

Data Analysis for Ancient Corpora

Cody Kingham and Dirk Roorda

FAMES, Cambridge, 2019-01-31

Data Archiving and Networked Services

DANS



UNIVERSITY OF
CAMBRIDGE

background

description

mini-study

new horizons

- Put researchers in control of their data.
- Empower researchers to fully harness the data available to them.
- Encourage a new paradigm in the humanities

 AVAILABLE
DISCOUNTS

► PRODUCTS

COMPARE
COLLECTIONS

CURRENT SPECIALS

NEW RELEASES

GIFTS & DISCOUNTS

► GRAPHIC
RESOURCES

► ENGLISH STUDIES

► GREEK STUDIES

► HEBREW STUDIES

► INTERNATIONAL
LANGUAGES

ACADEMIC BUNDLES

ADD-ON BUNDLES

► UPGRADES

► PURCHASE OPTIONS

Order By:



Academic Bundle Amber - Greek Initial Bundle (Accordance 12) | \$699.80 Our Price: \$660.00

A specially priced bundle for schools that focuses on English and Greek Bibles. Additional discounts are available to certain schools with special coupons and v ...

[ADD](#)  [WISHLIST](#)



Academic Bundle Amber - Hebrew Initial Bundle (Accordance 12) | \$707.70 Our Price: \$600.00

A specially priced bundle for schools that focuses on English and Hebrew Bibles. Additional discounts are available to certain schools with special coupons and ...

[ADD](#)  [WISHLIST](#)



Academic Bundle Amber - Level 1 (Accordance 12) | \$11,330 Our Price: \$998.00

A specially priced bundle for schools that focuses on English, Greek, and Hebrew Bibles. Additional discounts are available to certain schools with special coupons and ...

Academic Bundles

Logos 8 Launch Discounts End 2/7 at Midnight

[See What's New](#)

Already have Logos? [Sign in](#) to see your

John Piper

Founder and Teacher, Desiring God



Starter

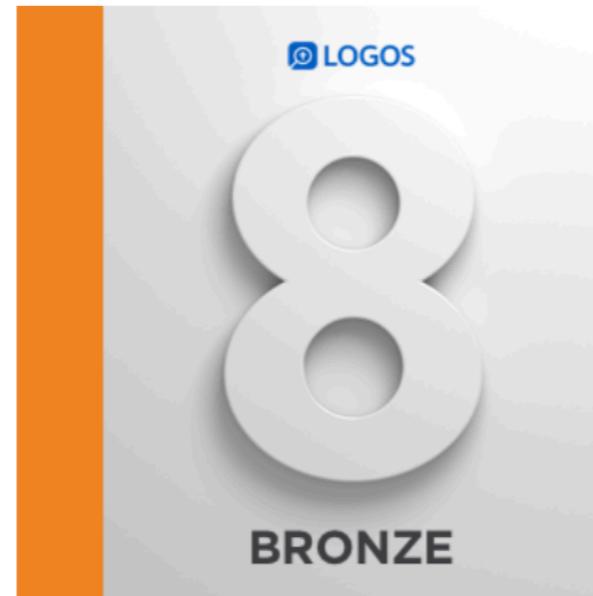
Library: 315+ resources

Features: Includes some Logos 8 features

Price: \$294.99

Your Price: \$265.49

[Learn more](#)



Bronze

Includes everything in Starter

Library: 545+ resources

Features: Includes most Logos 8 features

Price: \$629.99

Your Price: \$566.99

[Learn more](#)



Silver

Includes everything in Bronze

Library: 735+ resources

Features: Includes all Logos 8 features

Price: \$999.99

Your Price: \$899.99

[Learn more](#)



Gold

Includes everything in Silver

Library: 1,180+ resources

Features: Includes all Logos 8 features

Price: \$1,549.99

Your Price: \$1,394.99

[Learn more](#)

Products (1597) Articles (1237) Support (92)

Hebrew Bible

Search

Results per page: 20

1 2 3 4 5 6 7 8 9 10 >> ...80



HEBREW BIBLE ADVANCED ADD-ON (BHQ-6, BHS-T, MT-ETCBC, HMT-W4, HMT-W4.SYNTAX)

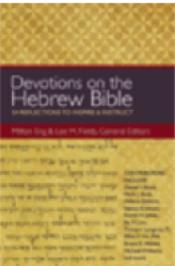
Our Price:

\$329.00

List: \$537.70

(Save \$208.70 or 37%)

ADD



DEVOTIONS ON THE HEBREW BIBLE

Price: \$16.90

ADD



HEBREW BIBLE APPARATUS

This item is not for individual sale. Please click the product link for purchase details.

The fully hypertexted apparatus to the Biblia Hebraica Stuttgartensia. See... [more](#)

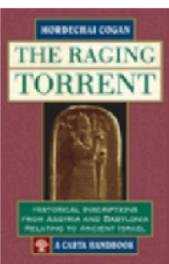


UNSETTLING GOD, AN: THE HEART OF THE HEBREW BIBLE (BRUEGEMANN)

Price: \$21.90

ADD

In the pages of the Hebrew Bible, ancient Israel gave witness to its encounter with a profound and uncontrollable reality experienced through ... [more](#)

[Products \(5\)](#) [Articles \(8\)](#) [Support \(0\)](#)
 Results per page: 

CARTA'S THE RAGING TORRENT: HISTORICAL INSCRIPTIONS FROM ASSYRIA AND BABYLONIA RELATING TO ANCIENT ISRAEL

Price: \$69.90

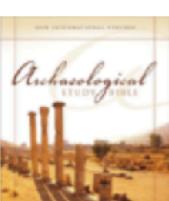
The Raging Torrent brings together all the cuneiform historical texts composed during the 9th to 6th centuries BCE that relate to Ancient Israel. This... [more](#)

[ADD](#) 

UGARITIC DATA BANK: THE TEXTS

Price: \$79.90

The Ugaritic Data Bank presents the texts excavated since 1928 at Ugarit (present day Ras-Shamra on the Syrian coast), a cosmopolitan port city in the... [more](#)

[ADD](#) 

ARCHAEOLOGICAL STUDY BIBLE NOTES

Price: \$59.90

To engage the Bible more, you need to enter the world from which it came. From the beginning of Genesis to the end of Revelation, The Archaeological... [more](#)

[ADD](#) 

Workspace

Instant Details Highlights User Notes User Tools Language Amplify

Workspaces Help

Hebrew Bible (Biblia Hebraica... LXX Rahlfs Tagged NA28 Greek NT BDB Complete Thayer Targum Pseu

Targum Pseudo-Jonathan WORDS

Verse 1 of 1268 Compare

Targum Pseudo-Jonathan

וַיֹּאמֶר אֱלֹקִים יְהִי נֹהָרָא לְאַנְהָרָא עַלְמָא וּמְנִידָה הָוֵה נֹהָרָא Gen. 1:3

וַיֹּאמֶר אֱלֹקִים יְהִי רְקִיעָה בְּמִצְעָה מִירָא וַיֹּהֵי מִפְרָשׁ בֵּין מָרֵי עַלְמָא לְבִנְיֵי מָרֵי תְּחָא Gen. 1:6

וַיֹּאמֶר אֱלֹקִים יְהִי חַדְשָׁן מִין תְּחָאֵן רְאֵישָׁתָאֵרָו מִזְלָרָעָ לְשָׂמִיא לְדוֹכְתָאֵחָר וְתְּנַגְּבָאֵרָעָ רְתְּחָמָיֵי בְּיִשְׁתָּאֵה וְהָוֵה כֵּן Gen. 1:9

וַיֹּאמֶר אֱלֹקִים חָרְבָּא אֲרָעָא דִּתְחָאֵי עִשְׁבָּא דְּבִזְרָהָא מִזְרָעָא וְאַילְעָן פְּרִירִי עַבְדִּי לִזְוָנָהָא דְּבִזְרָהָא בֵּיהֶה עַל אֲרָעָא וְהָוֵה כֵּן Gen. 1:11

וַיֹּאמֶר אֱלֹקִים יְהָוֵן נֹהָרִין בְּרְקִיעָה דְּשָׂמִיא לְאָפְרָשָׁא בֵּינֵי יִמְמָא וּבֵינֵי לִילְיָא וְיְהָוֵן לְסִימָנָן וְלְזָמָנִי מְוֻעָדִין וְלְמָמָנִי בְּהָוֵן חָשָׁבָן Gen. 1:14

▪ Search For
▪ Search In
▪ Look Up
▪ Search Back Linked Text
▪ Research
▪ My Workspaces
▪ My Groups

▪ Copy ⌘C
▪ Copy As
▪ Print Selection...
▪ Add Selection To
▪ Bookmark Selection ⌘7
▪ Edit User Note ⌘U

▪ Highlight
▪ Clear Highlight
▪ Highlight All Hits
▪ Clear Highlighted Hits

▪ Show Text As
▪ Set Text Pane Display... ⌘T

Vs Ch Bk AA Verb aph infinitiveConst to specify, (פְּרִישׁ) פרשׁ-0 אָפְרָשָׁא separate (pa.), set aside (aph.)

אלקים יהי נהורא לאנהרא עלמא ומנדיד הוה נהורא

ו אמר אלקים יהי רקייעא במציאות מיא ויהי מפרש בין מוי עלאי לבני מוי תהאי

ו אמר אלקים יתכנשווין חתאין דאישתארו מזדרע לשמייא לדוכתא חר ותתנגב ארעה דתתחמי יבישתא ויהה כן

ו אמר אלקים תרבי ארעה דיתאי עישבא דביזוריה מודרע ואילן פיריע עביד פיריע ליזניה דביזוריה ביה על ארעה ויהה כן

ו אמר אלקים יהונ נהורין ברקייעא לשמייא לאפרשתא בין ימא וביני ליליא ויהונ לסימני ולזמני מועדין ולמנני בהון חושבן

זמין ולמקדשא רישי יהחין וריש שנון עיבוריה יהחין ועיבוריה שנון ותקופות שמשא ומולד סירה ומחרוזין

ו אמר אלקים יהשון רקי מוי רחיש נפשא היהת ואופא דעתים ושרכפה על ארעה ושביל טיסיה על אויר רקייע לשמייא

ו אמר אלקים תהנפק גרגישתא דארעה נפשת בריחתא ליזנה זני דכין זני דלא דכין בערי וריחס ובריתת ארעה ליזנה ויהה כן

ו אמר אלקים למלאכיה דמשמשין קומי דאיתבריי[ז] ביום תנין לברית עלמא נעד אדם בצלמנה כדיזקננא וישלטן

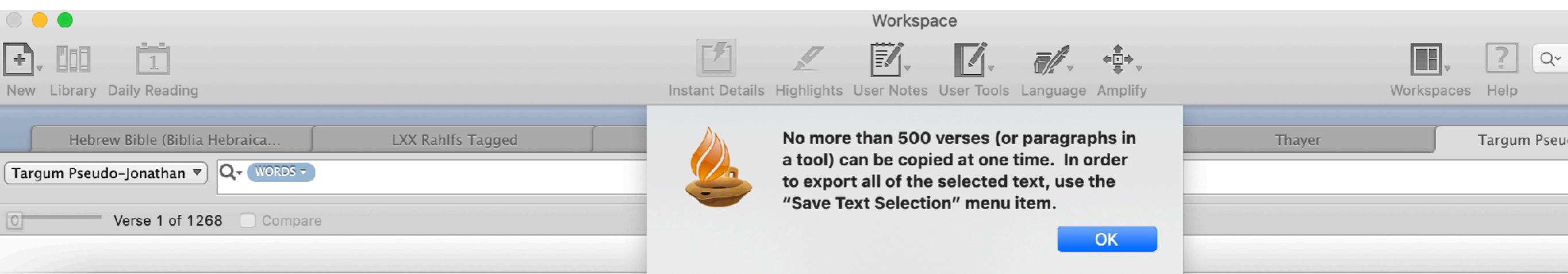
בונני ימא ובעוופא דבאoir שמייא ובעירא ובכל ארעה ובכל ריחשא דרחיש עליוי ארעה

ו בריך יהונ אלקים ו אמר להונ אלקים פושו וסנו ומלו ית ארעה בנין ובן ותקופו עלה בניכין ושלוטו בכורי ימא ובעוופא

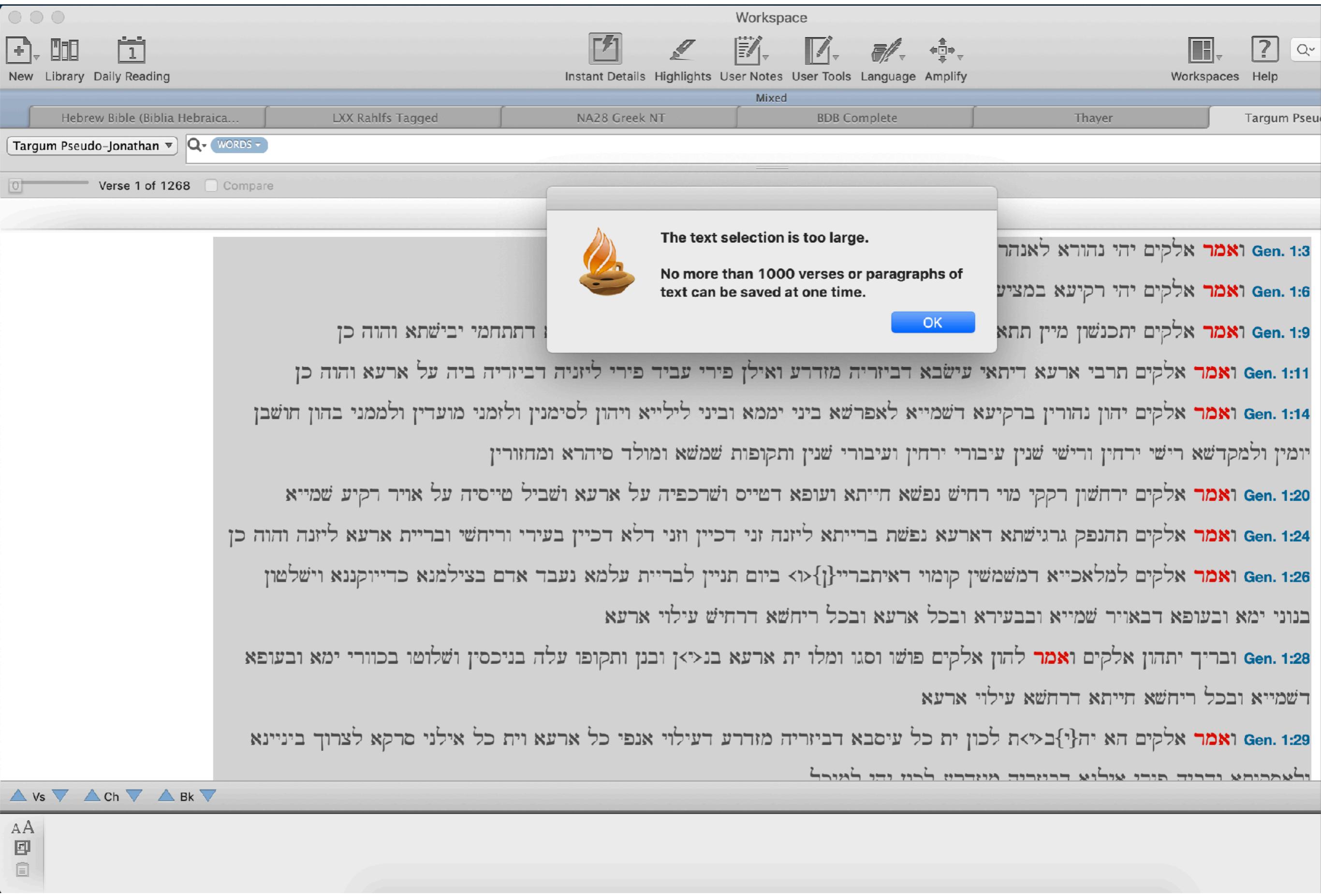
רשמייא ובכל ריחשא היהת דרחיש עליוי ארעה

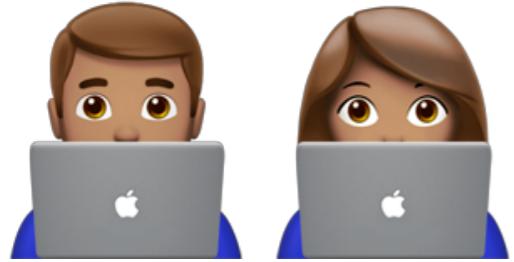
ו אמר אלקים הא יה[ז] ביגת לכו ית כל עיסבא דביזוריה מודרע דעלוי אנפי כל ארעה וית כל אילני סракא לצורך בינויא

ולאתהווא ודרהו פורי אילוא דביזוריה מידידי לריי יהי לאמירל



A screenshot of a digital Bible study application interface. At the top, there's a toolbar with icons for New, Library, Daily Reading, Instant Details, Highlights, User Notes, User Tools, Language, Amplify, Workspaces, and Help. Below the toolbar, tabs include Hebrew Bible (Biblia Hebraica...), LXX Rahlfs Tagged, Targum Pseudo-Jonathan, WORDS, Verse 1 of 1268, Compare, Thayer, and Targum Pseudo-Jonathan. A central text area displays a portion of the Hebrew Bible in Hebrew characters. A modal dialog box is open in the center, featuring a small icon of a flame in a pot. The text in the dialog reads: "No more than 500 verses (or paragraphs in a tool) can be copied at one time. In order to export all of the selected text, use the 'Save Text Selection' menu item." At the bottom right of the dialog is a blue "OK" button.

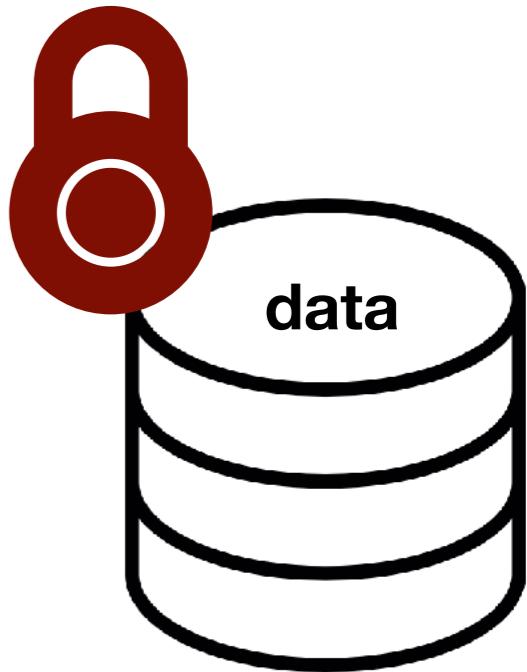




researchers

they decide

what's important
limits







ETCBC

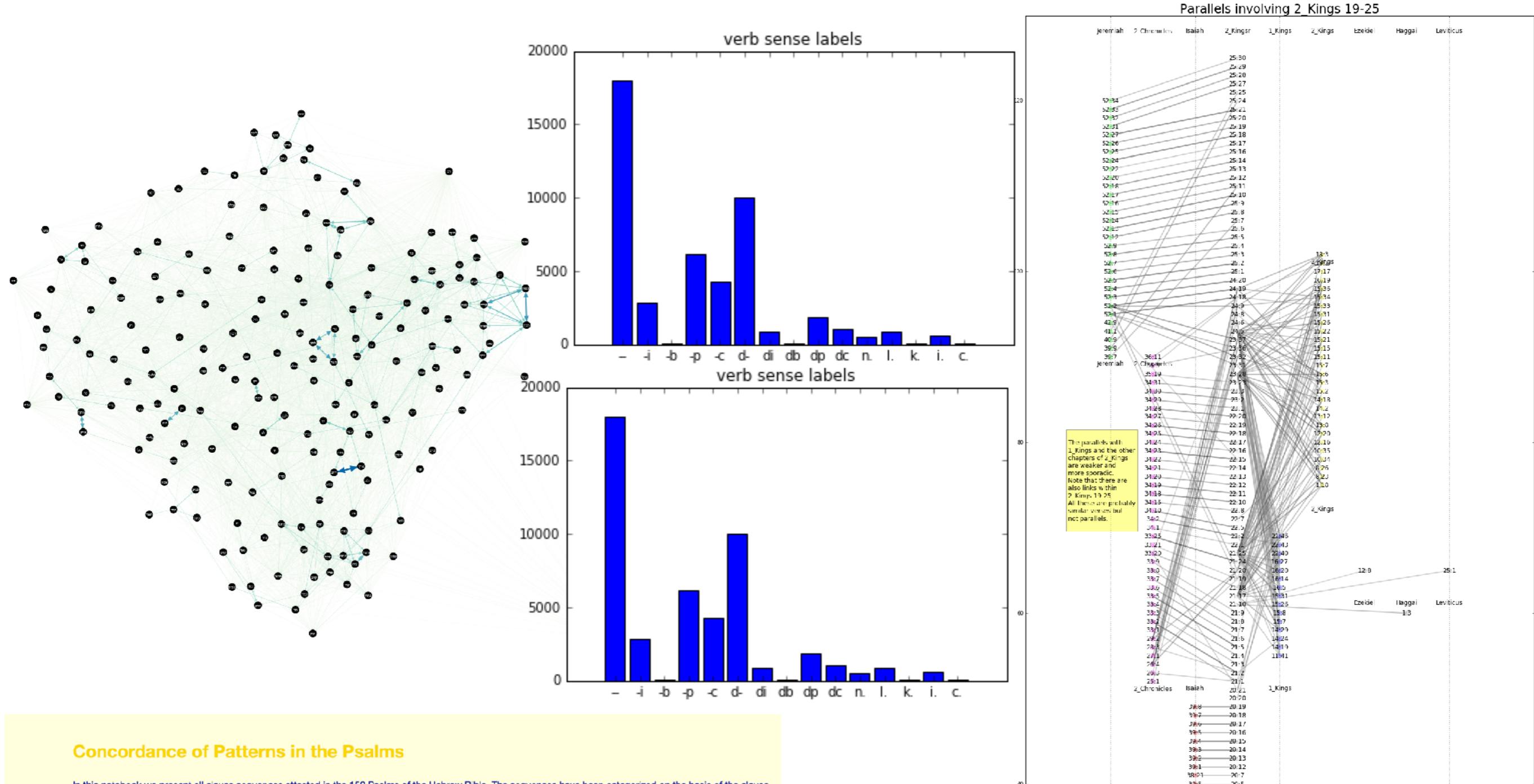
Data Archiving and Networked Services

DANS



Text-Fabric and Hebrew Data

- Free, accessible corpus annotation and analysis tool.
- Published the Amsterdam Hebrew data on Github with free, open-source license.
- Encouraged researchers to step out of their technological comfort zones.



Concordance of Patterns in the Psalms

In this notebook we present all clause sequences attested in the 150 Psalms of the Hebrew Bible. The sequences have been categorized on the basis of the clause types of the mother and daughter clause they consist of. Each category has an unique numerical identifier: 10, 20, 30, etc..

In addition, we have subdivided each category of clause sequences into three subcategories on the basis of the fact if both clauses are independent clause or not and, if so, if their subjects refer to the same or synonymous participants (see the [notebook on participants](#). The subcategorization is as follows:

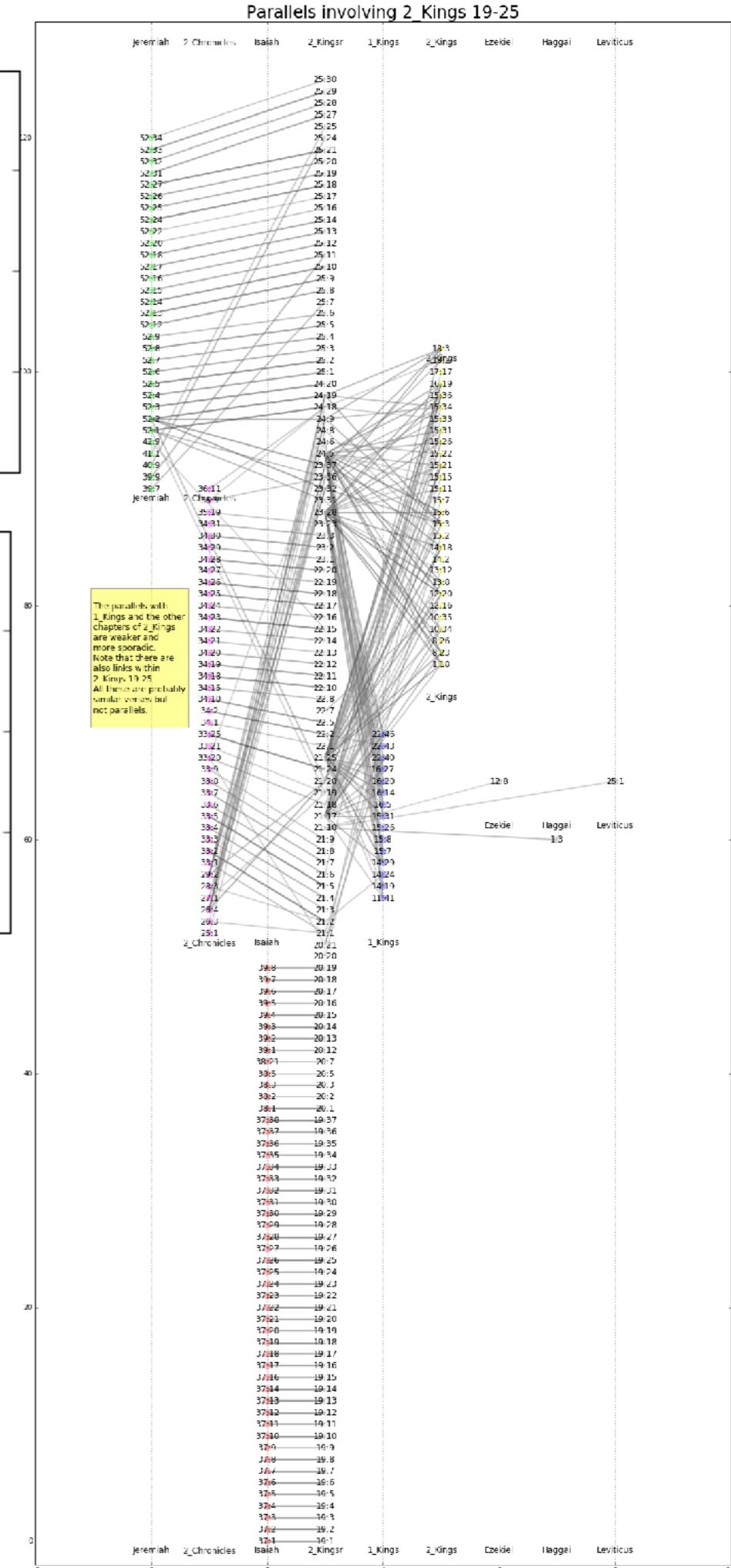
Subcat. 1	Subject continuity (SS, SR, or Sym)
Subcat. 2	No subject continuity (DS)
Subcat. 3	Daughter clause is a dependent clause (atom)(nformation with regard to subject reference is irrelevant for the assignment of functionality)

The identifiers of the individual patterns result from the summation of the numerical identifier of the pattern category (counted in tens) and the number of the subcategory to which the pattern belongs (1, 2, or 3).

Each of the individual occurrences of a pattern can be looked up by clicking on the verse label, which will bring the reader to the notebook of the appropriate Psalm.

```
In [2]: Psalms.print_patterns("Input/dataConcordanceOfPatterns.txt", 0, "yes")
```

#Pat	Vs	Ln	CITp (mouse-over for DefDiscFu)	Hebrew text (mouse over for Translation)	CCR	DefFu	Processes (hover over + sign)	FinFu	Mod- MDM	Ptpo
11	P3002.010	24	Zlm0	[חַסְכָּל]	imperat.	imperat.				SS
		25	Zlm0	[חוֹתֶם]	imperat.	imperat.				
11	P3002.011	27	Zlm0	[בְּבוֹדֵךְ] את־חוֹתֶם־[א]	imperat.	imperat.				SS
		28	Wlm0	[אָזְרָה]־[אָזְרָה]־[אָזְרָה]־[אָזְרָה]	imperat.	imperat.				
11	P3002.011	25	Zlm0	[חוֹתֶם]	imperat.	imperat.				SS
		27	Zlm0	[בְּבוֹדֵךְ]־[אָזְרָה]־[אָזְרָה]־[אָזְרָה]	imperat.	imperat.				
11	P3002.012	27	Zlm0	[בְּבוֹדֵךְ]־[אָזְרָה]־[אָזְרָה]־[אָזְרָה]	imperat.	imperat.				SS
		29	Zlm0	[שְׁלֹקְנָה]־[בְּ]	imperat.	imperat.				



A Different Vision

- Researchers are in charge of their data and set the agenda for its use.
- Researchers are empowered with the tools needed for powerful data analysis.
- Data is made open-source, freely available

Text-Fabric

- Graph model: words, phrases, etc. are “nodes,” relationships between them are edges.
- We can model complex data structures better than other methods (e.g. XML).
- All stored in easy-to-understand, plain-text files. No messy XML, SQL, etc.

&P005381 = MSVO 3, 70
#atf: lang qpc
@tablet
@obverse
@column 1
1.a. 2(N14) , SZE~a SAL TUR3~a NUN~a
1.b. 3(N19) , |GISZ.TE|
2. 1(N14) , NAR NUN~a SIG7
3. 2(N04)# , PIRIG~b1 SIG7 URI3~a NUN~a
@column 2
1. 3(N04) , |GISZ.TE| GAR |SZU2.((HI+1
2. , GU7 AZ SI4~f
@reverse
@column 1
1. 3(N14) , SZE~a
2. 3(N19) 5(N04) ,
3. , GU7
@column 2
1. , AZ SI4~f

Source data of a corpus

TEI, Markdown, ASCII, Database

Hebrew Ketiv-Qere (ETCBC)

(1:1:1:1)	bi	P	PREFIX bi+
(1:1:1:2)	somi	N	STEM POS:N LEM:{som R0OT:smw M GEN
(1:1:2:1)	{ll~ahi	PN	STEM POS:PN LEM:{ll~ah R0OT:Alh GEN
(1:1:3:1)	{l	DET	PREFIX Al+
(1:1:3:2)	r~aHoma`ni	ADJ	STEM POS:ADJ LEM:r~aHoma`n R0OT:rHm MS GEN
(1:1:4:1)	{l	DET	PREFIX Al+
(1:1:4:2)	r~aHiymi	ADJ	STEM POS:ADJ LEM:r~aHiym R0OT:rHm MS GEN
(1:2:1:1)	{lo	DET	PREFIX Al+
(1:2:1:2)	HamoduN	N	STEM POS:N LEM:Hamod R0OT:Hmd M NOM

Cuneiform Uruk (CDLI)

DEUT33,02	>C-	>;71C	1.000	>;71C-	>C-
DEUT33,02	DT	D.@73T	1.000	D.@73T	DT
DEUT33,09	BNW	B.@N@73JW	1.000	B.@N@73W	BNW
EST 01,16	MWMKN	M:MW.K@81N	1.000	M:WM.K@81N	MWMKN
EST 03,04	B-	K.:-	1.000	B.:-	B-
EST 03,04	>MRM	>@M:R@70M	1.000	>@M:R@70M	>MRM

Arabic Quran (Tanzil)

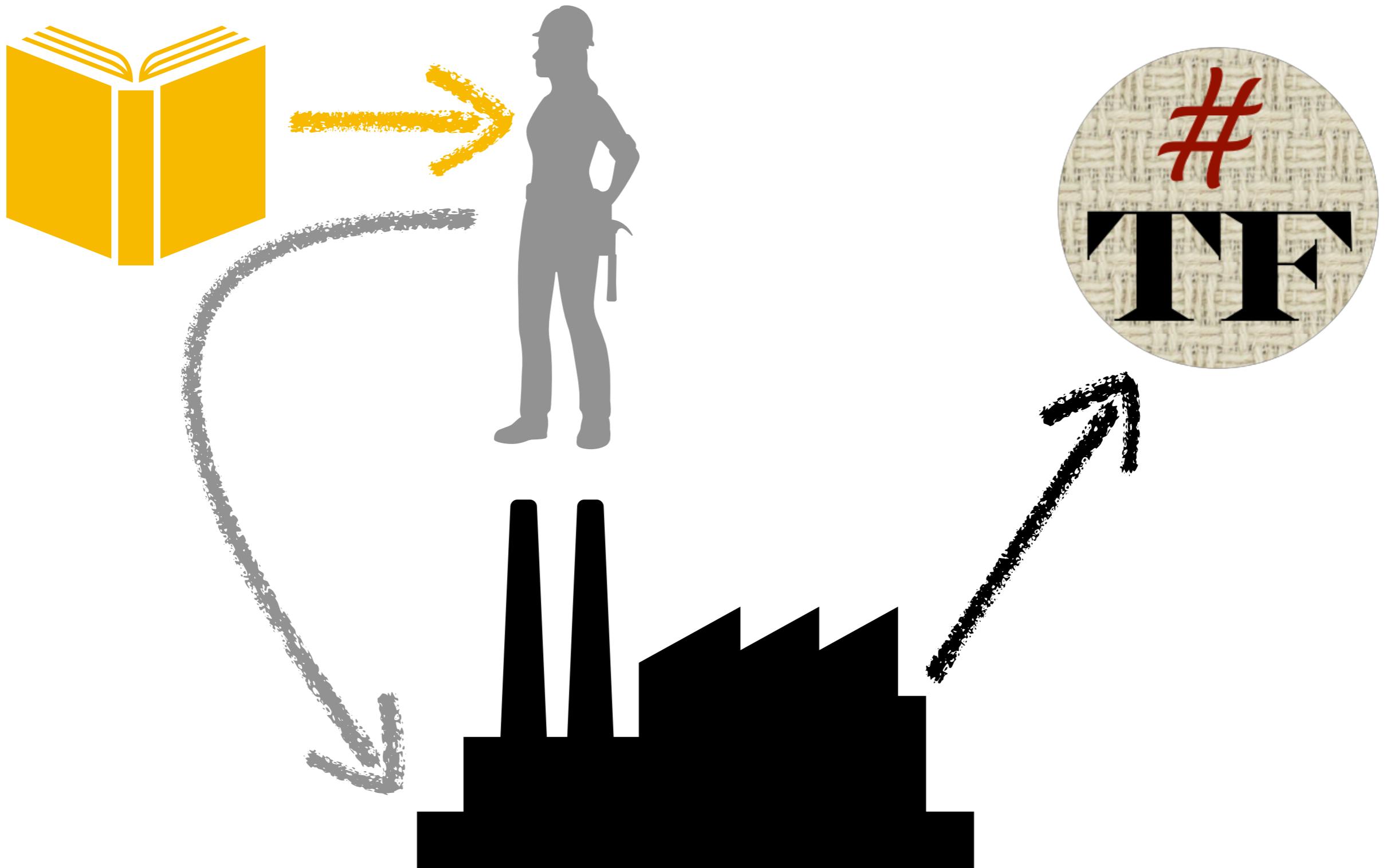
Syriac NT (Sedra database)

Data structure of TF - the IKEA spirit



Conversion to TF

TF does more than half of the work



Consider Phlebas
\$ author=Iain M. Banks

1

Everything about us,
everything around us,
everything we know [and can know of] is composed ultimately of
patterns of nothing;
that's the bottom line, the final truth.

So where we find we have any control over those patterns,
why not make the most elegant ones, the most enjoyable and good
ones,
in our own terms?

2

Besides,
it left the humans in the Culture free to take care of the things that
really mattered in life,
such as [sports, games, romance,] studying dead languages,
barbarian societies and impossible problems,
and climbing high mountains without the aid of a safety harness.

TF dataset

[banks/tf/](#)
[author.tf](#)
[gap.tf](#)
[letters.tf](#)
[number.tf](#)
[oslots.tf](#)
[otext.tf](#)
[otype.tf](#)
[punc.tf](#)
[terminator.tf](#)
[title.tf](#)

```
@node
@compiler=Dirk Roorda
@description=the letters of a word
@name=Culture quotes from Iain
Banks
@source=Good Reads
@url=https://www.goodreads.com/
work/quotes/14366-consider-phlebas
@valueType=str
@writtenBy=Text-Fabric
@dateWritten=2019-01-30T22:20:19Z
```

Everything
about
us
everything
around
us
everything
we
know
and
can
know
of
is
composed
ultimately
of
patterns
of
nothing
that's
the
bottom
line
the
final
truth
So

letters

```
@node
@compiler=Dirk Roorda
@description=the punctuation after
a word
@name=Culture quotes from Iain
Banks
@source=Good Reads
@url=https://www.goodreads.com/
work/quotes/14366-consider-phlebas
@valueType=str
@writtenBy=Text-Fabric
@dateWritten=2019-01-30T22:20:19Z
```

3 ,
6 ,
20 ;
24 ,
27 .
38 ;
45 ,
51 ,
55 ?
,

75 ,
78 ,
,

,

83 ,
88 ,
99 .

punc

otype

```
@node
@compiler=Dirk Roorda
@name=Culture quotes from Iain Banks
@source=Good Reads
@url=https://www.goodreads.com/work/quotes/14366-consider-phlebas
@valueType=str
@writtenBy=Text-Fabric
@dateWritten=2019-01-30T22:20:19Z
```

1–99	word
100	book
101–102	chapter
103–114	line
115–117	sentence

oslots

@edge
@compiler=Dirk Roorda
@name=Culture quotes from
@source=Good Reads
@url=https://www.goodread
@valueType=str
@writtenBy=Text-Fabric
@dateWritten=2019-01-30T2

100 1–99

1–55

56–99

1–3

4–6

7–9, 14–20

21–27

28–38

39–51

52–55

56

57–75

76–77, 81–83

84–88

89–99

1–27

28–55

56–99

1
Everything about us,
everything around us,
everything we know [and can know of] is composed ultimately of patterns of
nothing;
that's the bottom line, the final truth.

So where we find we have any control over those patterns,
why not make the most elegant ones, the most enjoyable and good ones,
in our own terms?

2
Besides,
it left the humans in the Culture free to take care of the things that really
mattered in life,
such as [sports, games, romance,] studying dead languages,
barbarian societies and impossible problems,
and climbing high mountains without the aid of a safety harness.

1–99	word
100	book
101–102	chapter
103–114	line
115–117	sentence

otext

```
@config
@compiler=Dirk Roorda
@fmt:text-orig-full={letters}{punc}
@name=Culture quotes from Iain Banks
@sectionFeatures=title,number
@sectionTypes=book,chapter
@source=Good Reads
@url=https://www.goodreads.com/work/quotes/14366-consider-phlebas
@writtenBy=Text-Fabric
@dateWritten=2019-01-30T22:20:19Z
```

Computing - Python - Jupyter notebooks



Feature statistics

`F` gives access to all features. Every feature has a method `freqList()` to generate a frequency list of its values, higher frequencies first. Here are the parts of speech:

```
In [19]: F.sp.freqList()
```

```
Out[19]: (('subs', 125581),  
          ('verb', 75450),  
          ('prep', 73298),  
          ('conj', 62737),  
          ('nmpr', 35696),  
          ('art', 30387),  
          ('adjv', 10052),  
          ('nega', 6059),  
          ('prps', 5035),  
          ('advb', 4603),  
          ('prde', 2678),  
          ('intj', 1912),  
          ('inrg', 1303),  
          ('prin', 1026))
```

Feature statistics

`F` gives access to all features. Every feature has a method `freqList()` to generate a frequency list of its values, higher frequencies first. Here are the parts of speech:

```
In [13]: F.pos.freqList()
```

```
Out[13]: (['pronoun', 29319),  
          ('noun', 29049),  
          ('verb', 19356),  
          ('particle', 13511),  
          ('preposition', 13006),  
          ('conjunction', 10134),  
          ('determiner', 8377),  
          ('adjective', 1961),  
          ('adverb', 1835),  
          ('prefix', 1641),  
          ('initials', 30))
```

Syriac NT

Feature statistics

`F` gives access to all features. Every feature has a method `freqList()` to generate a frequency list of its values, higher frequencies first. Here are the parts of speech:

```
In [14]: F.sp.freqList()
```

```
Out[14]: (('noun', 33717),  
          ('verb', 30441),  
          ('particle', 28486),  
          ('pronoun', 10247),  
          ('adjective', 4503),  
          ('numeral', 1620),  
          ('adverb', 434),  
          ('idiom', 192))
```

Frequency of the readings (Akkadian) and of the graphemes (Sumerian)

```
[6]: F.reading.freqList()[0:30]
```

```
[6]: (('a', 10941),  
      ('ma', 10749),  
      ('na', 7999),  
      ('i', 4620),  
      ('sza', 4447),  
      ('szu', 4270),  
      ('d', 3848),  
      ('li', 3638),  
      ('am', 3588),  
      ('ta', 3191),  
      ('u2', 3074),  
      ('ka', 3064),  
      ('ki', 2885),  
      ('um', 2811),  
      ('ni', 2681),  
      ('la', 2557),  
      ('bi', 2467),  
      ('u3', 2404),  
      ('ti', 2322),  
      ('disz', 2296),  
      ('ba', 2294),  
      ('ri', 2244),  
      ('ra', 2177),  
      ('nu', 2120),  
      ('asz', 2115),  
      ('im', 1946),  
      ('ia', 1703),  
      ('al', 1653),  
      ('di', 1541),  
      ('ku', 1505))
```

Old Babylon'

```
[7]: F.grapheme.freqList()[0:30]
```

```
[7]: (('x', 8267),  
      ('...', 1421),  
      ('ARAD', 117),  
      ('GAN2', 102),  
      ('ARAD2', 46),  
      ('BI', 46),  
      ('KI', 38),  
      ('KU', 37),  
      ('DI', 34),  
      ('GA', 28),  
      ('BU', 26),  
      ('NI', 26),  
      ('UD', 24),  
      ('ZI', 24),  
      ('KA', 23),  
      ('PI', 23),  
      ('AB', 21),  
      ('ZA', 19),  
      ('ZU', 19),  
      ('AD', 18),  
      ('NE', 18),  
      ('AH', 16),  
      ('ID', 15),  
      ('IG', 15),  
      ('SZI', 15),  
      ('DA', 13),  
      ('SU', 13),  
      ('AN', 12),  
      ('HI', 12),  
      ('IA', 12))
```

1050

MI Hierarchies. p.18n49. First Person Verbs in Narrative

Cody Kingham Southeastern Baptist Theological Seminary The Mesha Inscription and the Hebrew Bible

The Mesha Stele (MS) is written from the first person perspective, beginning with the personal pronoun 'ink and continued through 1cs verbs. The frequent use of the wayyiqtol in MS seems to indicate that the text should be marked as narrative. Are there places in

> shared ✓ 2016-01-14 21:38:45

> c never executed

> 2017 never executed

4b MQL

```
SELECT ALL OBJECTS WHERE
[book
  [clause txt = 'N'
    [word FOCUS sp = verb
      [word no = n1
```

status	results up to date
published	✓ 2016-01-14 21:38:50
results	273
words	273
modified	2016-01-13 20:09:36
executed	2016-01-13 20:09:37
engine	3.4.1.pre12

> 4 300 results up to date

Display a menu

bhsa i c 2017 4b 4 English Cody Kingham: MI Hierarchies. p.18n49. First Person Verbs in Narrative 1 → hebrew phonetic text Syntax (2/3)

273 words in 177 verses on 9 pages

Exodus 32: 33

אָלֶּה מֹשֶׁה Cmpl Subj Pred Conj נִיאָמֵר יְהוָה Frnt מַיִּ לְיִּ אֲשֶׁר חָטָאת Pred Rela מִמְחֻנָּה PreO מִסְפָּרִים Cmpl

Deuteronomy 2: 1

גַּפְנָה Pred Conj נִבְנֵנָה Cmpl Pred Conj הַמִּדְבָּרָה Adju Cmpl נִסְעָה Pred Conj כִּי אֲשֶׁר דִּבֶּר יְהוָה Pred Conj אַלְיִי Subj Pred Conj נִסְבָּה Pred Conj אֶת הַר שְׁעִיר Time Obj Pred Conj

Deuteronomy 2: 8

גַּעֲבָר Pred Conj מִאֵת אֲחִינוּ בְּנֵי עָשָׂו Pred Conj הַשְׁבִּים PreC Rela בְּשְׁעִיר Cmpl מִלְךָה עֲרָבָה מִאֵילָת וּמִעַצְיָן גָּבָר ס Pred Conj גַּפְנָה Pred Conj נִיְתְּבָה Adju Pred Conj

Deuteronomy 2: 13

גַּנְחָל זָרֶד

Deuteronomy 2: 14

זָרֶד בְּרַגְעַל

This query has been published. If that happened more than a week ago, it can be safely cited. It will not be changed anymore.

The results of this query have been obtained after the query body has been last modified

SHEBANQ

Computing - more power!

BHSA

[5]:

```
query = ...
book
    clause txt=N|?N
        word sp=verb ps=p1
...
results = A.search(query)
A.table(results, end=5)
```

0.55s 272 results

n	p	book	clause	word
1	Deuteronomy 2:1	Deuteronomy	נִפְנֵן	נִפְנֵן
2	Deuteronomy 2:1	Deuteronomy	נִפְשָׁע הַ מִדְבָּרֶה בְּרֵךְ יִמְסֹדֶף	נִפְשָׁע
3	Deuteronomy 2:1	Deuteronomy	נִסְבָּה אֶת־הָרֶה שְׁעִיר יְמִים רַבִּים ס	נִסְבָּה
4	Deuteronomy 2:8	Deuteronomy	נִעֲבָר מִאֵת אֶת־אֱתָנִי בְּנֵי־עַשְׂוֵה מִלְאִילָת הַ עֲרָבָה מִלְדָּךְ וְמִעֲצָנוֹ גָּבָר ס	נִעֲבָר
5	Deuteronomy 2:8	Deuteronomy	נִפְנֵן	נִפְנֵן

Quran

```
[9]: query = '''
aya
word pos=verb
<: word pos=noun posx=proper root=Alh
'''
```

We run it with `A.search()`:

```
[10]: results = A.search(query)
0.31s 529 results
```

```
[11]: A.table(results, start=10, end=20)
```

n	p	aya	word	word
10	2:80	وَقَالُوا لَنْ تَمَسَّنَا النَّارُ إِلَّا أَيَّامًا مَعْدُودَةً قُلْ أَتَخَذْتُمْ عِنْدَ اللَّهِ عَهْدَهُ أَمْ تُقُولُونَ عَلَى اللَّهِ 2:80 يُخْلِفَ اللَّهُ ما لَا تَعْلَمُونَ	يُخْلِفَ	اللَّهُ
11	2:90	بِئْسَمَا أَشْرَرُوا بِهِ أَنفُسُهُمْ أَن يَكْفُرُوا بِمَا أَنْزَلَ اللَّهُ مِنْ فَضْلِهِ عَلَى مَن يَشَاءُ مِنْ عِبَادِهِ 2:90 أَنْزَلَ اللَّهُ فَبَآءُوا بِغَضْبٍ عَلَى غَضَبٍ وَلِلْكَافِرِينَ عَذَابٌ مُهِينٌ	أَنْزَلَ	اللَّهُ
12	2:90	بِئْسَمَا أَشْرَرُوا بِهِ أَنفُسُهُمْ أَن يَكْفُرُوا بِمَا أَنْزَلَ اللَّهُ بَغْيًا أَن يُنْزَلَ اللَّهُ مِنْ فَضْلِهِ عَلَى مَن يَشَاءُ مِنْ عِبَادِهِ 2:90 يُنْزَلَ اللَّهُ فَبَآءُوا بِغَضْبٍ عَلَى غَضَبٍ وَلِلْكَافِرِينَ عَذَابٌ مُهِينٌ	يُنْزَلَ	اللَّهُ

[12]: A.show(results, start=10, end=14, withNodes=True)

result 10

2:80

205735

2228	2227	2226	2225	2224	2223	2222	2221	2220	2219
أياماً yawom ywm noun	إلاً <il-aA particle exceptional	نَارٌ naAr nwr noun	آل determiner	نَا pronoun personal	تَمَسَّ mas-a mss verb imperfect	لَنْ lan particle negative	وَأْ pronoun personal	قَالُ qaAla qwl verb perfect	وَ particle resumption
لَنْ lan particle negative	فَ particle resumption	عَهْدًا Eahod Ehd noun	اللهِ {ll-ah Alh noun proper	عِنْدَ Eind End adverb location	تُمْ pronoun personal	تَخَذُّ {t~axa*a Ax* verb VIII perfect	أَ particle interrogative	قُلْ qaAla qwl verb imperative	مَعْدُودَةً m-aEoduwdap Edd adjective participle
مَا maA pronoun relative	اللهِ {ll-ah Alh noun proper	عَلَى Ealay` preposition	وَنَ pronoun personal	تَقُولُ qaAla qwl verb imperfect	>ام >am conjunction coordinating	أَمْ pronoun personal	عَهْدٌ Eahod Ehd noun	اللهُ {ll-ah Alh noun proper	يُخْلِفَ >axolafu xlf verb IV imperfect
وَنَ pronoun personal	تَعْلَمُ Ealima Elm verb imperfect	لَا laA particle negative							

Syriac NT

```
[7]: query = ''  
      word lexeme_etcbc=KTB> nu=p  
      ...  
      results = A.search(query)  
      A.show(results, end=10)
```

0.12s 28 results

result 1

Matthew 22:29

ରୂପରୀତି	ରୂପ	ରୂପ	ରୂପରୀତି	ରୂପ	ରୂପ	ରୂପ	ରୂପ	ରୂପ	ରୂପ	ରୂପ	ରୂପ	ରୂପ	ରୂପ	ରୂପ
D>LH>	XJLH	WL>	KTB>	>NTWN	JD<JN	DL>		>NTWN	V<JN	LHWN	W>MR	JCW<	<N>	
>LH>	XJL>	L>	KTB>	>NT	JD<	L>		>NT	V<>	L	>MR	JCW<	<N>	
noun	noun	particle	noun	pronoun	verb	particle	pronoun	verb	particle	verb	peal	noun	verb	
nu=s	nu=s		nu=p	nu=p	peal		nu=p	peal		peal	perfect	nu=s	peal	perfect
					participle			participle						nu=s
					nu=p			nu=p						

result 2

Matthew 26:54

የሰጠና	እወል	የሰጠና	ክተክ	አሳሳሁ	ከተማ	የሰጠና
DNHW>	WL>	DHKN>	KTB>	NTMLWN	HKJL	>JKN>
HW>	WL>	HKN>	KTB>	ML>	HKJL	>JKN>
verb	adjective	particle	noun	verb	particle	particle
peal	peal		nu=p	ethpael		
imperfect	participle			imperfect		
nu=s	nu=s			nu=p		

Proper nouns

List of sign stretches that occur between um-ma and ma .

```
query = '''
line
  sign reading=um
  <: sign reading=ma
    < sign reading=ma
...
results = sorted(S.search(query))
print(f'{len(results)} results')
```

1622 results

```
results[0:5]
```

```
[(206905, 12, 13, 21),
(206909, 63, 64, 67),
(206939, 357, 358, 364),
(206957, 508, 509, 514),
(206994, 792, 793, 799)]
```

```
introNouns = collections.Counter()

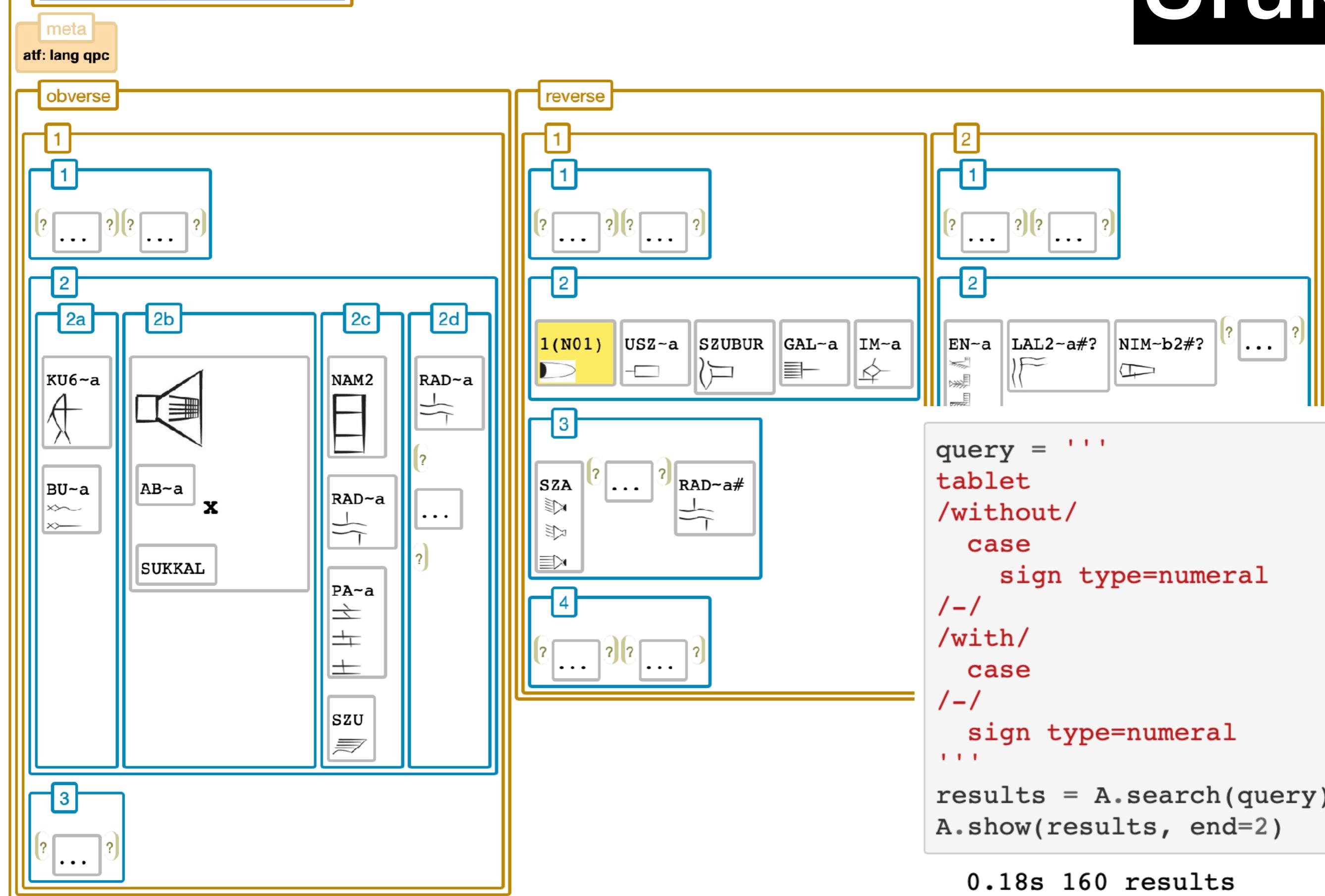
for (line, um, ma1, ma2) in results:
    betweenNodes = range(ma1 + 1, ma2)
    introNoun = ''.join(
        f'{F.reading.v(n) or F.grapheme.v(n)}{F.after.v(n)}'
        for n in betweenNodes
    )
    introNouns[introNoun] += 1
```

Old Babylon'

```
for (proper, amount) in sorted(
    introNouns.items(),
    key=lambda x: (-x[1], x[0]),
)[0:100]:
    print(f'{proper:<30} {amount:>4} x')
```

ha-am-mu-ra-bi-	108	x
szu-	83	x
szu-u2-	82	x
at-ta-	59	x
a-na-ku-	55	x
at-ta-a-	49	x
a-na-ku-u2-	36	x
szu-nu-	28	x
a-hu-um-	23	x
ha-am-mu-ra-pi2-	18	x
a-bi-e-szu-uh-	17	x
d-marduk-mu-sza-lim-	15	x
at-ti-	14	x
lu2-igi-sa6-	13	x
s,i-li2-d-utu-	13	x
am-mi-s,a-du-qa2-	12	x
sa-am-su-i-lu-na-	12	x
d-utu-na-s,i-ir-	11	x
d-iszkur-ra-bi-	10	x
d-suen-i-din-nam-	10	x

tablet P006438 AOST 109 uruk-iii



Power to you! (without the programming)

Uruk

```
~ — -bash
dirk:~ > text-fabric uruk
```

```
This is Text-Fabric 7.4.5
Using TF app uruk commit 3d71ccd1f5003889cae53378a4a93e71d17281ac
  in /Users/dirk/text-fabric-data/__apps__/uruk
Cleaning up remnant processes, if any ...
Loading data for uruk. Please wait ...
Using TF app uruk commit 3d71ccd1f5003889cae53378a4a93e71d17281ac
  in /Users/dirk/text-fabric-data/__apps__/uruk
Setting up TF kernel for uruk
Using Nino-cunei/uruk/tf/uruk - 1.0 r1.2 in /Users/dirk/text-fabric-data
Using Nino-cunei/uruk/sources/cdli/images r1.2 in /Users/dirk/text-fabric-data
Found 2095 ideograph linearts
Found 2724 tablet linearts
Found 5495 tablet photos
TF setup done.
Opening uruk in browser
Listening at port 18984
Using TF app uruk commit 3d71ccd1f5003889cae53378a4a93e71d17281ac
  in /Users/dirk/text-fabric-data/__apps__/uruk
onDocker=False
* Running on http://localhost:8104/ (Press CTRL+C to quit)
```

URUK Character table

Feature docs Search Reference

App tutorial

► Text-Fabric Browser

► Tablet pad

► Node Pad

▼ Search Pad

Enter a search template here. See the buttons on top for the docs.

Click **Go** to fetch the results.

The results of the search appear in the result list, with a positive sequence number.

► Result list

► Jobs

► Options

▼ Export

Export your results. Provide your name, title, and description (markdown is supported).

Click . The exported page opens in a new window or tab, formatted for saving as PDF. Use your browser to export this page to PDF. The PDF will contain a complete description of your work, with persistent links to the corpora and the tools, with additional metadata, and with the information you specify. It reflects the results as you see them on your screen.

Click . Additional data will be exported as files in a subdirectory corresponding to the current job:

(1) a file `resultsx.tsv` with all your query results (uncondensed). **This file can directly be opened in Excel.** Every row corresponds with a result tuple of nodes. It contains result sequence number, a verse reference to where the result occurs, and fields for the member nodes. Every member of the result tuple corresponds to a few

tablet pad

P005381

node pad

100,101,102

search pad

```
tablet
/without/
  case
    sign type=numeral
/-/
/with/
  case
/-/
  sign type=numeral
```

TF uruk-case



expand all tablets
expand all nodes
expand all results

1 current position

20 results per page

1	1
10	1
20	2
30	3
40	4
50	5
60	6
70	7
80	8
90	9
160	100

n tablet sign

► 1 P006438 tablet P006438 | 1(N01)

► 2 P002082 tablet P002082 | 1(N01)

► 3 P002082 tablet P002082 | 2(N01)

► 4 P002082 tablet P002082 | 1(N14)

► 5 P002082 tablet P002082 | 1(N01)

► 6 P002082 tablet P002082 | 1(N34)

► 7 P002120 tablet P002120 | 1(N02)

► 8 P003718 tablet P003718 | 1(N01)

► 9 P003718 tablet P003718 | 3(N01)

► 10 P003718 tablet P003718 | 3(N14)

► 11 P003718 tablet P003718 | 1(N01)

► 12 P003718 tablet P003718 | 5(N14)

► 13 P003718 tablet P003718 | 2(N01)

► 14 P003820 tablet P003820 | 1(N01)

► 15 P003820 tablet P003820 | 1(N01)

► 16 P004461 tablet P004461 | 2(N01)

► 17 P004461 tablet P004461 | 3(N39~a)

► 18 P004461 tablet P004461 | 1(N24)

**TF**

tablet pad
P005381

node pad
100,101,102

search pad
 tablet
 /without/
 case
 sign type=numeral
 /-/
 /with/
 case
 /-/
 sign type=numeral

Help

Jobs

Options

Export

n tablet sign

expand all tablets
 expand all nodes
 expand all results

1 current position
20 results per page

1 1 10 2 20 3 30 4 40 5 160 100 50 6 7 8 9 90

tablet P006438 AOST 109 uruk-iii

meta
atf: lang qpc
type=meta

obverse
type=obverse

1 1 ... type=ellipsis ... type=ellipsis

2 2a KU6-a type=ideograph
A 2b AB-a type=ideograph
X 2c NAM2 type=ideograph
| 2d RAD~a type=ideograph
| RAD~a type=ideograph
| PA-a type=ideograph
| SZU type=ideograph

reverse
type=reverse

1 1 ... type=ellipsis ... type=ellipsis

2 2 1(N01) type=numeral USZ-a type=ideograph SZUBUR type=ideograph

3 SZA type=ideograph ... type=ellipsis RAD-a# type=ideograph

4 ... type=ellipsis ... type=ellipsis

tablet P002082 1(N01)

tablet P002082 2(N01)

Display a menu

Mini-Study: Atnachs and Phrase Divisions

- How often do atnach accents disagree with the ETCBC phrase divisions?
- Why?

Sharing and re-using data

Text-Fabric has been developed by a DANS-employee
as a consequence:

Data export is built in 

Provenance tracking is built in 

Redistribution of newly created data is built in 

Data Archiving and Networked Services



sharing #1: GitHub & NBviewer

work done in a Jupyter Notebook inside a GitHub repository
is very *sharable*

[Code](#) [Issues 0](#) [Pull requests 0](#) [Projects 0](#) [Wiki](#) [Insights](#) [Settings](#)

Branch: master

[primers / oldbabylonian / OB-primer1.ipynb](#)[Find file](#) [Copy path](#) justincalejohnson2 Add files via upload

2b875a0 a day ago

1 contributor

1286 lines (1285 sloc) | 41.7 KB

[Code](#) [Raw](#) [Blame](#) [History](#) [!\[\]\(2f1827fcacb2b2cce95d0ea67f936e78_img.jpg\)](#) [!\[\]\(e55e6ddcd36d021d4bc319aea739016a_img.jpg\)](#) [!\[\]\(c0f80464aa245d104e9ca61fde665d57_img.jpg\)](#)

Old Babylonian Letters (AbB) primer 1

This is a primer for using Text-Fabric to analyse Old Babylonian letters. For instructions on installing Text-Fabric on your own computer, see:

<https://annotation.github.io/text-fabric/>

Text-Fabric has been used to work on a number of ancient scripts and languages, including Proto-Cuneiform, Biblical Hebrew, Syriac, Quranic Arabic, and others. A number of tutorials are available at Dirk Roorda's website:

<https://github.com/annotation/tutorials>

In this tutorial we are using approximately half of the corpus of Old Babylonian letters published in *Altbabylonische Briefe in Umschrift und Übersetzung*, which is abbreviated as AbB. The first volume was published by F.R. Kraus in 1964, the most recent by K.R. Veenhof in 2005. These transliterations derive from CDLI and can be downloaded by searching under "AbB" in both "Primary Publication" and "Secondary Publication" and combining the results.

First, we need to import Text-Fabric and ...

In [1]:

```
import os
import collections
import re
from tf.fabric import Fabric
https://github.com/Nino-cunei/primers/blob/master/oldbabylonian/OB-primer1.ipynb
```

sharing #2: Export from TF-browser

[BHSA](#) [Character table](#) [Feature docs](#) [Search Reference](#) [App tutorial](#)

Job: ministudy
Author: Cody Kingham
Created: 2019-02-01T13:13:28+01:00
Data: BHSA = Biblia Hebraica Stuttgartensia Amstelodamensis

version	c
release	v1.6
download	etccbc/bhsa v:c (rv1.6)
DOI	10.5281/zenodo.1007624

Data: Phonetic Transcriptions

version	c
release	1.2
download	etccbc/phono v:c (r1.2)
DOI	10.5281/zenodo.1007636

Data: Parallel Passages

version	c
release	1.2
download	etccbc/parallels v:c (r1.2)
DOI	10.5281/zenodo.1007642

Tool: Text-Fabric 7.4.5 [10.5281/zenodo.592193](#)
TF App: bhsa
commit [7f353d587f4befb6efe1742831e28f301d2b3cea](#)



Atnach ministudy

Cody Kingham

In this study, we will compare the athnach divisions of the Tiberian vocalization of the Hebrew Bible against the phrase divisions of the Eep Talstra Centre for Bible and Computer. In doing so, we want to compare the differences between the logical groupings of the two traditions.

n	phrase_atom	word	word
► 1	Genesis 1:21	וְנֶרֶלִים	וְנֶרֶלִים
► 2	Genesis 8:18	וְגַתָּה	וְגַתָּה
► 3	Genesis 9:9	אַתָּה כֹּהֵן	אַתָּה כֹּהֵן
► 4	Genesis 10:2	וְמִנְגָּדָה וְמִנְגָּדָה	וְמִנְגָּדָה וְמִנְגָּדָה
		תְּבִלָּה	תְּבִלָּה
► 5	Genesis 10:4	כְּפָתִים	כְּפָתִים
► 6	Genesis 10:22	וְאַשְׁר וְאַשְׁר	וְאַשְׁר וְאַשְׁר
		אֲרָם:	אֲרָם:
		עִילָּם	עִילָּם
		וְאַשְׁׁוּר	וְאַשְׁׁוּר
		וְאַרְפָּכָשָׂר	וְאַרְפָּכָשָׂר
		וְאַרְמָם	וְאַרְמָם
		וְאַלְיָה	וְאַלְיָה

sharing #3: Zenodo

zenodo uruk  Upload Commu 

Found 3 results.

< 1 >

Sort by: Best match

November 11, 2018 (1.2) Software Open Access

Nino-cunei/uruk: Renamed release binaries

Dirk Roorda; Cale Johnson;
Cuneiform corpora in Text-Fabric
Uploaded on November 11, 2018
2 more version(s) exist for this record

December 21, 2011 (7702) Journal article Open Access

Mining and Visual Management of XML-Based Image Collections

Khalil Shihab; Nida Al-Chalabi;
This article describes Uruk, the virtual museum of Iraq that we developed for visual exploration and management of large collections. The system largely exploits the loosely-structured hierarchy of XML documents that provides a flexible representation method to store semi-structured or unstructured data.
Uploaded on August 7, 2018

September 3, 2016 (10005710) Journal article Open Access

A Multi-Modal Virtual Walkthrough of the Virtual Past and Present Based on Panoramic View, Crowd Simulation and Acoustic Heritage on Mobile Platform

Lim Chen Kim; Tan Kian Lam; Chan Yi Chee; 

Versions

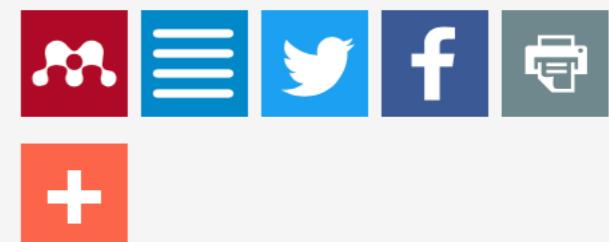
Version 1.2 Nov 11, 2018
10.5281/zenodo.1482791

Version 1.1.0 Jul 4, 2018
10.5281/zenodo.1304911

Version v1.0.0 Mar 7, 2018
10.5281/zenodo.1193842

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.1193841](https://doi.org/10.5281/zenodo.1193841). This DOI represents all versions, and will always resolve to the latest one. [Read more](#).

Share



Cite as

Dirk Roorda, & Cale Johnson. (2018, November 11). Nino-cunei/uruk: Renamed release binaries (Version 1.2). Zenodo. <http://doi.org/10.5281/zenodo.1482791>

sharing #4: Create new features

- [etcbc/valence/tf](#) : the results of the *verbal valence* work of Janet Dyk in the SYNVAR project;
- [etcbc/lingo/heads/tf](#) : head words for phrases, work done by Cody Kingham;
- [ch-jensen/Semantic-mapping-of-participants/actor/tf](#) : participant analysis in progress by Christian Høygaard-Jensen;
- [cmerwich/bh-reference-system/tf](#): participant analysis in progress by Christiaan Erwich;
- or whatever you have in the making!
- **HINT: semantic/fuzzy/plurality for collective nouns (Chip Hardy?)**



↔ Code

! Issues 0

Pull requests

Branch: master ▾

lingo / easter / tf / c /



dirkroorda ordinary

..

📄 egg.tf

b^ə
prep
in
egg=Everything

ši^{tək'ā}
verb
sit
qal
inf_c
egg=about

b^ə
prep
in
egg=us,
vêt'ekā
subs
house
egg=everything

word egg

clause Coor InfC

phrase Conj CP

û
conj
and
egg=around

phrase PreS VP

v^ə
prep
in
egg=us,
lekt^{ək'ā}
verb
walk
qal
inf_c
egg=everything

phrase Cmpl PP

va
prep
in
egg=we
art
the
egg=know
dd'erek
subs
way
egg=and

clause Coor InfC

phrase Conj CP

'û
conj
and
egg=can

phrase PreS VP

v^ə
prep
in
egg=know
šokb^{ək'ā}
verb
lie down
qal
inf_c
egg=is

clause Coor InfC

phrase Conj CP

û
conj
and
egg=composed

phrase PreS VP

v^ə
prep
in
egg=ultimately

qûm'ekā .
verb
arise
qal
inf_c
egg=of



Open Science Rocks

thank you

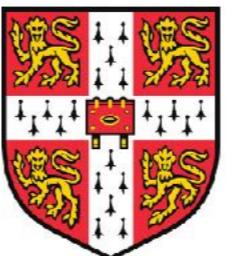
Cody Kingham codykingham@icloud.com

Dirk Roorda dirk.roorda@dans.knaw.nl



Data Archiving and Networked Services

DANS



UNIVERSITY OF CAMBRIDGE