

Hiob-Ludolf-Zentrum
für Äthiopistik



A Short Introduction to XML: Mark Up

Eugenia Sokolinski

What is it all about

Bausi, A. 2021. “Paleografia quale scienza dello spirito”: Once more on the Gə‘əz inscription of Ham (*RIÉ* no. 232)’, in J. B. Quenzer, ed., *Exploring Written Artefacts: Objects, Methods, and Concepts*, I, *Studies in Manuscript Cultures*, 25 (Berlin–Boston, MA: De Gruyter, 2021), 3–33.

Gori, A. 1993. ‘Islam in Etiopia’, *Rassegna di Studi Etiopici*, 37 (1993), 45–87 (see <<http://www.jstor.org/stable/41299787>>) (DOI: 10.2307/41299787).

Article	Author		Publication	Issue/ Volume	Year	Page Numbers	URL
Title	Last name	First name					
Islam in Etiopia	Gori	Alessandro	Rassegna di studi etiopici	37	1993	45-87	http://....
“Paleografia quale scienza dello spirito”: Once more on the Gə‘əz inscription of Ham (<i>RIÉ</i> no. 232)	Bausi	Alessandro	J. B. Quenzer, ed., <i>Exploring Written Artefacts: Objects, Methods, and Concepts</i>	1	2021	3-33	

Part 1

Things about things (or Words about words)

Metadata

Metalanguage

Markup

Markup language

Tag

Explaining things: metadata

Metadata: data (information) about things (or other data)

- e.g. A book in a library has a description in the catalogue = metadata
- e.g. People and companies have short descriptions (address, phone number) in a telephone book = metadata
- e.g. Persons, places, objects, events are described in an encyclopaedia = metadata
- e.g. Words are described in a dictionary = metadata
- e.g. A manuscript in a repository, or a digital copy of a manuscript in a computer has a description in a catalogue (printed or digital) = metadata
- e.g. A text has been copied many times, with changes, and a collation/critical edition have been produced to describe it = metadata
- can be descriptive (title/name, content/profession), administrative (access, rights, use), structural (composition)

Explaining things: metalanguage

- Metalanguage:** a language (words or symbols = terms) to describe another language
- e.g. grammatical categories sets (noun, verb, case, etc.) used to describe a human language
 - e.g. programming language, that is a set of terms to get a computer produce the results you want

Explaining things: markup

Markup: is needed to structure text to make it understandable (to humans and to computers, but computers of course need more formal logic)

bf Feature - Researching to Add Richness

Sometimes you're writing nothing but a *Blistering* Hot love scene, or a character study, or a drabble. In those cases, you're canon is all you'll want or need, but *there are other* times Yuletide, a Big Bang, your own original novel-in-progress, that sprawling Au epic that's been taking up all your spare imagination for weeks when you're going to want the sort of *little* details that add texture and depth to a story! You may be blessed with a library that includes just the reference books you want or a local library with sympathetic librarians. If not, there's always the Intarwebs ... but sifting treasure from trash can be a challenge.

Let's take a look at some ways to make the world wide web cough up the details that you need to write the story that you want.

Stet *ital* With examples involving the historical novel Frontier Wolf.

Explaining things: markup language

Markup language: a set of rules governing what markup information may be included in a document and how it is combined with the content of the document

bf Feature - Researching to Add Richness
~~~~~  
 Sometimes you're writing nothing but a Blistering Hot love scene, or a character study, or a drabble. In those cases, you're canon is all you'll want or need, but there are other times Yuletide, a Big Bang, your own original novel-in-progress, that sprawling epic that's been taking up all your spare imagination for weeks when you're going to want the sort of little details that add texture and depth to a story! You may be blessed with a library that includes just the reference books you want or a local library with sympathetic librarians. If not, there's always the Intarwebs ... but sifting treasure from trash can be a challenge.  
 Let's take a look at some ways to make the world wide web cough up the details that you need to write the story that you want.  
 With examples involving the historical novel Frontier Wolf.

| Marking                                                                                | Location          | Meaning                                                                                                                                                        |
|----------------------------------------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <span style="border: 1px solid red; border-radius: 50%; padding: 2px;">align</span>    | margin            | Align text vertically or horizontally, depending on the guidelines drawn in the text.                                                                          |
| <span style="border: 1px solid red; border-radius: 50%; padding: 2px;">bf</span>       | margin            | Use boldface type for the marked text.                                                                                                                         |
| Add <span style="color: red;">~~~~~</span>                                             | inline            | Use boldface type on the text underlined with the zigzags.                                                                                                     |
| Au <span style="color: red;">≡</span>                                                  | inline            | Capitalize the marked text.                                                                                                                                    |
| <span style="border: 1px solid red; border-radius: 50%; padding: 2px;">@wide@eb</span> | inline            | Capitalize the circled text (alternate version).                                                                                                               |
| <span style="border: 1px solid red; border-radius: 50%; padding: 2px;">Cap</span>      | margin            | Capitalize the indicated text.                                                                                                                                 |
| ∑                                                                                      | margin and inline | Insert a comma.                                                                                                                                                |
| <span style="color: red;">e</span>                                                     | margin and inline | Delete the marked text. When used inline, the longer "tail" of the symbol crosses through the text to be deleted.                                              |
| <span style="color: red;">E</span>                                                     | margin and inline | Delete the marked text and close up the characters on either side to fill the space left.                                                                      |
| <span style="border: 1px solid red; border-radius: 50%; padding: 2px;">em dash</span>  | margin            | Insert an em dash (the insertion of an en dash is handled similarly)                                                                                           |
| <span style="color: red;">-</span>                                                     | inline            | Change the hyphen to an em dash (the change of a hyphen to an en dash is handled similarly).                                                                   |
| <span style="color: red;">M</span>                                                     | inline            | Insert an em dash. Note the caret marking the insertion point, a notation that can be used for other insertions as well, such as a single character or a word. |
| <span style="border: 1px solid red; border-radius: 50%; padding: 2px;">ital</span>     | margin            | Use italic type for the marked text.                                                                                                                           |
| <u>novel</u>                                                                           | inline            | Italicize the text that's underlined.                                                                                                                          |

# Explaining things: markup language

**Markup language:** a set of rules governing what markup information may be included in a document and how it is combined with the content of the document

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>bf</b> Feature - Researching to Add Richness</p> <p>Sometimes you're writing nothing but a <b>Blistering</b> Hot love scene, or a character study, or a drabble. In those cases, you're canon is all you'll want or need, but there are other times Yuletide, a Big Bang, your own original novel-in-progress, that sprawling Au epic that's been taking up all your spare imagination for weeks when you're going to want the sort of <b>little</b> details that add texture and depth to a story. You may be blessed with a home library that includes just the reference books you want or a local library with sympathetic librarians. If not, there's always the Intarwebs ... but sifting treasure from trash can be a challenge.</p> <p>Let's take a look at some ways to make the world wide web cough up the details that you need to write the story that you want.</p> <p>With examples involving the historical novel Frontier Wolf.</p> | <p><b>Feature - Researching to Add Richness</b></p> <p>Sometimes you're writing nothing but a blistering hot love scene, or a character study, or a drabble. In those cases, your canon is all you'll want or need. But there are other times Yuletide, a Big Bang, your own original novel-in-progress, that sprawling AU epic that's been taking up all your spare imagination for weeks — when you're going to want the sort of details that add texture and depth to a story. You may be blessed with a home library that includes just the reference books you want or a local library with sympathetic librarians. If not, there's always the Intarwebs ... but sifting treasure from trash can be a challenge.</p> <p>Let's take a look at some ways to make the World Wide Web cough up the details that you need to write the story that you want.</p> <p>With examples involving the historical novel Frontier Wolf.</p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|



**Manuscripts**

The term M. (from Lat. *manu scriptum*, 'written by hand') is applied to artefacts (codices, unbound fascicles, scrolls, or even single leaves), usually on parchment, much more rarely on paper (writing on papyrus is not attested in the Ethio-Eritrean region), which primarily carry texts and/or paintings; their contents are themselves of no relevance to the definition of M. as such. Epigraphic writings (inscriptions on stone, metal, wood), occasional writings on paintings (captions and legends), on pottery, textiles, and other objects (such as *mānbārā tabot*), are not to be considered M. On the other hand, scrolls (whether magic or not), although sometimes considered separately, can be labelled M.

Christian, Islamic and Betä ʿĪsraʿel cultures have developed and been transmitted in Ethiopia and Eritrea mainly through the production and preservation, in M. form, of numerous texts in various fields (theological, liturgical, literary, historical, etc.) and in various languages (esp. ʿGəʿəz and ʿArabic, ʿAmharic, and, to a lesser degree, ʿHarari, ʿTəgrəñña and others).

The peculiar and exceptionally relevant importance of M. in Ethiopian culture is determined by the very late introduction of an extensive use of printing, which does not antedate the 19<sup>th</sup> cent. (Incunabula). Ethiopia and Eritrea are still among the few countries where a tradition of parchment M. (Manuscript preparation) is still alive. The M. tradition is rooted in the peculiar past of the area, where – contrary to most African countries south of the Sahara – written tradition (starting from the 1<sup>st</sup> millennium B.C. in the so-called “South-Arabian phase” of Ancient Ethiopian history) has played a central role in the development of civilization.

**Manuscript studies**

The study of M. is of major importance to research in various fields of Ethiopian studies. Modern research, esp. recent trends starting from the second half of the 20<sup>th</sup> cent., has clearly distinguished various branches within M. studies, or “manuscriptology”, each one with its own methodology and scope, based upon the fundamental distinction between text and what supports the text.

Codicology studies hand-written books as artefacts, with specific reference to their material constituents and techniques of production (“archaeology of M.”, cp., e.g., Muzerelle 1985; Ma-

niaci 2002). Illuminations and drawings in M. are the subject of art history.

Palaeography considers the peculiarities of writing and its general development and dating (historically the first aim of palaeography), but also the general “act of writing” in all its general contexts and aspects (UhPal).

Text-criticism, or philology in the narrow sense (cp. Bausi in PICES 15b), has refined its methods in the search for reliable criteria to establish the original reading of texts. It takes advantage of recent developments in other disciplines, witnessed, for example, by the emergence of such specific sub-disciplines as “codicological stemmatics”, or “material philology”. These new research directions establish the relationships between the M. and the texts they transmit, also by tracing the history of each single M., its material structure and/or the influence specific copies may have exerted within a single textual tradition. So far these recent disciplinary subdivisions have seldom been applied to Ethiopian M. studies.

**Ethiopic manuscripts**

I. General topics in Ethiopic M. studies  
 There are fundamental issues in M. studies which still need a definitive assessment, such as:

(1) The total number of manuscripts, inside and outside Ethiopia and Eritrea

(a) Manuscripts inside Ethiopia and Eritrea  
 The exact number of Ethiopic M. in the Ethio-Eritrean region is not known. The estimate that there should be ca. 200,000 (excluding scrolls; cp. Sergew Hable Selassie 1981:35) seems realistic if one considers the minimum number of M. necessary for every church for religious and liturgical practice (the number of parishes ranging from at least 13,000 to 32,350), and the occasionally large M. collections still preserved in monastic libraries (e.g., ʿDābrā Hayq ʿĪstifanos, ca. 200 M.; ʿDābrā Bizān, ca. 570 M.; ʿGundā Gunde, formerly 800 [?] and now ca. 220 M.; ʿAksum Šayon, several hundred M. at least). Although still relatively little exploited by scholars and not yet furnished with printed catalogues, two institutional Ethiopian libraries have rich M. collections: the Library of the ʿInstitute of Ethiopian Studies, ca. 1,500 M.; and the ʿNational Archives and Library of Ethiopia (NALE), ca. 850 M.

(b) Manuscripts outside of Ethiopia and Eritrea  
 The number of M. abroad may be estimated to be several thousands, most of them catalogued

# Explaining things: markup

**Markup:** is needed to structure text to computers, but computers of course ne

**bf** Feature - Researching to Add Richness

Sometimes you're writing nothing but a Blistering Hot love scene, or a character study, or a drabble. In those cases, your canon is all you'll want or need, but there are other times Yuletide, a Big Bang, your own original novel-in-progress, that sprawling epic that's been taking up all your spare imagination for weeks when you're going to want the sort of little details that add texture and depth to a story! You may be blessed with a library that includes just the reference books you want or a local library with sympathetic librarians! If not, there's always the Intarwebs ... but sifting treasure from trash can be a challenge.

Let's take a look at some ways to make the world wide web cough up the details that you need to write the story that you want.

With examples involving the historical novel Frontier Wolf.

# Explaining things: markup

**Markup:** is needed to structure text to computers, but computers of course need

Note: the markup example on the right (from the Encyclopaedia Aethiopica) uses the markup convention of the Encyclopaedia Aethiopica. Other books may use e.g. italics for the main title, quotation marks for foreign words, indentation for subtitle, ↑↑ or bold type for cross-references, etc., etc., etc.

When we speak of books, the markup conventions are usually specific for each publication project (single book, series), publishing house and must be consistent so that the human eye can easily understand what is what = parse the description.

Manuscripts

Manuscripts

**Bold = title**

*Italics = foreign*

The term M. (from Lat. *manu scriptum*, 'written by hand') is applied to artefacts (codices, unbound fascicles, scrolls, or even single leaves), usually on parchment, much more rarely on paper (writing on papyrus is not attested in the Ethiopian region), which primarily carry texts and/or paintings; their contents are themselves of no relevance to the definition of M. as such. Epigraphic writings (written on metal, wood), occasional writings on paintings (captions and legends), on pottery, textiles, and other objects (such as *manbärä tabot*), are not to be considered M. On the other hand, scrolls (whether magic or not), although sometimes considered separately, can be labelled M.

**Arrow = see also**

Christian, Islamic and Betä ʿĪsraʿel cultures have developed and been transmitted in Ethiopia and Eritrea mainly through the production and preservation, in M. form, of numerous texts in various fields (theological, liturgical, literary, historical, etc.) and in various languages (esp. ʾGəʿəz and ʾArabic, ʾAmharic, and, to a lesser degree, ʾHarari, ʾTəgrəñña and others).

The peculiar and exceptionally relevant importance of M. in Ethiopian culture is determined by the very late introduction of an extensive use of printing, which does not antedate the 19<sup>th</sup> cent. (ʾIncunabula). Ethiopia and Eritrea are still among the few countries where a tradition of parchment M. (ʾManuscript preparation) is still alive. The M. tradition is rooted in the peculiar past of the area, where – contrary to most African countries south of the Sahara – written tradition (starting from the 1<sup>st</sup> millennium B.C. in the so-called “South-Arabian phase” of Ancient Ethiopian history) has played a central role in the development of civilization.

Manuscript studies

**Line break = subtitle**

The study of M. is of major importance to research in various fields of Ethiopian studies. Modern research, esp. recent trends starting from the second half of the 20<sup>th</sup> cent., has clearly distinguished various branches within M. studies, or “manuscriptology”, each one with its own methodology and scope, based upon the fundamental distinction between text and what supports the text.

Codicology studies hand-written books as artefacts, with specific reference to their material constituents and techniques of production (“archaeology of M.”, cp., e.g., Muzerelle 1985; Ma-

niaci 2002). Illuminations and drawings in M. are the subject of art history.

ʾPalaeography considers the peculiarities of writing and its general development and dating (historically the first aim of palaeography), but also the general “act of writing” in all its general contexts and aspects (UhPal).

Text-criticism, or ʾphilology in the narrow sense (cp. Bausi in PICES 15b), has refined its methods in the search for reliable criteria to establish the original reading of texts. It takes advantage of recent developments in other disciplines, witnessed, for example, by the emergence of such specific sub-disciplines as “codicological stemmatics”, or “material philology”. These new research directions establish the relationships between the M. and the texts they transmit, also by tracing the history of each single M., its material structure and/or the influence specific copies may have exerted within a single textual tradition. So far these recent disciplinary subdivisions have seldom been applied to Ethiopian M. studies.

Ethiopic manuscripts

I. General topics in Ethiopic M. studies

There are fundamental issues in M. studies which still need a definitive assessment, such as:

(1) The total number of manuscripts, inside and outside Ethiopia and Eritrea

(a) Manuscripts inside Ethiopia and Eritrea

The exact number of Ethiopic M. in the Ethiopian region is not known. The estimate that there should be ca. 200,000 (excluding scrolls; cp. Sergew Hable Selassie 1981:35) seems realistic if one considers the minimum number of M. necessary for every church for religious and liturgical practice (the number of parishes ranging from at least 13,000 to 32,350), and the occasionally large M. collections still preserved in monastic libraries (e.g., ʾDäbrä Hayq ʾĪstifanos, ca. 200 M.; ʾDäbrä Bizän, ca. 570 M.; ʾGundä Gunde, formerly 800 [?] and now ca. 220 M.; ʾAksum Šayon, several hundred M. at least). Although still relatively little exploited by scholars and not yet furnished with printed catalogues, two institutional Ethiopian libraries have rich M. collections: the Library of the ʾInstitute of Ethiopian Studies, ca. 1,500 M.; and the ʾNational Archives and Library of Ethiopia (NALE), ca. 850 M.

(b) Manuscripts outside of Ethiopia and Eritrea

The number of M. abroad may be estimated to be several thousands, most of them catalogued

**Bottom page = page number**

# Explaining things: using markup

**Presentational markup:** used for display, as in the examples above. This is the markup invisibly applied the text editors you use. (And you may be surprised to know that Word, OpenOffice and even Photoshop actually apply their markup in the same way, only that you do not see it from the graphic user interface – but it is always there.)

**Bold:** Manuscripts

**Plain:** The term M. (from Lat. **Italic:** manu scriptum) ...

**Manuscripts**

The term M. (from Lat. *manu scriptum*) ...

**Descriptive markup:** assigns labels to describe the actual content; such markup can be used both for displaying the data but also for looking for something in particular.

**Title:** Manuscripts

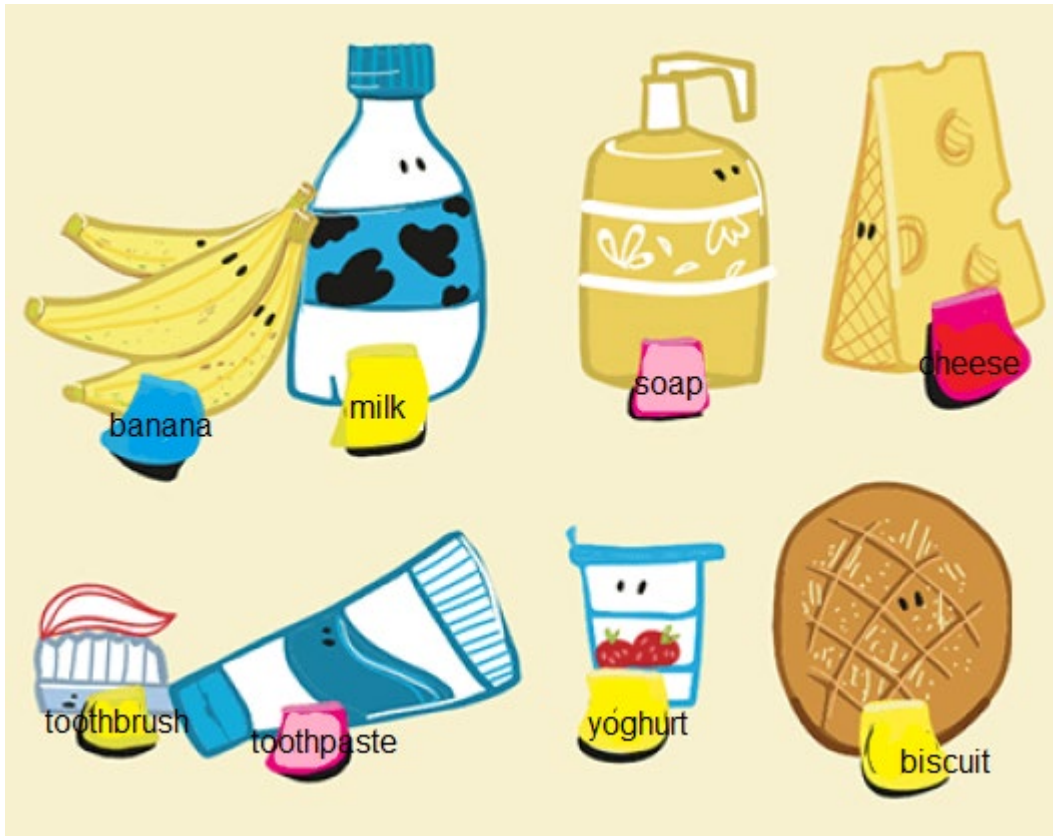
**Paragraph:** The term M. (from Lat. **Foreign:** manu scriptum) ...

Ask computer: look for all Foreign elements  
Answer: *manu scriptum*

**Manuscripts**

The term M. (from Lat. *manu scriptum*) ...

# Markup languages



1. Assign descriptions to each word = label (sticky note) = **tag**
2. Use a predefined vocabulary = **set of tags**
3. Can be understood (processed) by humans and computers
4. Must be logical and consistent