A. De Santis, I. Rossi
Scuola Normale Superiore di Pisa

Is still Arabia at the margins of digital epigraphy?

Challenges in the digitization of the pre-Islamic inscriptions in the project DASI

6th EAGLE International Event
Off the beaten track. Epigraphy at the borders

Bari, September 24th to 25th 2015

DASI - Digital Archive for the Study of pre-Islamic Arabian Inscriptions

- **Grant:** ERC – Advanced Grant, 7th Framework programme
- **Duration:** 5 years (2011-2016)
- **Principal Investigator:** Alessandra Avanzini
- **Host institution:** University of Pisa – Dip. Civiltà e Forme del Sapere
- **Additional participant:** Scuola Normale Superiore di Pisa
- **Objects:** pre-Islamic inscriptions from the Arabian peninsula

Ancient North Arabian

Aramaic

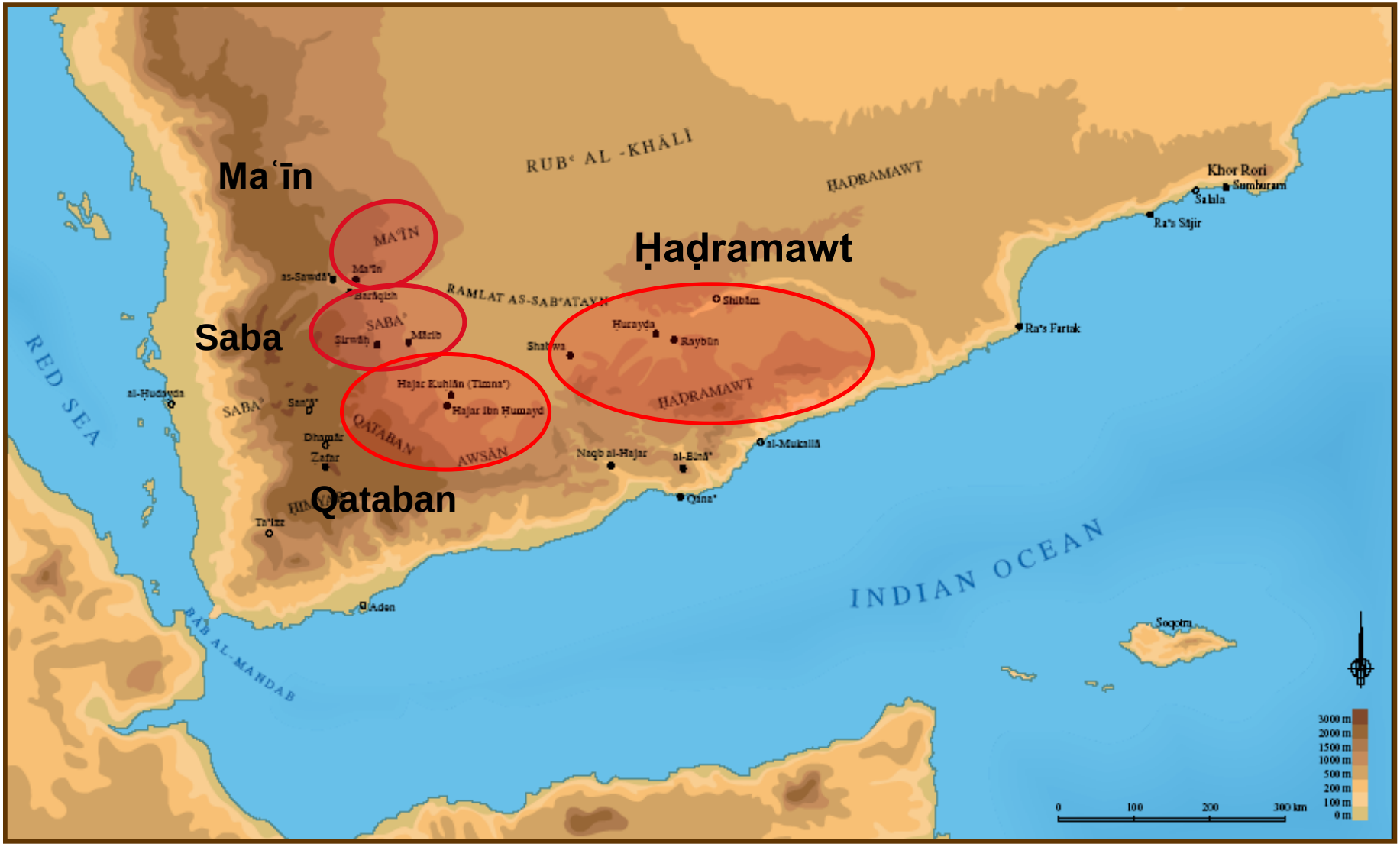
Ancient South Arabian

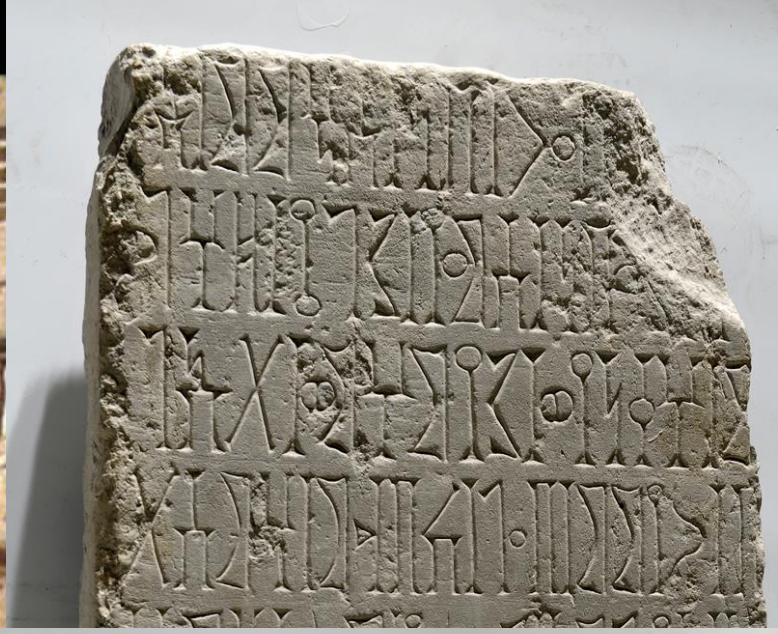
A. De Santis I.Rossi

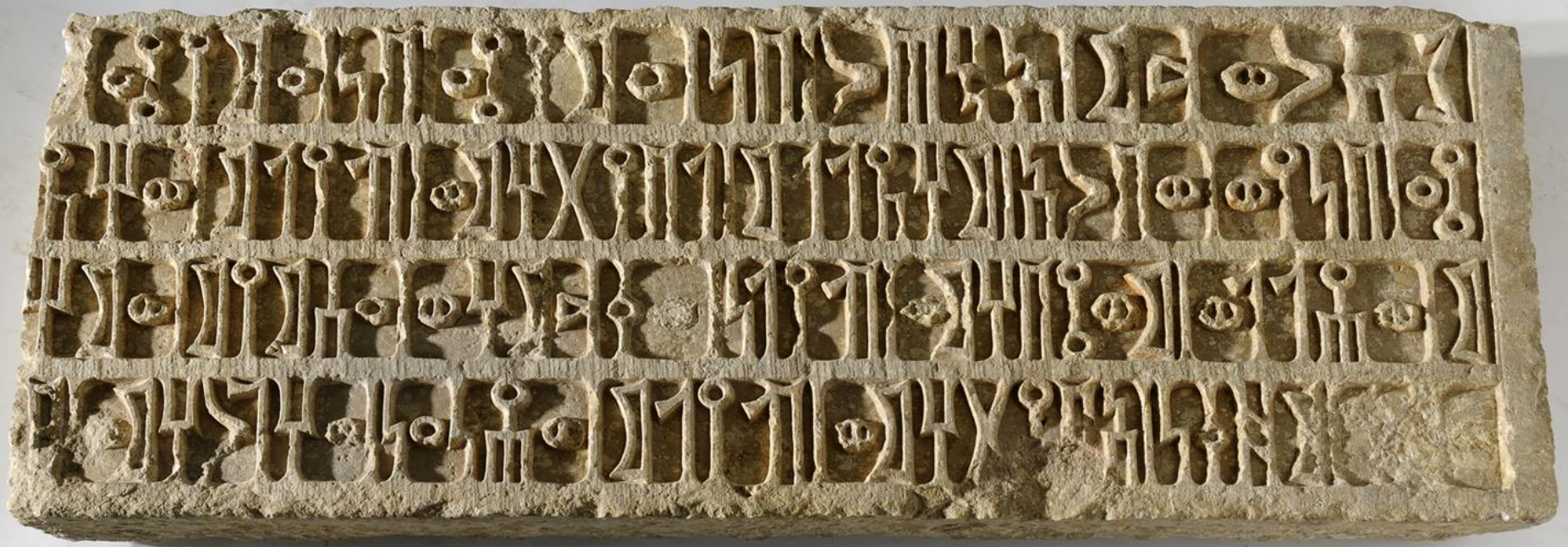
MEDINA - ***Mediterranean Network for the valorization and fruition of the inscriptions preserved in museums***

funded by EU with the ENPI – CBCMed programme:

- South Arabian collection of the Museo Nazionale d'Arte Orientale di Roma
- Phoenician inscriptions of the National Museum of Beirut
- Nabataean inscriptions of the Yarmouk University Museum


Maʿīn
Saba
Qataban
Ḥaḍramawt
RUBʿ AL -KHĀLĪ
ḤAḌRAMAWT
RAMLAT AS-SABʿATAYN
MAʿĪN
SABAʾ
QATABAN
AWSĀN
ḤIMYAR
RED SEA
INDIAN OCEAN
RAB AL-MANDAB
Khor Rori
Sumhuram
Salala
Ra's Sājir
Ra's Fartak
Shibām
Ḥurayda
Raybūn
Shabwa
Ma'īn
as-Sawdā'
Barāqish
Sirwāḥ
Mārib
Hajar Kuhlān (Timna')
Hajar ibn Ḥumayd
al-Ḥudayda
San'ā'
Dhamār
Ẓafar
Ta'izz
Aden
Naqb al-Hajar
al-Binā'
Qana'
al-Mukallā
Soqotra
3000 m
2000 m
1500 m
1000 m
500 m
200 m
100 m
0 m
0
100
200
300 km





Is still Arabia at the margins of digital epigraphy?

6th EAGLE International Event
September 24th to 25th, 2015

A. De Santis I.Rossi

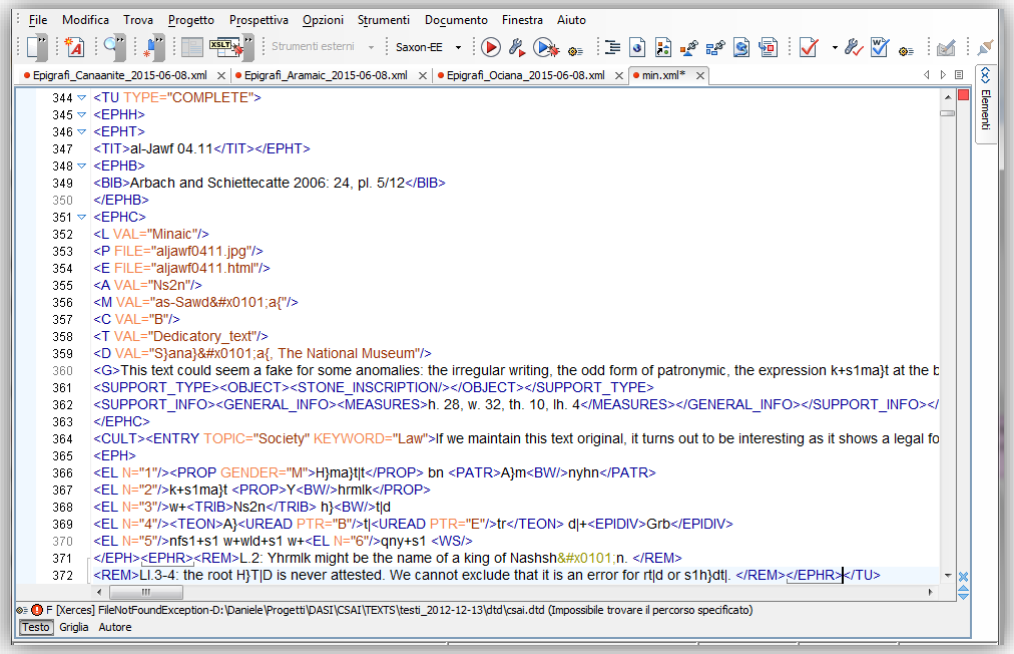
DASI – Specific objectives

- continuing the mass digitization of the pre-Islamic Arabian epigraphic heritage;
- enhancing historic and cultural knowledge of Ancient Arabia;
- strengthening the linguistic study of texts (reconstruction of lexicon and grammar)

The experience of CSAI

Website's home

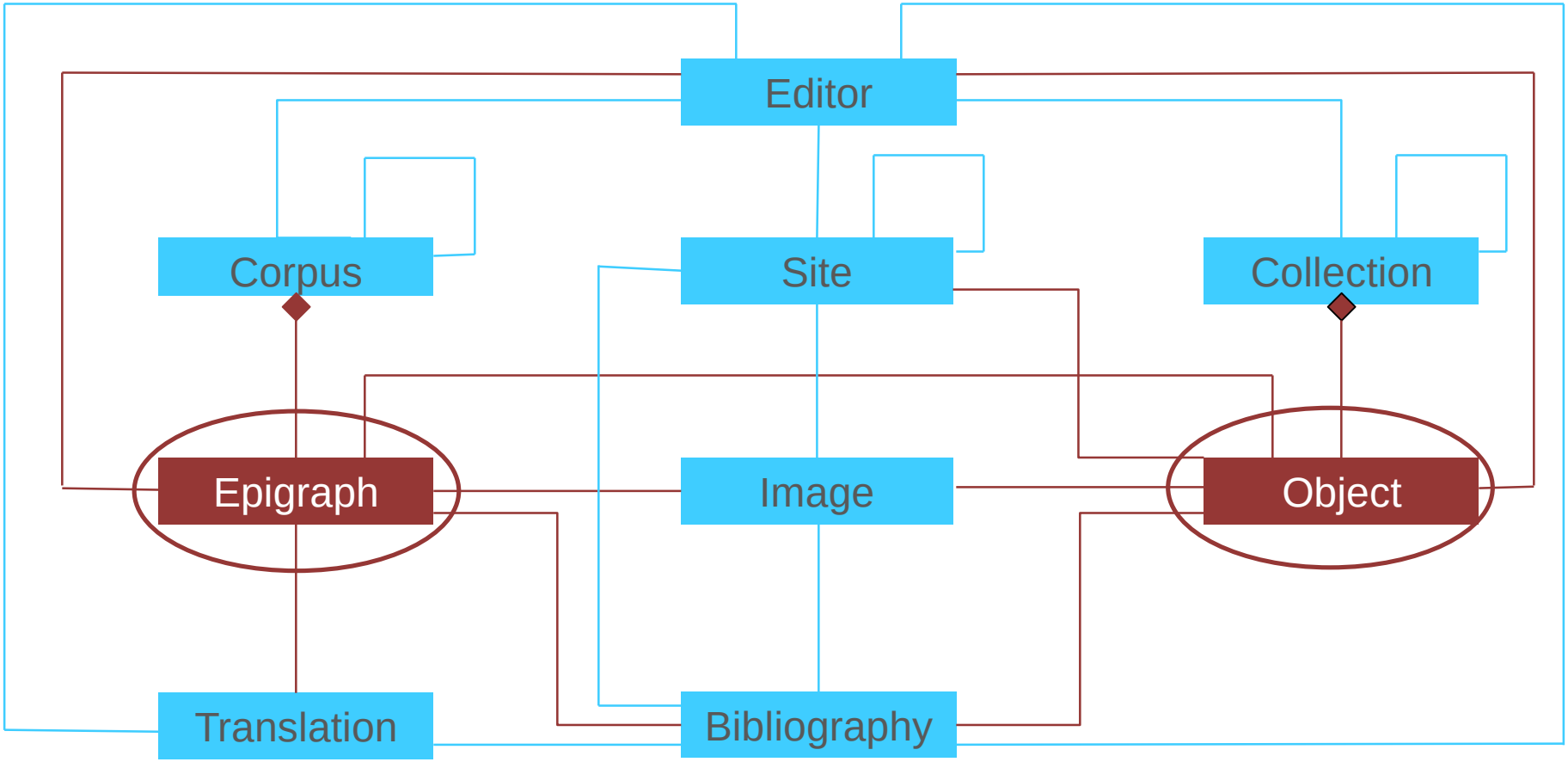
Encoding method



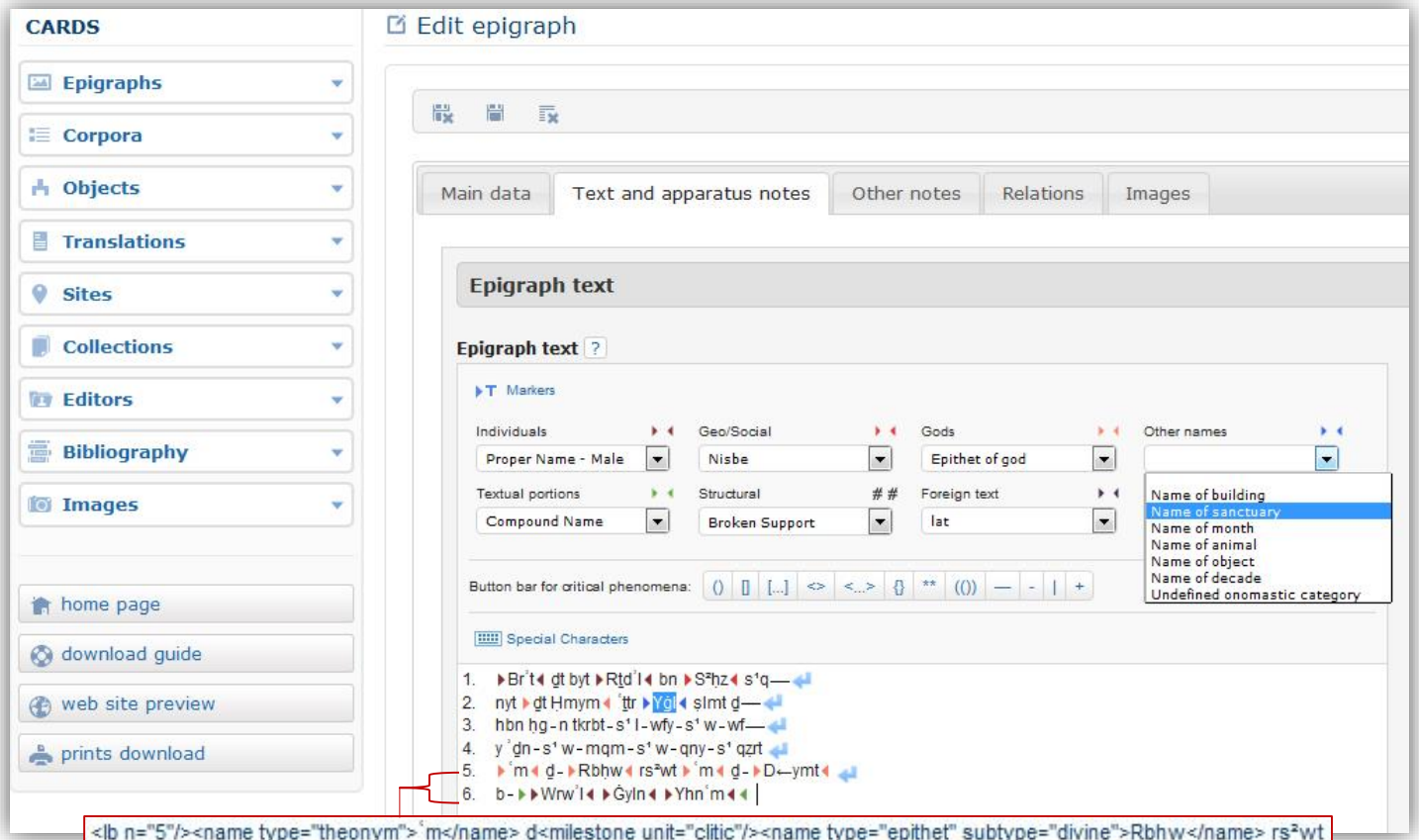
DASI SYSTEM

- **Relational database**
  - metadata
  - XML text
- **Data entry**
- **Front end**
  - content of the database is extracted in XML by the web service XMLManager, in order to construct the dynamic sections of the front end

DASI Conceptual model



“Epigraph” with DASI textual editor module




```
<lb n="5"/><name type="theonym">ʿm</name> ḏ<milestone unit="clitic"/><name type="epithet" subtype="divine">Rbḫw</name> rs²wt
<name type="theonym">ʿm</name> ḏ<milestone unit="clitic"/><name type="epithet" subtype="divine">D<milestone unit="turnSide"
rend="l"/>ymt</name>
<lb n="6"/>b<milestone unit="clitic"/><rs type="complex"><persName type="royal" subtype="withoutTitle" role="king"
n="1">Wrwʾl</persName> <persName type="royal" subtype="withoutTitle" role="king" n="2">Ġyln</persName> <persName
type="royal" subtype="withoutTitle" role="king" n="3">Yhnʿm</persName></rs> <g type="wordSeparator">|</g>
```

DASI – Encoding of text

- structural phenomena (line breaks and in-word line breaks)
- phenomena concerning the relationship between text and support (broken support, turn of side)
- transcription phenomena (ambiguous and lost characters, lacunae, text omitted)
- editorial interventions (restored text, superfluous characters, corrections)
- onomastics (names of individuals and social groups, toponyms, etc.)
- textual portions (signature, eponyms, etc.)
- symbols (word divider, paraph)
- grammatical phenomena

“Object” entity

Object identification

Status	approved
User	(import)
Work group	Csai
Object id	518
Object title	CIAS 47.11/o 1/F 72

Support description

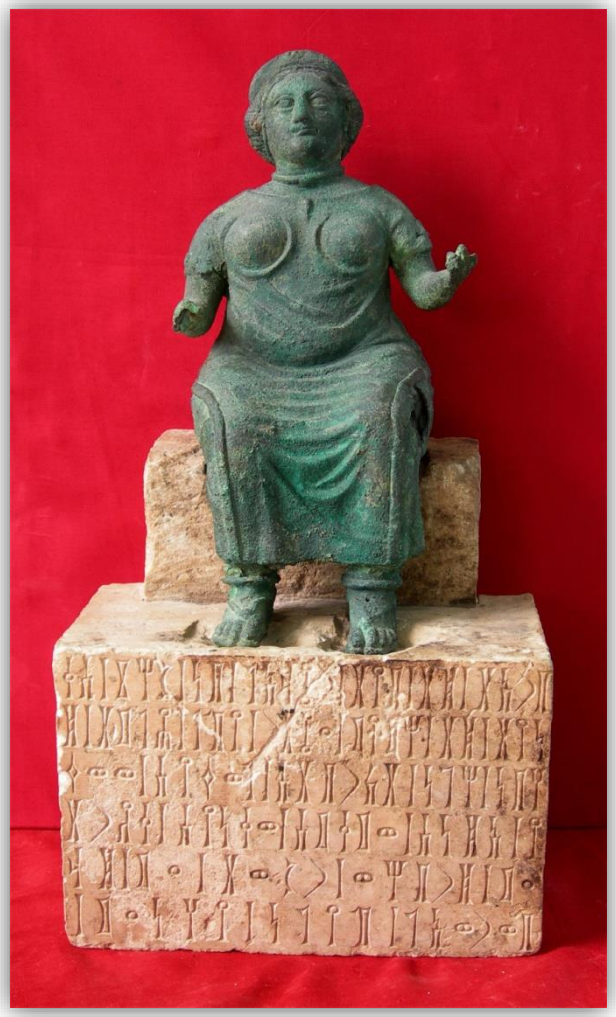
Support type	Object » Artefact » Base » Of statue with dedicatory inscription
Measures	Stone base: h. 17, w. 26, th. 5.8; bronze statuette: h. 18

▲ Materials (Number of records: 2)

Material	Bronze
Material	Stone

▲ Decorations (Number of records: 1)

Decoration Id	190
Figurative subject	Human figure
Human gender	Woman
Part of human body	Complete figure
Human gesture(s)	Sitting, Arms stretching out
Human clothes and attributes	Short sleeved tunic, Sandals
Human jewellery	Diadem, Necklace, Ankle band



“Site” entity

Site identification	
Status	approved
User	()
Site id	67
Modern name	Hajar Kuḥlān
Ancient name	Tmnʿ

Site data	
Latitude	15° 01′ 11″
Longitude	45° 48′ 16″
Country	Yemen
Governorate	Shabwa
Geographical area	wādī Bayḥān
Chronology	Probes carried out by the Italian Mission have revealed that the settlement has been inhabited since the beginning of the 1st millennium BC. A hiatus has been detected and dated between the 6th and the 5th centuries BC, while the apogee was during the 5th – 1st centuries BC. During the 1st century AD Timnaʿ suffered two major destructions, all archaeologically documented (fire traces) and dated thanks to the finding of imported pottery (mainly Arretine ware and Near Eastern Sigillata). At the second half of the 1st century AD the site has probably already lost its capital status, which was transferred to the nearby site of Hajar ibn Ḥumayd. The kingdom was soon afterwards definitely conquered by Ḥaḍramawt.
General description	Extension: 23 ha. The site was encircled by a defence wall measuring about 680 x 420 m, and whose length has been estimated in 1850 m. The mound has a height of 20 m in respect to the surrounding valley (to be partly attributed to a natural elevation). It has remained abandoned since its last destruction in antiquity, and only its SE edge is now occupied by a modern settlement.
Classical sources	Strabo , Geographia, 16.4.2 (Tamna) [citing Erathostenes of Cyrene] Plinius the Elder, Naturalis Historia, 6.32.153; 12.32.64 (Thomna) Claudius Ptolemaeus, Geographia, 6.7.37 (Thumna) [?]
Identification	1895: E. Glaser
Travellers	1899: G.W. Bury 1936: S. Perowne

City walls and gates

South-West gate: Two massive bastions (11 x 5 m) frame the entrance. They have both an ample vane at their interior and they have been realized using imposing granite blocks (e.g. 1.70 x 2.80 m) which have been also incised with important legal inscriptions (e.g. RES 3878, 3879) or recalling construction activities. The passage is 5.50 m large (externally) and 4.20 (at the interior of the town) and was originally paved and with two benches on each lateral side, while a small staircase was put against the eastern bastion. No traces of a closing device of the doorway have been detected. North-West gate: Only construction inscriptions (MAFYS-Timna 2-5 and MAFYS-Timna 1) allow so far the reconstruction of this gate, which has not been excavated scientifically and has moreover suffered an intense pillaging. According to the texts, these gates were formed by a massive granite basement (like that still visible in the SW Gate) which sustained an upper structure made in wood and mud bricks. Limestone and calcite-alabaster blocks enriched this upper part.

South Gate

South Gate

DASI front-end

DASI DIGITAL ARCHIVE FOR THE STUDY OF PRE-ISLAMIC ARABIAN INSCRIPTIONS

DASI | Home | Projects

Corpus of South Arabian Inscriptions
Online Corpus of the Inscriptions of Ancient North Arabia

TOOLS
Map
Help
Editorial Criteria

DASI

Digital Archive for the Study of pre-Islamic Arabian Inscriptions

DASI is a five-year project directed by Alessandra Avanzini of the University of Pisa, which has been funded by the European Community within the Seventh Framework Programme "Ideas", Specific Programme "ERC – Advanced Grant".

DASI seeks to gather all known pre-Islamic Arabian epigraphic material into a comprehensive online database, with the aim to make available to specialists and to the broader public a wide array of documents often underestimated because of their difficulty of access.

By means of a digitization process through a hybrid data entry/xml system developed by the Scuola Normale Superiore di Pisa according to international encoding standards, DASI gives access at present to nearly 7,500 Ancient South Arabian inscriptions recorded by the University of Pisa team under the direction of Alessandra Avanzini.
Thanks to the fruitful collaboration with other two major European centres for the study of the epigraphic documentation of the Arabian Peninsula, also sections of the two corpora of Ancient North Arabian inscriptions (supervision by Michael C.A. Macdonald, University of Oxford) and Aramaic inscriptions (supervision by Laila Nehmé, UMR 8167, CNRS-Paris) are being prepared for consultation.

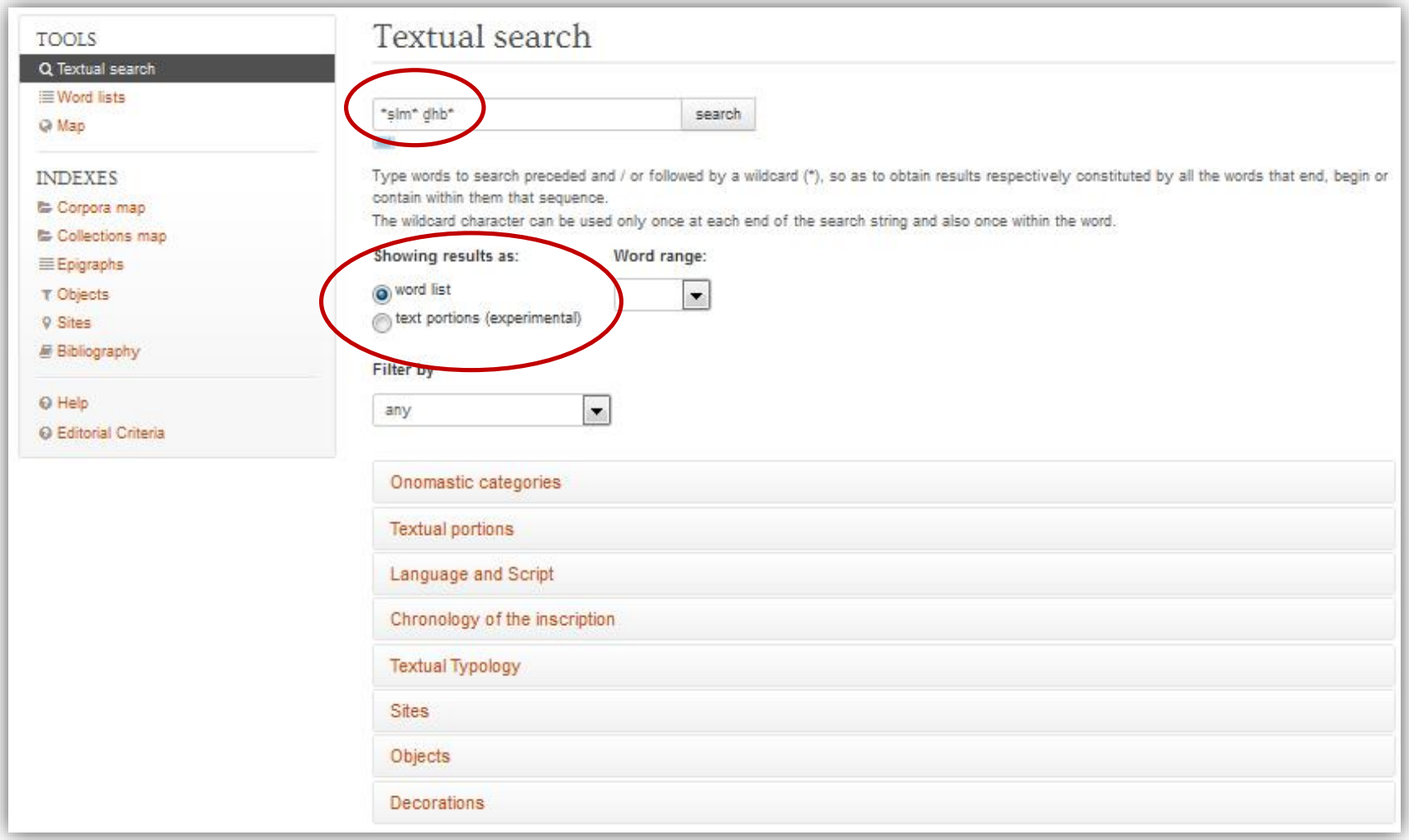
Project details »

Projects

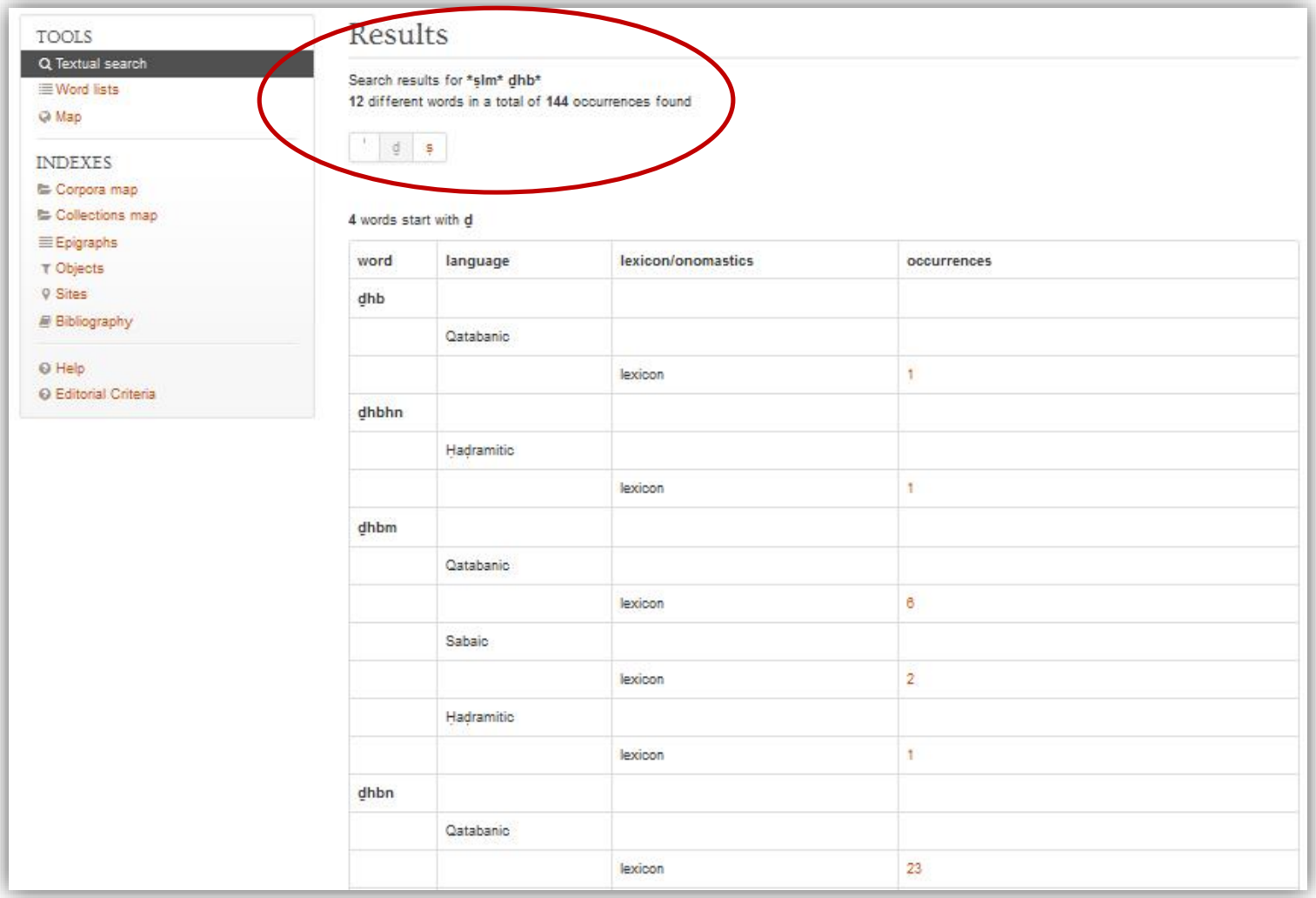
Corpus of South Arabian Inscriptions

The ASA (Ancient South Arabian) corpus is one of the most interesting collections of epigraphic documents of the Semitic world, first and foremost for its vastness. With its over 15,000 inscriptions, it is the first-hand, written documentation of the culture that flourished in South Arabia from the late second millennium BC to the sixth century AD. At present, CSAI contains a collection of some 7,500 texts digitized by the team of the University of Pisa directed by A. Avanzini.

Textual search

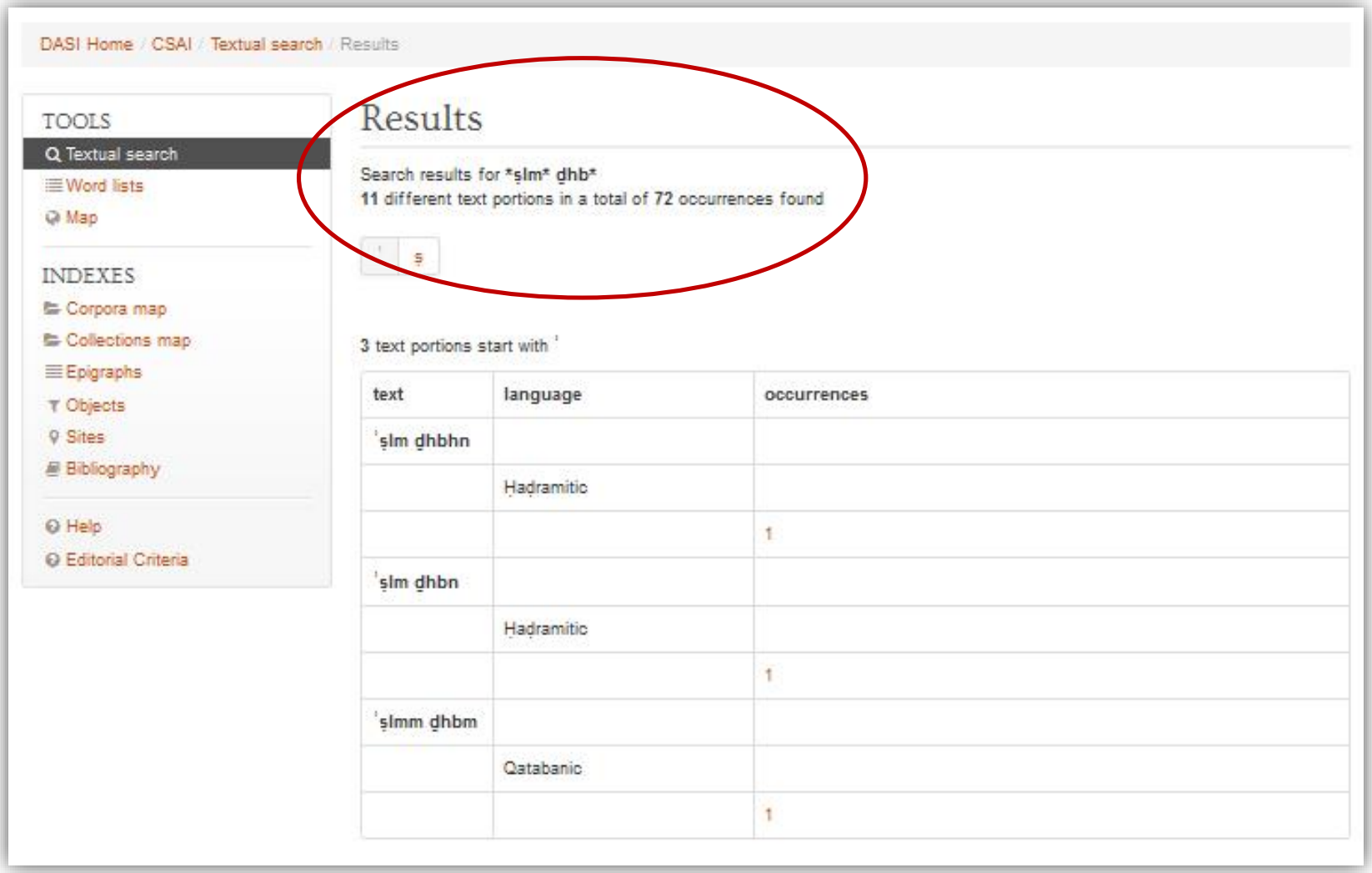


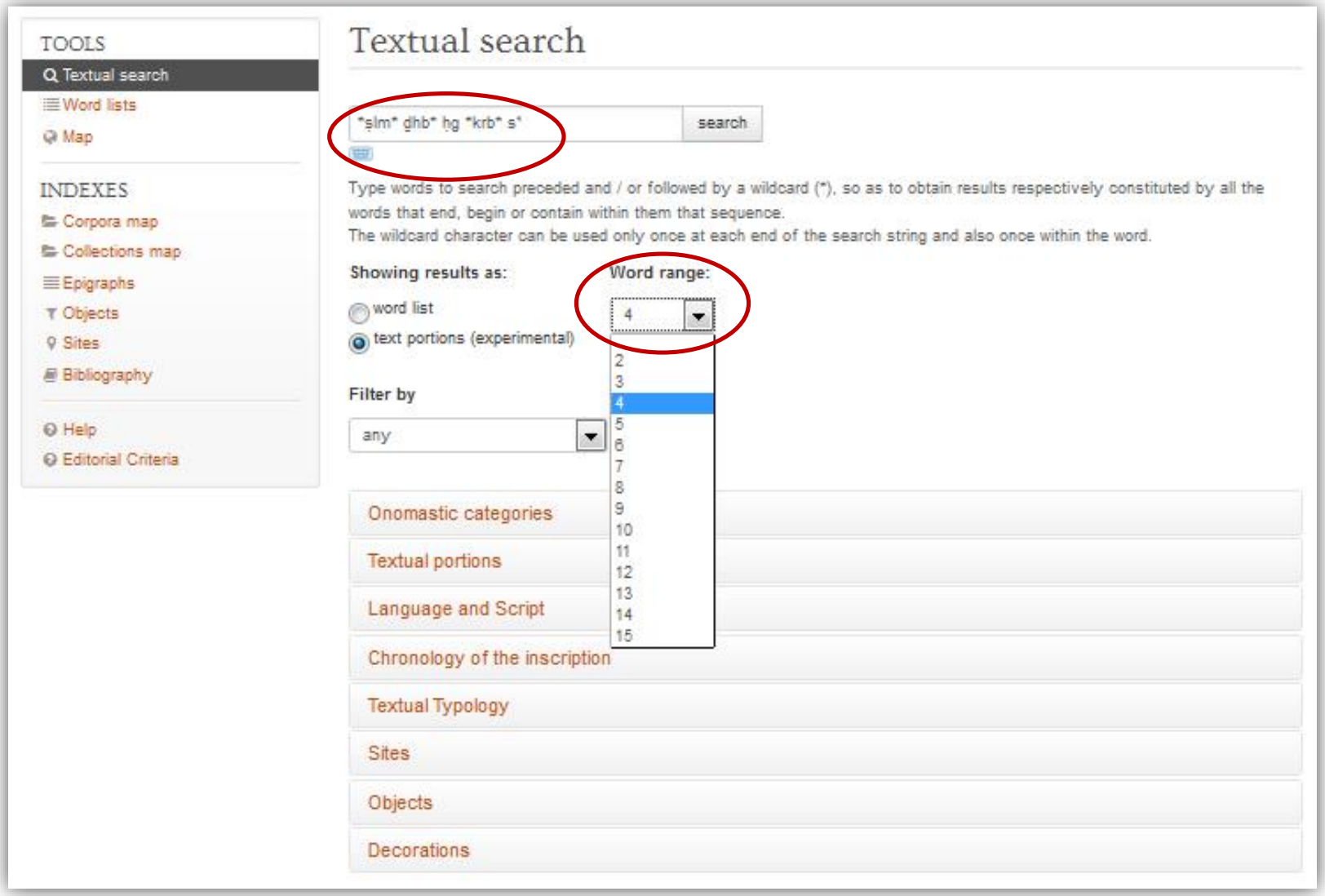
Textual search



word	language	lexicon/onomastics	occurrences
ḏhb			
	Qatabanic		
		lexicon	1
ḏhbhn			
	Ḥaḍramitic		
		lexicon	1
ḏhbm			
	Qatabanic		
		lexicon	6
	Sabaic		
		lexicon	2
	Ḥaḍramitic		
		lexicon	1
ḏhbn			
	Qatabanic		
		lexicon	23

Textual search



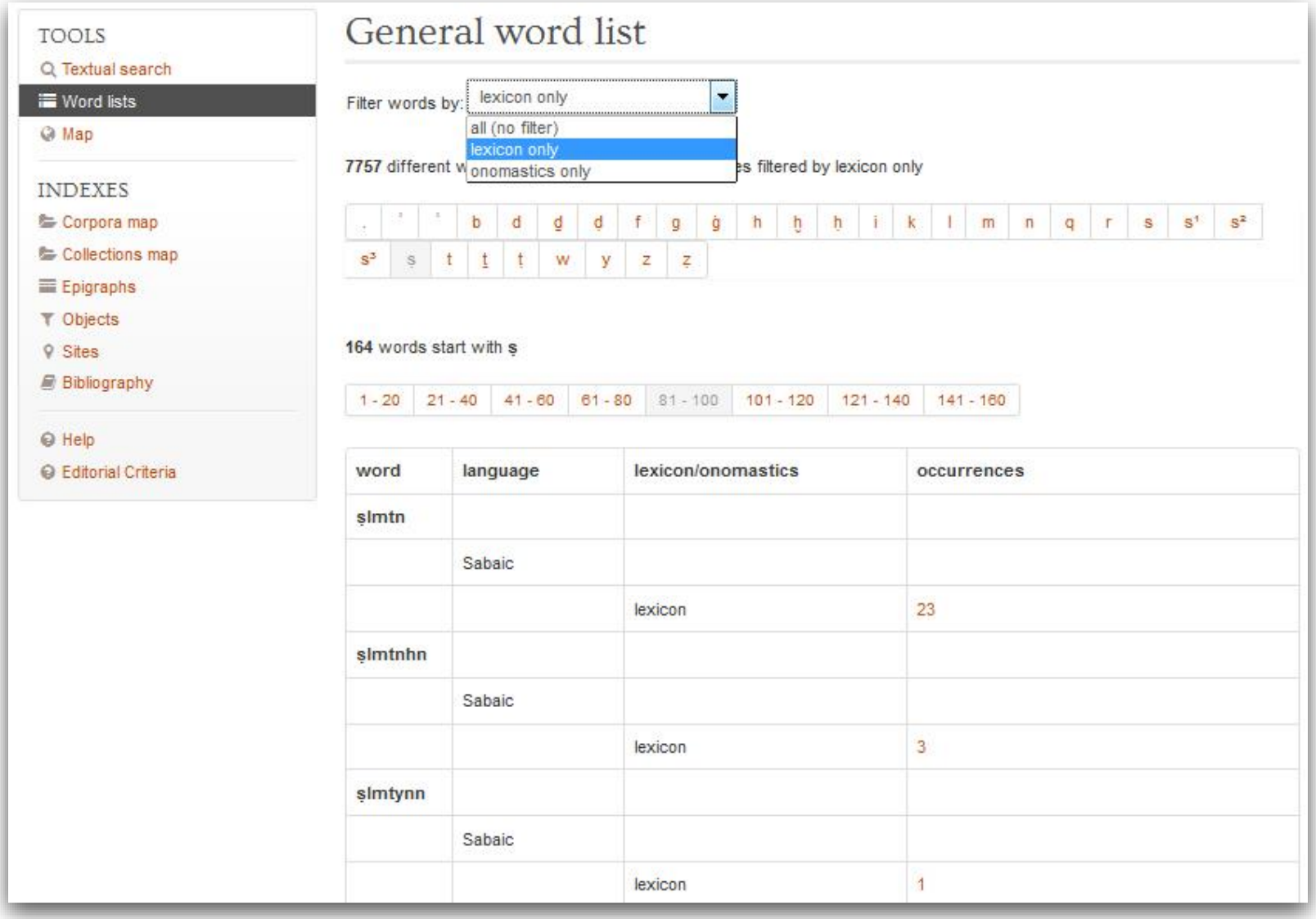

TOOLS
Textual search
Word lists
Map
INDEXES
Corpora map
Collections map
Epigraphs
Objects
Sites
Bibliography
Help
Editorial Criteria
Textual search
ṣlm ḏhb* ḥg *krb* s*
search
Type words to search preceded and / or followed by a wildcard (*), so as to obtain results respectively constituted by all the words that end, begin or contain within them that sequence.
The wildcard character can be used only once at each end of the search string and also once within the word.
Showing results as:
word list
text portions (experimental)
Word range:
4
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Filter by
any
Onomastic categories
Textual portions
Language and Script
Chronology of the inscription
Textual Typology
Sites
Objects
Decorations

Tokenization rules

- Spaces

- Tags indicating:
  - divider symbols (<g type="wordSeparator"/>: mlk|hw)
  - geometric signature (<g type="paraph">: #)
  - line breaks (<lb n="n"/>)
  - clitic components (<milestone unit="clitic"/>: mlk-hw < *mlkhw*)
  - variants (<rdg>: mlk((mlʾ))hw)

E.g.: SEARCH: “f ʾtww” > “f ʾtww”
“f|ʾtww”
“f-ʾtww”, etc



TOOLS
- Textual search
- Word lists
- Map

INDEXES
- Corpora map
- Collections map
- Epigraphs
- Objects
- Sites
- Bibliography

- Help
- Editorial Criteria

General word list

Filter words by: lexicon only

all (no filter)
lexicon only
onomastics only

7757 different w…es filtered by lexicon only

ʾ ʿ b d ḏ ḍ f g ġ h ḥ ḫ i k l m n q r s s¹ s² s³ ṣ t ṯ ṭ w y z ẓ

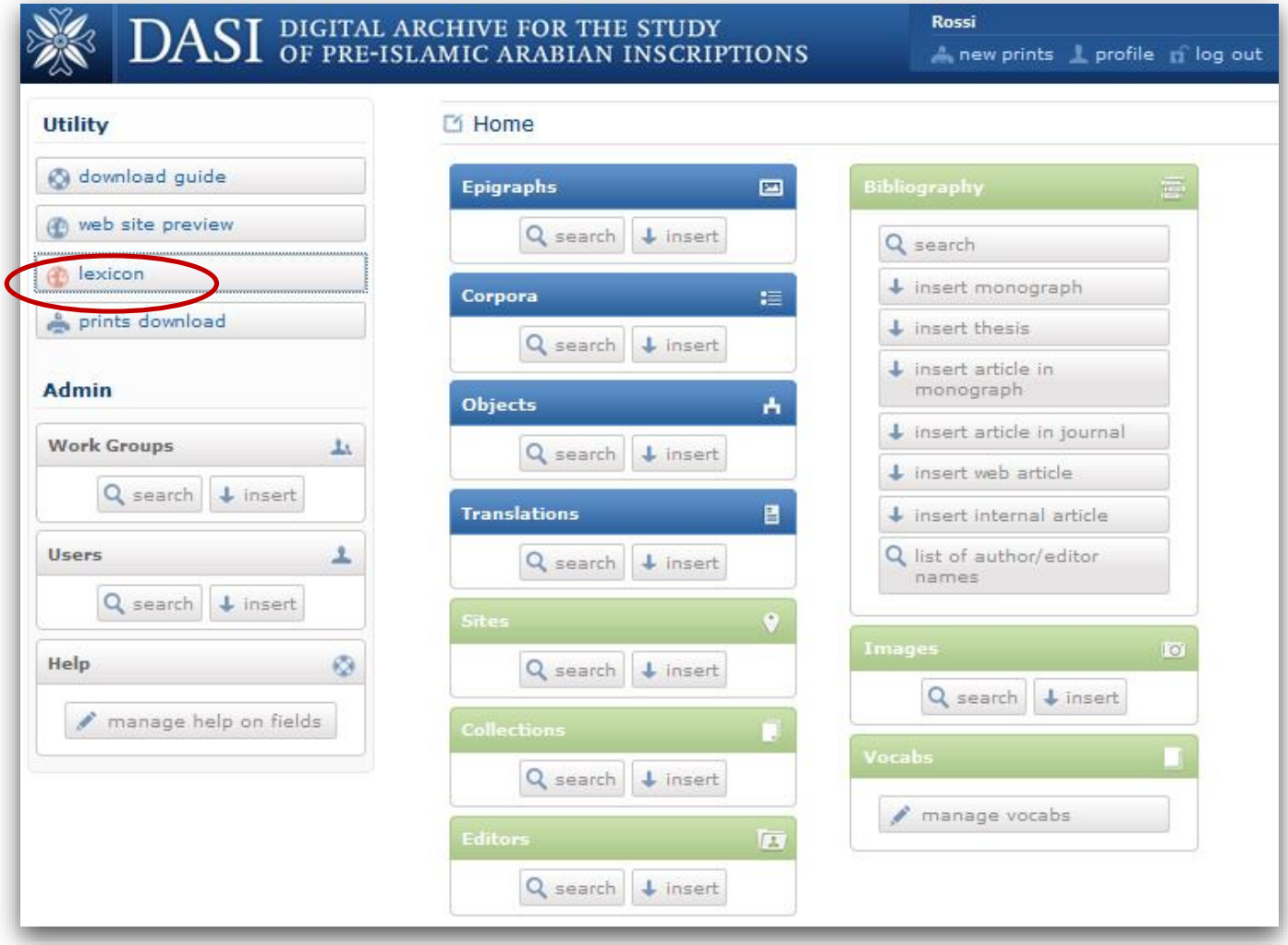
164 words start with ṣ

1 - 20 | 21 - 40 | 41 - 60 | 61 - 80 | 81 - 100 | 101 - 120 | 121 - 140 | 141 - 160

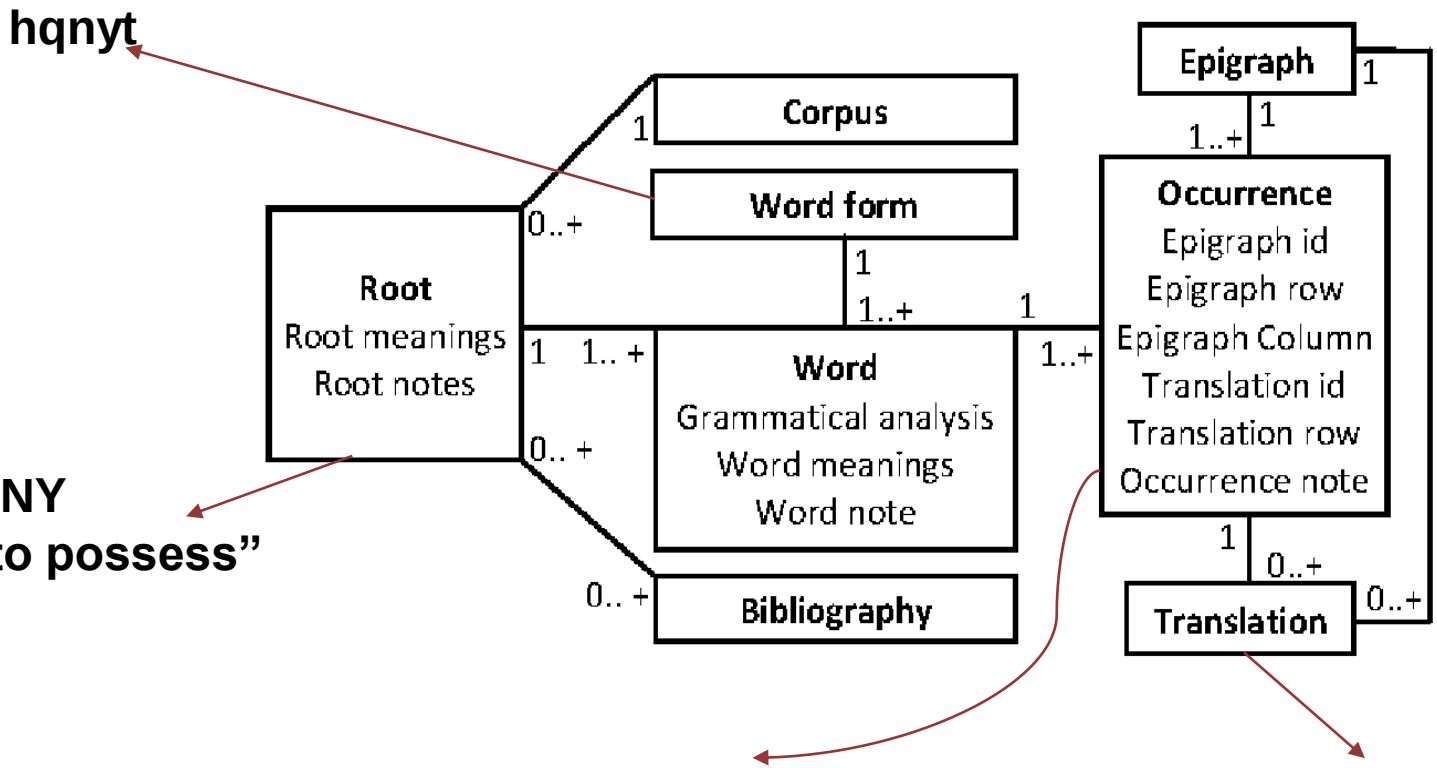
word	language	lexicon/onomastics	occurrences
ṣlmtn			
	Sabaic		
		lexicon	23
ṣlmtnhn			
	Sabaic		
		lexicon	3
ṣlmtynn			
	Sabaic		
		lexicon	1

A. De Santis I.Rossi

DASI Lexicon

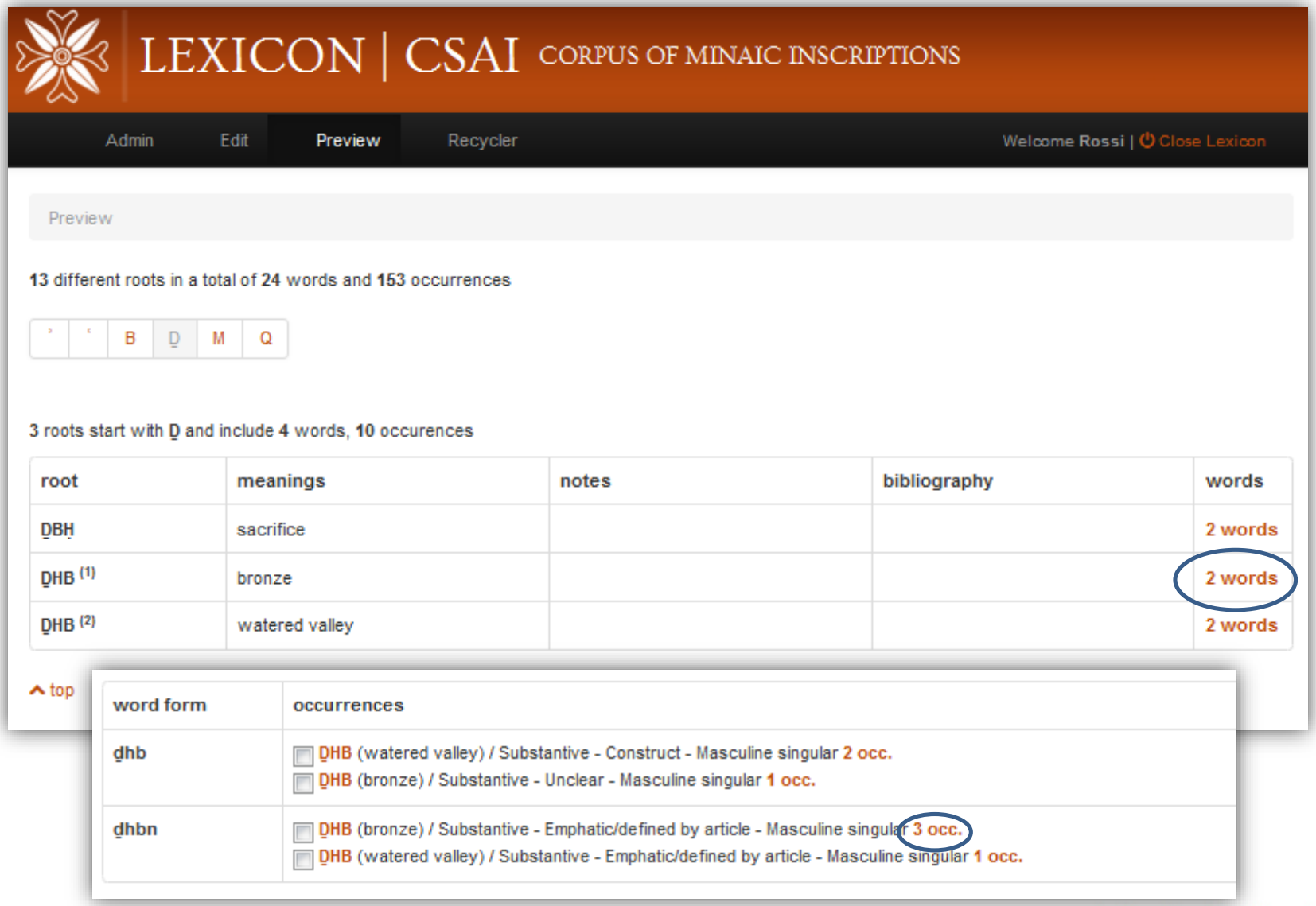


Conceptual model of DASI Lexicon



CIAS 39.11/o 3 n° 1 » w-rṯdw **hqnyt**-hmw "and committed their dedication"

CIAS 39.11/o 3 n° 10 » Ddḥ(m)d ʾmh bn Ḫḏw(t) **hqnyt** ʾlmqh "Ddḥmd servant of the family Ḫḏwt dedicated to ʾlmqh"


LEXICON | CSAI CORPUS OF MINAIC INSCRIPTIONS
Admin
Edit
Preview
Recycler
Welcome Rossi | Close Lexicon
Preview
13 different roots in a total of 24 words and 153 occurrences
ʾ
ʿ
B
Ḏ
M
Q
3 roots start with Ḏ and include 4 words, 10 occurences
root
meanings
notes
bibliography
words
ḎBḤ
sacrifice
2 words
ḎHB (1)
bronze
2 words
ḎHB (2)
watered valley
2 words
top
word form
occurrences
ḏhb
ḎHB (watered valley) / Substantive - Construct - Masculine singular 2 occ.
ḎHB (bronze) / Substantive - Unclear - Masculine singular 1 occ.
ḏhbn
ḎHB (bronze) / Substantive - Emphatic/defined by article - Masculine singular 3 occ.
ḎHB (watered valley) / Substantive - Emphatic/defined by article - Masculine singular 1 occ.

LEXICON | CSAI CORPUS OF MINAIC INSCRIPTIONS

Admin Edit Preview Recycler Welcome Rossi | Close Le

Preview / Words / Occurrences

ḏhbn

root: ḎHB | **meanings:** bronze
grammar: Substantive - Emphatic/defined by article - Masculine singular
word meanings: bronze

3 total occurrences

language	epigraph	context	translations
Central Minaic	Moussaieff 10	» ʾlʾlt Ḥnḏr((Ḥnḏr)) mʿq⁴b **ḏh(b)n** w-rṯd ʾhl ʿqrn «	• **CSAI - English (rows: 3)** [3] for ʿttr ḏ-Qbḍ, god of Ḥnḏr, and the gods of Ḥnḏr, this bronze
Central Minaic	YM 24942+YM 24943	» ʿ(w)hn ms¹lm w-mfḥmy **ḏhbn** «	• **CSAI - English (rows: 2)** [2] to ʿttr Mtbqbḍ, god of ʿwhn, and ʿwhn an altar and two incense burners in bronze.
Marginal Minaic	Riyāḍ 302F8	» kl ʾlʾlt M⁶ʿn s¹qnyt **ḏhbn** b-gntn | ⁷b-Qryt «	• **Robin in Al-Ghabban 2010 - French (rows: 6)** [6] ʿin l'offrande de bronze dans l'Oasis

A manual approach. The reasons

- Epigraphic nature of the textual sources
  - Restricted range of topics, formulaic contexts > limited set of words
  - *Hapax legomena* > data sparsity
- Morphology of the Semitic languages
  - Word formation based on the “root-and-pattern” mechanism + morphemes
  - Orthographic unit composed by more than one lexical item (wmlk < w-mlk)
  - Morphological and spelling ambiguity

  > more than one analysis for a given word form
  (*ḥqnyt* “dedication of” or “she dedicated”)
  > more than one word form for one morphological analysis and meaning
  (*ḫyl* or *ḫl* “power”; *s¹bʾt* or *s¹byʾ* “expeditions”)

- Semantic and morphological study is still in progress
  - avoid to encode grammatical and lexicographic information on the texts
  - need to modify specific occurrences or groups of them in a reasonable lapse of time
  - need to verify/elaborate rules and correct integrations and interpretations

DASI Lexicon is a preliminary step before an automatic word extractor and a morphological analyzer can be attempted

Thank you for your kind attention!

DASI

dasi.humnet.unipi.it

http://www.dasiproject.eu/