THE ONOMASTICS PROJECT

A MODERN DATA STRUCTURE FOR ANCIENT NAMES

Elizabeth Robar
Tyndale House
Cambridge, UK

MAIN TIME PERIODS

Alalakh VII, etc.	*Ugarit, Amarna, Alalakh IV, etc.*	
Middle Bronze Age (2100-1550 BC)	**Late Bronze Age (1550-1200 BC)**	**Iron Age I (1200-1000 BC)**
Abraham (ca. 2000 BC)	*Moses (mid-15th century or mid-13th century)*	*David (ca. 1000 BC)*

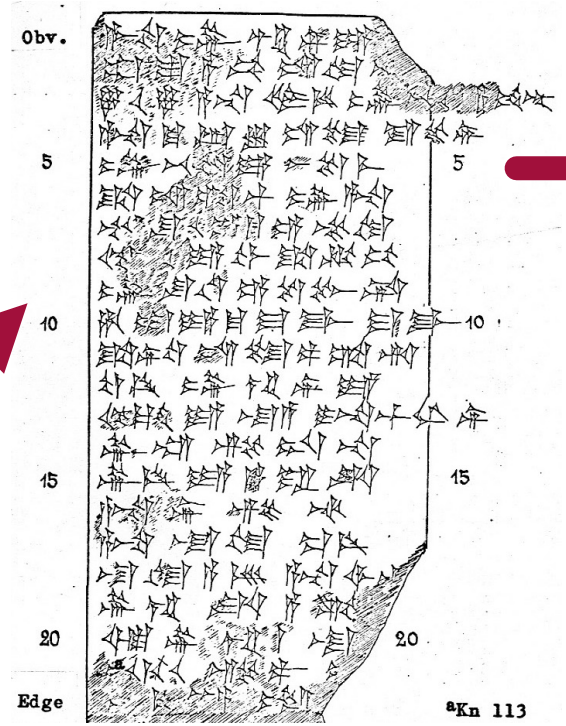
MIDDLE BRONZE AGE LOCATIONS

LATE BRONZE AGE LOCATIONS: NORTHERN ISRAEL

LATE BRONZE AGE LOCATIONS: SOUTHERN ISRAEL

LATE BRONZE AGE LOCATIONS: AMARNA

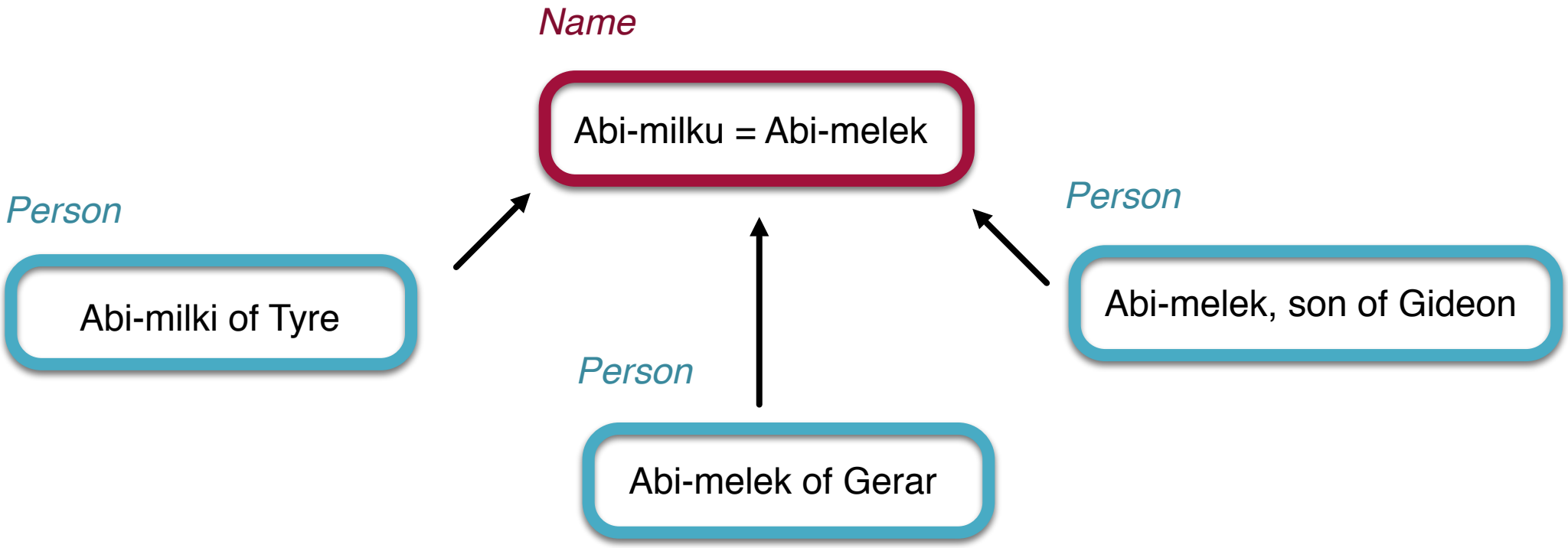
FROM ARTEFACT TO AN ATTESTED NAME



a-na LUGAL EN-*lí-ia* [. . .]
um-ma *A-bi-mil-ki* A[RAD-*ka*]
7 *u* 7 *a-na* GÌR.MEŠ LUGAL ⌜EN-*ia*⌝ *am-qut*

To the king, my lord . . . ,
Abi-milki (ʾAbī-milkī),
your servant (speaks) as follows:
“7 (times) and 7 (times) I fall at the feet of the king, my lord. ”

MULTIPLE PEOPLE WITH THE SAME NAME



ARTEFACTS, NAMES AND PEOPLE

Artefacts

Amarna 146

Amarna 147

Amarna 148

attest a name

Name

Abi-milki

referring to a person

Person

Abi-milki of Tyre

DEFINING THE RELATIONSHIP(S)

FROM PRINCIPLES TO DECISIONS

PRIMARY PRINCIPLES

- The scholars should do what only scholars can do

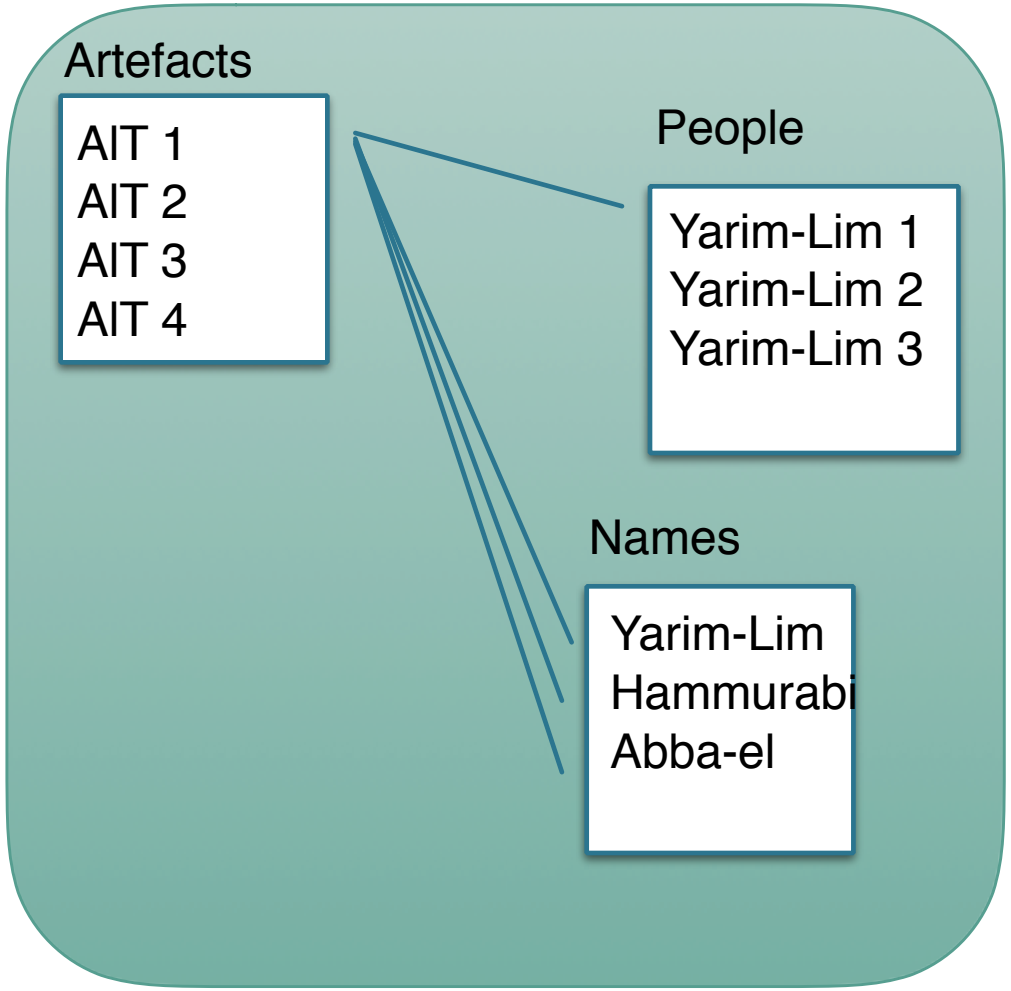
- The software is in the service of the scholars
  - Tailored to their workflow (sequence and process of entering data)
  - Requiring minimal training
  - Flexible to adapt to the evolving data model as work progresses

- Tracking all data changes; no data loss
- Separation of primary data from scholarly interpretation
- Interpretation only when necessary (if a “ya-” prefix is our reason for considering a name Amorite, then tag it as having a “ya-” prefix and a process can label it “Amorite” until we know more)

CHOOSING A DATABASE MODEL

RELATIONAL V SEMANTIC

RELATIONAL DATABASE MODEL WITH SQL




```
select
        Artefacts.Text_Description,
        Names.Full_Name,
        People.Title
from
        Artefacts
        join Attestations
            on Artefacts.Artefact_pk = Attestations.Artefact_fk
        join Names
            on Attestations.Name_fk = Names.Name_pk
        join People
            on Attestations.Person_fk = People.Person_pk
```

FROM DATA ENTRY TO SEMANTIC ANNOTATIONS

AIT 1 (“Alalakh Text 1”)

```
{{Artefact
  | CDLI=P347968
  | Manuscript=BM 131446
  | Tablet type=landscape
}}
{{Artefact/Attestation
  | At line number=o 8
  | Attests reading={diš}-ia-ri-im-li-im
  | For person=Yarim-Lim 4
}}
{{Artefact/Attestation
  | At line number=o 11.2
  | Attests reading=ia-ri-im-li-im
}}
```

Attestation

In text::AIT 1
At line number::o 8
Attests reading:: {diš}-ia-ri-im-li-im
For name::Yarim-Lim
Has determinative:: {diš}

Name

Is name:: Yarim-Lim
Is attested in text::AIT 1

Attestation

In text::AIT 1
At line number::o 11.2
Attests reading:: ia-ri-im-li-im
For name::Yarim-Lim

SEMANTIC QUERIES LINKING MULTIPLE OBJECTS

(searching for all attestations for the name Yarim-Lim)

Category:Attestation
For name::Yarim-Lim

(for all attestations in a text that has the reading “ia-ri”)

Category:Attestation
In text.Has attestation.Attests reading::~*ia-ri*

(for all names both in a text attesting the person Yarim-Lim 3 and in a text attesting the person Hammurabi 2)

Category:Name
Is attested in text.Has attestation.For person::Yarim-Lim 3
Is attested in text.Has attestation.For person::Hammurabi 2

RELATIONAL V SEMANTIC DATABASES: DATA ENTRY

- Relational databases: tables & columns and cross-reference tables
- Primary key / ID fields that maintain relationships and are in themselves semantically empty meaningless
- Permit links before the full semantic content of the links is understood
- Rigid structure; strongly typed
- Painful and time-consuming to alter

- Semantic databases: self-contained objects with properties that may link to other objects
- There is no such thing as a meaningless property: all links are by definition semantic
- No link can be established until there is semantic content
- Nearly unlimited flexibility in structure
- Simple to modify structure
- Weakly typed (risk of duplicating property names, etc. without standardization)

RELATIONAL V SEMANTIC DATABASES: DATA ANALYSIS

- Relational databases: SQL (structured query language)
- Very powerful
- Requires a thorough understanding of the relationships
- Rarely can scholars penetrate it
- Many relational databases are under-utilised because they are so inaccessible

- Semantic databases: generated properties pre-digest relationships
- Semantically named properties (with “dot” syntax) make for a pseudo-natural language
- Scholars can understand:

  artefact.attests name(“Yarim-Lim”).from period(“Middle Bronze”)

RELATIONAL V SEMANTIC DATABASES: OUR CONCLUSION

- Many databases go unused because scholars are either frightened away by the interface or simply cannot penetrate it
- We could not anticipate ahead of time the full extent of our data structure; we knew it would emerge over time
- A semantic database would permit us to enter data initially and, over time, generate an ever richer layer of semantic annotations which would constitute our research base
- Because the semantic annotations were programmatically generated, they could be refreshed every time we came up with a better interpretation
- The original data thus was kept separate from the interpretation, though both equally searchable

WHICH SEMANTIC DATABASE?

- We chose Semantic MediaWiki
  - Fundamentally a wiki = full version history, no data loss, complete tracking
  - Familiar (same platform as Wikipedia)
  - Quick and easy interface for non-programmers to use
    - Standard styling makes for an acceptable interface, even for a programmer to develop!
  - Free, open source

- We did not assume we would stay with MediaWiki for the duration of the project; we did not know if it would scale with the project

- Our end goal is a serviced database openly available; MediaWiki remains a viable option

AN EVER-EVOLVING DATA STRUCTURE

“ACTUALLY, THAT WON’T WORK”

FIRST DATA STRUCTURE:

AKKADIAN TABLETS & TEXTS

- 3 conceptual objects: Text / Name / Person
- Composition (descriptive title, language, dialect)
  - Manuscript (scribe)
    - Artefact (provenance, museum number)
- Name (e.g. Nebuchadnezzar; normalisation, linguistic analysis)
  - Orthography (e.g. {d}MUATI-ku-dúr-PAB; attestation list, ascribed language, ethnicity)
    - Linguistic element (e.g. Nabû; loanword, logogram, roots)
- Person (title, profession, chronology, geography)

- 2 physical realities: Attesting Artefact / Attested orthography

- Artefact (museum number, CDLI number, publication information)
  - Attestation (of a given name)
    - Orthography (= attestation stripped of diacritics)

- Orthography
  - Normalisation (defines a name)
    - Linguistic analysis
    - Person

- A normalisation is a property of the orthography

SECOND DATA STRUCTURE:

ATTESTATION BEFORE ORTHOGRAPHY

THIRD DATA STRUCTURE:

NAME AS OF YET UNKNOWN

- Work flow: the text of an Artefact may have multiple attestations referring to a name. However, the name itself may not be known yet or be known only provisionally.

- If multiple orthographies belong to the same normalisation, they should be grouped together, but how to do this before the normalisation itself has yet been determined?

THIRD DATA STRUCTURE:

SOFT LINK: COLLOQUIAL NAME

- Initially, we conceived of the orthography to contain the normalisation, which defined a name
- But the name cannot be determined with certainty until much of the data has been entered
- Even if the name can be determined, the normalisation generally remains provisional until after multiple passes of analysis

- A semantic database does not permit a link from an attestation to a normalisation, based on a name, if the name is not yet known, because the name only exists once it is known

- Practical solution: the colloquial name, which becomes a soft link to connect attestations, names and people:
  - colloquial “Niqmepa” instead of normalised Niqmîpa or Niqmī-yipuʕ

- Colloquial names are freely provisionally assigned as provisional connections that became the sandbox in which to explore and determine proper scholarly connections

- Artefact: a physical object with a text
- Attestation: a written instance of a personal name
  - uniquely identified by line number and ordinal (e.g. o 1.1, o 1.2)
- Each attestation is attributed to a colloquial name

- Name: uniquely identified by a colloquial name (perhaps multiple; splits and mergers to be accommodated) with multiple possible normalisations

- Linguistic Analysis: one coherent means of analysing a particular name as interpreted within a given language and dialect
  - composed of inflected words
    - composed of morphemes

- Person: a flesh-and-blood person (mostly) who held a name (or multiple names) at a point in history

THIRD DATA STRUCTURE:

AS THE SCHOLAR LEARNS

- Attestation: a written reference to a person
  - whether by name (e.g. “Yarim-Lim”)
  - or by designation (e.g. “my brothers”)
  - which may refer to one or multiple people

- Additional designation: a means of referring to a person without the use of a personal name; this is the main source of prosopographical data (e.g. ‘LUGAL’)

- Names v designations: ‘Johnson’ vs ‘Son of John’

THIRD DATA STRUCTURE:

PEOPLE WITHOUT NAMES

AN EXAMPLE, ARTEFACT TO ANALYSIS

ALT 1 (PLAIN TEXT OF WEB PAGE)

```
{{Artefact
|CDLI=P347968
|Manuscript=BM 131446
|Tablet type=landscape
}}
{{Artefact/Attestation
|At line number=o 1.1
|For person=Brothers of Abba-el 1
|Attests designation=aḫ-ḫu-šu
}}
{{Artefact/Attestation
|At line number=o 1.2
|Attests name={diš}-ab-ba-DINGIR
|For person=Abba-el 1
|Attests designation=be-el-šu-nu
|More notes=The 3mp pronominal suffix on bēlšunu refers to Abba-el's brother,
mentioned earlier in the line (aḫ-ḫu-šu).
}}
```

This can be HTML interspersed with wikitext, including templates (indicated by the double curly braces, a template name, then parameters separated by pipes).

This is NOT searchable any more than a flat web page.

SEMANTIC OBJECTS & ANNOTATIONS GENERATED

Template:Artefact
(code to generate semantic annotations)

Template:Artefact/Attestation
(code to generate semantic annotations)

Category:Attestation
In text::AlT 1
At line number::o 1.1
Designated::aḫ-ḫu-šu
For person::Brothers of Abba-el 1
Links to::o 1.2

Category:Artefact
Is text::AlT 1
Has CDLI number::P347968
Has table type::Landscape
Has attestation count::20

Category:Attestation
In text::AlT 1
At line number::o 1.2
Attests name:: {diš}-ab-ba-DINGIR
For name::Abba-el
Designated::be-el-šu-nu
Has determinative:: {diš}

Run query: Basic Search

Enter the information for which you want to search. Use an asterisk (*) as a wildcard, if you are searching for discontinuous letters. Note that it is **case-sensitive**. Maximum number of results to display: 50000

1. The first checkbox is for **case-sensitive searching** (otherwise, case does not matter).
2. The second checkbox is for **exact matches** (otherwise, additional text before or after is considered acceptable.)

Artefact Catalog Data [edit]

CDLI

Primary Publication

Author

Publication Date

Secondary Publications

Collection

General Provenience

Specific Provenience

Google Maps Location

Museum No.

Accession No.

Excavation No.

Archaeol. Period

Assyr. Period

Dates Referenced

Absolute Dates

Object Type

Height

Width

Thickness

Tablet Type

Remarks

Material

Language

Genre

Subgenre

Archive

Seal Name

Seal Bearer

Description

Text Date

Manuscript

Composition

Keywords

Description

Attestations [edit]

Artefact

Normalization

Designated

Search for: ○ **Artefacts** ○ **Attestations** ○ **Names** ● **Name Analyses** ○ **Words** ○ **Morphemes**

Search Conditions

```
With element.Glossed as::God
```

Run query

Name Analyses: Names with analyses [edit]

-Part of name analysis.Glossed as::God

Returns 12 Name Analyses

For name	Translated as	In language	Normalized as	Part of speech pattern	Syntactic pattern	Word count	Morpheme count
Abba-el	Abba is the god.	Akkadian	Abba-ilu	Noun-Noun	Predicate-Subject	2	3
Ili-Kubaba	Kubaba is my god.	Akkadian	Ilī-Kubaba			0	0
Abba-el	Abba is the god.	Amorite	ʔAbba-ʔel	Noun-Noun	Subject-Predicate	2	3
Abdi-ili	Servant of the god.	Amorite	ʕAbdu-ʔili			0	0
Ili-Dagan	Dagan is my god.	Amorite	ʔIlī-Dagan			0	0
Ili-Kubaba	Kubaba is my god.	Amorite	ʾilī-Kubaba			0	0
Irḫam-ila	The god has shown mercy.	Amorite	ʾIrḥam-ʾila			0	0
Uštanni-ila	The hero is a god.	Hurrian and Semitic	Ušt=a=nni-il=a			0	0
Eḫlum-eni	The god(ess) has saved (the child).	Hurrian	Eḫl=o=m-en(i)=ni			0	0
Bendi-ili	Make (the child) good, O god!	Hurrian-Semitic	Fend=i-ili=Ø			0	0
Bendi-ili	The god makes (the child) good.	Hurrian-Semitic	Fend=Ø=i-ili=Ø			0	0
Talmali	The god is great.	Hurrian-Semitic	Talm=a-ili			0	0

UPON REFLECTION

- We started with the conceptual data, as if we had come to all our scholarly conclusions
- We had to re-structure to accommodate slowly evolving scholarly conclusions (the endpoint, rather than starting point, of the scholarly process!)
- A continually evolving structure was made possible by a semantic database model

- Chosen software (MediaWiki)
  - Solves data tracking (keeps unlimited version history)
  - Semantic annotations provide semantic database capability
  - Interface makes for quick and easy modifications and user comfort
  - Interpreted “wikitext” constrains heavily data manipulation
  - Now that we want to explore and analyse… Python and data visualisation?!? R?!? Machine learning?