متمانه یه
- ✓ زور بهربلاو له ئه ورۆی ویدا
- × خووه سفکه ر نین و پیکهاته کان به پیی مهنتقی په یوه ندییه وه له خشته دا ریک ده خرین
- به نیوبانگترین بنکه دراوه کان په یوه ندیدارن (relational) PostgreSQL ، MySQL و Oracle Database



چوارچێۆه‌ی وه‌سفکردنی سه‌رچاوه‌ (Resource)  
(Description Framework (.ttl))

```
:lex_kmr_0278943883 a ontolex:
  LexicalEntry, ontolex:Word ;
ontolex:canonicalForm :
  form_kmr_0278943883 ;
rdfs:label "bend"@kmr-latn .

:form_kmr_0278943883 a ontolex:Form ;
dct:language <www.lexvo.org/page/
  iso639-3/kmr> ;
ontolex:writtenRep "bend"@kmr-latn ;
lexinfo:partOfSpeech lexinfo:noun ;
lexinfo:gender lexinfo:feminine ;
lexinfo:number lexinfo:singular ;
ontolex:sense :kmr_9120343779_sense.
```



```

:lex_kmr_0278943883 a ontolex:
  LexicalEntry, ontolex:Word ;
ontolex:canonicalForm :
  form_kmr_0278943883 ;
rdfs:label "bend"@kmr-latn .

:form_kmr_0278943883 a ontolex:Form ;
dct:language <www.lexvo.org/page/
  iso639-3/kmr> ;
ontolex:writtenRep "bend"@kmr-latn ;
lexinfo:partOfSpeech lexinfo:noun ;
lexinfo:gender lexinfo:feminine ;
lexinfo:number lexinfo:singular ;
ontolex:sense :kmr_9120343779_sense.

```

- چوارچییۆه‌ی وه‌سفکردنی سه‌رچاوه‌ (Resource) (Description Framework (.ttl))
- ✓ خۆوه‌سفکه‌ر و دارپێژراو



```

:lex_kmr_0278943883 a ontolex:
  LexicalEntry, ontolex:Word ;
ontolex:canonicalForm :
  form_kmr_0278943883 ;
rdfs:label "bend"@kmr-latn .

:form_kmr_0278943883 a ontolex:Form ;
dct:language <www.lexvo.org/page/
  iso639-3/kmr> ;
ontolex:writtenRep "bend"@kmr-latn ;
lexinfo:partOfSpeech lexinfo:noun ;
lexinfo:gender lexinfo:feminine ;
lexinfo:number lexinfo:singular ;
ontolex:sense :kmr_9120343779_sense.

```

- چوارچییۆه‌ی وه‌سفکردنی سه‌رچاوه‌ (Resource) (Description Framework (.ttl)
- خۆوه‌سفکه‌ر و دارپێژراو ✓
- زمانی ویب و پیدراوه‌ په‌یوه‌ندی‌داره‌کان ✓



```

:lex_kmr_0278943883 a ontolex:
  LexicalEntry, ontolex:Word ;
ontolex:canonicalForm :
  form_kmr_0278943883 ;
rdfs:label "bend"@kmr-latn .

:form_kmr_0278943883 a ontolex:Form ;
dct:language <www.lexvo.org/page/
  iso639-3/kmr> ;
ontolex:writtenRep "bend"@kmr-latn ;
lexinfo:partOfSpeech lexinfo:noun ;
lexinfo:gender lexinfo:feminine ;
lexinfo:number lexinfo:singular ;
ontolex:sense :kmr_9120343779_sense.

```

چوارچییۆه‌ی وه‌سفکردنی سه‌رچاوه‌ (Resource) (Description Framework (.ttl))

✓ خۆوه‌سفکه‌ر و دارپێژراو

✓ زمانی ویب و پیدراوه‌ په‌یوه‌ندی‌داره‌کان

✓ خیرا و به‌ربلاو



```

:lex_kmr_0278943883 a ontolex:
  LexicalEntry, ontolex:Word ;
ontolex:canonicalForm :
  form_kmr_0278943883 ;
rdfs:label "bend"@kmr-latn .

:form_kmr_0278943883 a ontolex:Form ;
dct:language <www.lexvo.org/page/
  iso639-3/kmr> ;
ontolex:writtenRep "bend"@kmr-latn ;
lexinfo:partOfSpeech lexinfo:noun ;
lexinfo:gender lexinfo:feminine ;
lexinfo:number lexinfo:singular ;
ontolex:sense :kmr_9120343779_sense.

```

چوارچییۆه‌ی وه‌سفکردنی سه‌رچاوه‌ (Resource)  
(Description Framework (.ttl)

✓ خۆوه‌سفکه‌ر و دارپێژراو

✓ زمانی ویب و پیدراوه‌ په‌یوه‌ندی‌داره‌کان

✓ خیرا و به‌ربلاو

× هه‌ندی‌ک جار ئه‌ندازه‌ی

پیدراوه‌کان زۆر ده‌کاته‌وه‌



```

:lex_kmr_0278943883 a ontolex:
  LexicalEntry, ontolex:Word ;
ontolex:canonicalForm :
  form_kmr_0278943883 ;
rdfs:label "bend"@kmr-latn .

:form_kmr_0278943883 a ontolex:Form ;
dct:language <www.lexvo.org/page/
  iso639-3/kmr> ;
ontolex:writtenRep "bend"@kmr-latn ;
lexinfo:partOfSpeech lexinfo:noun ;
lexinfo:gender lexinfo:feminine ;
lexinfo:number lexinfo:singular ;
ontolex:sense :kmr_9120343779_sense.

```

چوارچییۆهی وه سفکردنی سه رچاوه (Resource)  
(Description Framework (.ttl)

✓ خۆ وه سفکه ر و دارپژراو

✓ زمانی ویب و پیدراوه په یوه نیدیاره کان

✓ خیرا و بهربلاو

× هه ندیک جار ئه ندازه ی

پیدراوه کان زۆر ده کاته وه

✓ پیشنیاری W3C (بهسته ر)





```

:lex_kmr_0278943883 a ontolex:
  LexicalEntry, ontolex:Word ;
ontolex:canonicalForm :
  form_kmr_0278943883 ;
rdfs:label "bend"@kmr-latn .

:form_kmr_0278943883 a ontolex:Form ;
dct:language <www.lexvo.org/page/
  iso639-3/kmr> ;
ontolex:writtenRep "bend"@kmr-latn ;
lexinfo:partOfSpeech lexinfo:noun ;
lexinfo:gender lexinfo:feminine ;
lexinfo:number lexinfo:singular ;
ontolex:sense :kmr_9120343779_sense.

```

- چوارچێۆه‌ی وه‌سفکردنی سه‌رچاوه‌ (Resource) (Description Framework (.ttl))
- ✓ خۆوه‌سفکه‌ر و دارپێژراو
- ✓ زمانی ویب و پیدراوه‌ په‌یوه‌ندی‌داره‌کان
- ✓ خیرا و به‌ربلاو
- × هه‌ندی‌ک جار ئه‌ندازه‌ی پیدراوه‌کان زۆر ده‌کاته‌وه‌
- ✓ پێشنیاری W3C (به‌سته‌ر)
- لێره‌دا ده‌توانن له‌گه‌ڵ RDF دا زۆرتر ئاشنا بن: به‌سته‌ر.



# به کارهینانی سه رچاوه زمانیه کان



- هاوشیۆهی بنه ماکانی بهرپۆه بردنی پیدراوه زانستییهکان [19]، سه رچاوه یه کی زمانه وانی باش ده بی:



- هاوشیۆهی بنه ماکانی به پڕیۆه بردنی پیدراوه زانستییهکان [19]، سه رچاوه یه کی زمانه وانى باش ده بێ:
- Findable: بدۆزریته وه



- هاوشیۆهی بنه ماکانی به پڕیۆه بردنی پیدراوه زانستییهکان [19]، سه رچاوه یه کی زمانه وانى باش ده بی:
- **F**indable: بدۆزریته وه
- **A**ccessible: به رده ست بی



- هاوشیۆه ی بنه ماکانی به پڕیۆه بردنی پیدراوه زانستییه کان [19]، سه رچاوه یه کی زمانه وانى باش ده بی:
- **F**indable: بدۆزریته وه
- **A**ccessible: به رده ست بی
- **I**nteroperable: کاری له گه ل بکری



- هاوشیوهی بنه ماکانی بهر یوه بردنی پیدراوه زانستییه کان [19]، سهرچاوه یه کی زمانه وانی باش ده بی:
- **F**indable: بدۆزریته وه
- **A**ccessible: بهرده ست بی
- **I**nteroperable: کاری له گه ل بکری
- **R**eusable: به که لک بیته وه



- هاوشیۆهی بنه ماکانی به پڕیۆه بردنی پیدراوه زانستییهکان [19]، سه رچاوه یه کی زمانه وانى باش ده بی:
  - **F**indable: بدۆزریته وه
  - **A**ccessible: به رده ست بی
  - **I**nteroperable: کاری له گه ل بکری
  - **R**eusable: به که لک بیته وه
- گرینگایه تیی پیدراوه زمانه وانیه په یوه ندیداره کان (linguistic linked data) [20]





- هاوشیوهی بنه ماکانی بهرپوه بردنی پیدراوه زانستییه کان [19]، سهرچاوه یه کی زمانه وانى باش ده بی:
  - **F**indable: بدۆزریته وه
  - **A**ccessible: بهرده ست بی
  - **I**nteroperable: کاری له گه ل بکری
  - **R**eusable: به که لک بیته وه
- 
- گرینگایه تیی پیدراوه زمانه وانیه په یوه ندیداره کان (linguistic linked data) [20]
  - پیدراوه بناغه ی تیگه یشتن له زمان و گرینگایه تیشیان له زانسته مروییه دیجیتالییه کاند (digital humanities) هه یه.



## بەشی پراکتیک



ئەم بەشە لە خولەكە، بەشیکى پراکتیکیشی هەیه که خویندکار دەتوانی لە رینگای مالپەری  
خولەكەوه شوینی بکا.



- [1] Unsupervised cross-lingual representation learning.  
<https://ruder.io/unsupervised-cross-lingual-learning>.  
 Accessed: 2022-07-30.
- [2] Wheeler M. Thackston.  
*Sorani Kurdish—A Reference Grammar with Selected Readings*.  
 Harvard University, 2006.
- [3] Michael L Chyet and Martin Schwartz.  
*Kurdish-English Dictionary (Kurmanji)*.  
 Yale University Press, 2003.
- [4] George Miller, Christiane Fellbaum, Judy Kegl, and Katherine Miller.  
 Wordnet: An electronic lexical reference system based on theories of lexical memory.  
*Revue quebecoise de linguistique*, 17(2):181–212, 1988.
- [5] Purya Aliabadi, Mohammad Sina Ahmadi, Shahin Salavati, and Kyumars Sheykh Esmaili.  
 Towards building Kurdnet, the Kurdish Wordnet.  
 In *Proceedings of the Seventh Global Wordnet Conference*, pages 1–6, 2014.
- [6] Collin F Baker, Charles J Fillmore, and John B Lowe.  
 The Berkeley Framenet project.  
 In *Proceedings of the 17th international conference on Computational linguistics-Volume 1*, pages 86–90.  
 Association for Computational Linguistics, 1998.

- [7] Jacqueline Léon.  
Claimed and unclaimed sources of corpus linguistics.  
*Henry Sweet Society for the History of Linguistic Ideas Bulletin*, 44(1):36–50, 2005.
- [8] Guillaume Desagulier, Guillaume Desagulier, and Amboy.  
*Corpus linguistics and statistics with R*.  
Springer, 2017.
- [9] Kyumars Sheykh Esmaili, Donya Eliassi, Shahin Salavati, Purya Aliabadi, Asrin Mohammadi, Somayeh Yosefi, and Shownem Hakimi.  
Building a test collection for Sorani Kurdish.  
In *2013 ACS International Conference on Computer Systems and Applications (AICCSA)*, pages 1–7. IEEE, 2013.
- [10] Sina Ahmadi, Hossein Hassani, and Kamaladdin Abedi.  
A Corpus of the Sorani Kurdish Folkloric Lyrics.  
In *Proceedings of the 1st Joint Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL) Workshop at the 12th International Conference on Language Resources and Evaluation (LREC)*, Marseille, France, 2020.
- [11] Hadi Veisi, Mohammad MohammadAmini, and Hawre Hosseini.  
Toward Kurdish language processing: Experiments in collecting and processing the AsoSoft text corpus.  
*Digital Scholarship in the Humanities*, 35(1):176–193, 2020.

- [12] Sina Ahmadi.  
Building a corpus for the zaza–gorani language family.  
In *Proceedings of the Seventh Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2020)*, 2020.
- [13] Roshna Omer Abdulrahman, Hossein Hassani, and Sina Ahmadi.  
A Rule-Based Kurdish Text Transliteration System.  
2019.
- [14] Yaron Matras.  
Revisiting Kurdish dialect geography: findings from the Manchester Database.  
*Current issues in Kurdish linguistics*, 1:225, 2019.
- [15] Nicola Guarino and Pierdaniele Giaretta.  
Ontologies and knowledge bases.  
*Towards very large knowledge bases*, pages 1–2, 1995.
- [16] John P McCrae, Julia Bosque-Gil, Jorge Gracia, Paul Buitelaar, and Philipp Cimiano.  
The ontalex-lemon model: development and applications.  
In *Proceedings of eLex 2017 conference*, pages 19–21, 2017.
- [17] Mustafa Jarrar.  
The Arabic ontology—an Arabic wordnet with ontologically clean content.  
*Applied ontology*, (Preprint):1–26, 2021.



- [18] Thomas M Connolly and Carolyn E Begg.  
*Database systems: a practical approach to design, implementation, and management.*  
 Pearson Education, 2005.
- [19] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al.  
 The fair guiding principles for scientific data management and stewardship.  
*Scientific data*, 3(1):1–9, 2016.
- [20] Fahad Khan, Christian Chiarcos, Thierry Declerck, Daniela Gifu, Elena González-Blanco García, Jorge Gracia, Max Ionov, Penny Labropoulou, Francesco Mambrini, John McCrae, Émilie Pagé-Perron, Marco Passarotti, Salvador Ros, and Ciprian-Octavian Truica.  
 When Linguistics Meets Web Technologies. Recent advances in Modelling Linguistic Linked Open Data.  
<http://www.semantic-web-journal.net/content/when-linguistics-meets-web-technologies-recent-advances-modelling-linguistic-linked-open-0>, 2021.
- [21] Steven Bird and Gary Simons.  
 Extending dublin core metadata to support the description and discovery of language resources.  
*Computers and the Humanities*, 37(4):375–388, 2003.

