

TALOS ERA CHAIR IN ARTIFICIAL INTELLIGENCE FOR HUMANITIES AND SOCIAL SCIENCES



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE



LEAF

Linked Editing
Academic Framework

An open-source tool for text
encoding

Rachel Milio

PhD Candidate, RT1

TALOS-AI4SSH

rachel.milio@outlook.com



Funded by the
European Union



<https://talos-ai4ssh.uoc.gr/>

Horizon ERA Chair TALOS AI4SSH Project funded by the European Commission
Grant Agreement n° 101087269, <https://cordis.europa.eu/project/id/101087269>



What is text encoding?

Making human
readable text
machine readable

AKA
“markup”

Types of mark-up:

- Structural
- Presentational
- Semantic



What is XML?

XML = eXtensible Markup Language

- Structure and label data and metadata
- Extensible = no predefined tags

Extensible Markup

Pros	Cons
Flexibility and freedom for users	Less interoperability/reusability

```
<place>Crete</place>  
  <loc>Crete</loc>  
<placename>Crete</placename>
```

What is TEI?

Text Encoding Initiative

Humanistic
standard for
XML

Used for literary & linguistic texts
Manuscripts, historical archives,
critical editions, anthologies,
inscriptions...

Why?

More consistent text
encoding
Increase interoperability and
reusability



```
<persName  
  ref='https://www  
  .wikidata.org/wi  
  ki/Q164061'>
```

```
<forename>Helen</  
forename>of  
<placeName  
  ref='https://www  
  .wikidata.org/wi  
  ki/Q5690'>Sparta
```

```
</placeName>  
</persName>
```

<https://tei-c.org/>

“Helen of Sparta”

persName
= full name
of person

Original
text string

forename =
first name
("Helen")

placeName =
location
("Sparta")

Attribute = describes
an element

ref = attribute
for external
reference

Element = unit of text = tag

LEAF-Writer

- Web-based text editor
- Free; no configuration or installation needed!
- Supports sharing and collaboration
- TEI schema support
- On-the-fly validation
- Entity tagging



LEAF

**Linked Editing
Academic Framework**

<https://leaf-writer.leaf-vre.org/>

SIGN IN WITH



Github

OPEN



From the cloud



From your device



Import document



Templates



Samples



or paste your XML here

SAMPLES



Bertrand Russell to Patric...



To Congress on women's s...



A.M. Klein, "Political Mee...



Virginia Woolf, "Mrs. Dall...



- teiHeader
- text
 - body
 - div
 - head
 - opener
 - p
 - p
 - pb
 - p
 - p
 - closer

Bertrand Russell to Patricia Spence - October 21, 1935

NOTE:

Bad writing due to shaky train

In train

Oslo to Bergen

21.10.35

Dearest -

I have had no letter from you since I left Stockholm, but I had a nice one from John in an envelope you had sent him. I had sent him one addressed to Copenhagen but he hadn't used it.

When I reached Oslo yesterday evening, Brynjulf Bull should have been there to meet me, but wasn't. He is not on the telephone, so I took a taxi to his address, which turned out to be a students' club with no one about on Sundays, so I went to a hotel feeling rather non-plussed. But presently he turned up. He had got the

2

time of my arrival wrong, and when [sic] when he had found he had missed me he phoned to every hotel in Oslo till he hit on the right one. He left me at 10, and then I had to do a Sunday Referee article. Today my journey lasts from 9 till 9 - fortunately one of the most beautiful railway journeys in the world. Tomorrow I lecture at Bergen to the Anglo-Norwegian Society. Next day I go back to Oslo, lecture there Fri. and Sat. and then start for home via Bergen.

3

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="https://tei-c.org/release/
xml/tei/custom/schema/relaxng/tei_all.rng"
type="application/xml"
schematypens="http://relaxng.org/ns/
structure/1.0"?>
<?xml-stylesheet type="text/css" href="https://
cwrc.ca/templates/css/tei.css"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Sample Document Title</
        title>
      </titleStmt>
      <publicationStmt>
        <p/>
      </publicationStmt>
      <sourceDesc sameAs="http://www.
cwrc.ca">
        <p>Created from original
        research by members of
        CWRC/CS C unless otherwise
        noted.</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <div type="letter">
        <head>
          <title>Bertrand Russell to
          Patricia Spence -
          October 21, 1935</
          title>
        </head>
        <opener>
          <note>
            <p>Bad writing due to
            shaky train</
            p><p>In train</
            p><p>
            Oslo to Bergen</p>
          </note>
        </opener>
      </div>
    </body>
  </text>
</TEI>
```

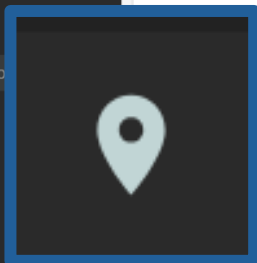
EDIT





TEI

- teiHeader
- text
 - body
 - div
 - head
 - opener
 - p I Have Had No Letter From You
 - p
 - pb
 - p
 - closer



Bertrand Russell to Patricia Spence - October 21, 1935

NOTE:

Bad writing due to shaky train

In train

Oslo to Bergen

21.10.35

Dearest -

I have had no letter from you since I left **Stockholm**, but I had a nice one from John in an envelope you had sent him. I had sent him one addressed to Copenhagen but he hadn't used it.

When I reached Oslo yesterday evening, Brynjulf Bull should have been there to meet me, but wasn't. He is not on the telephone, so I took a taxi to his address, which turned out to be a students' club with no one about on Sundays, so I went to a hotel feeling rather non-plussed. But presently he turned up. He had got the

2

time of my arrival wrong, and **when [sic] when** he had found he had missed me he phoned to every hotel in Oslo till he hit on the right one. He left me at 10, and then I had to do a Sunday Referee article. Today my journey lasts from 9 till 9 - fortunately one of the most beautiful railway journeys in the world. Tomorrow I lecture at Bergen to the Anglo-Norwegian Society. Next day I go back to Oslo, lecture there Fri. and Sat. and then start for home via Bergen.

3

Stockholm

EDIT



Find Place

Stockholm



WIKIDATA

Stockholm

capital and largest city of Sweden

Stockholm

municipality in Stockholm County, Sweden

Stockholm University

state university of Stockholm, Sweden

Stockholm

family name

Stockholm County

county in Sweden

Stockholm City

urban municipality in Sweden until the end of 1970

1912 Summer Olympics

Games of the V Olympiad, celebrated in Stockholm (Sweden) in 1912

Wikidata 10

LINCS 2

Other

CANCEL

TAG WITHOUT LINKING

SELECT

Tag Place

SELECTED TEXT

Stockholm

TAG AS



Stockholm

source: [wikidata](#)

LEVEL OF CERTAINTY

High

Medium

Low

Unknown

PRECISION OF LOCATION OF PLACE NAME

High

Medium

Low

Unknown

CERT Certainty

high

(certainty) signifies the degree of certainty associated with the intervention or interpretation.

KEY

Stockholm

Attributes

ana

Analysis

cert

Certainty

change

copyOf

corresp

Corresponds

datingMethod

datingPoint

evidence

exclude

facs

Facsimile

from

from-custom

from-iso

full

CANCEL

OK





editing ▼

Table of Contents	Markup	Entities
-------------------	--------	----------

Stockholm

Standard: Stockholm

URI: <http://www.wikidata.org/entity/Q1754>

cert: high

precision: high

 Tag
 












Bad writing due to shaky train

In train

Oslo to Bergen

21.10.35

Dearest -

I have had no letter from you since I left **Stockholm**, but I had a nice one from John in an envelope you had sent him. I had sent him one addressed to Copenhagen but he hadn't used it.

When I reached Oslo yesterday evening, Brynjulf Bull should have been there to meet me, but wasn't. He is not on the telephone, so I took a taxi to his address, which turned out to be a students' club with no one about on Sundays, so I went to a hotel feeling rather non-plussed. But presently he turned up. He had got the

2

time of my arrival wrong, and **when** [sic]**when** he had found he had missed me he phoned to every hotel in Oslo till he hit on the right one. He left me at 10, and then I had to do a Sunday Referee article. Today my journey lasts from 9 till 9 - fortunately one of the most beautiful railway journeys in the world. Tomorrow I lecture at Bergen to the Anglo-Norwegian Society. Next day I go back to Oslo, lecture there Fri. and Sat. and then start for home via Bergen.

2

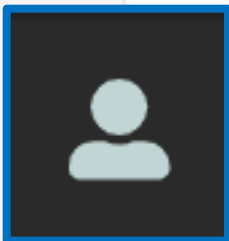
Scrape Candidate Entities

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="https://tei-c.org/release/
  xml/tei/custom/schema/relaxng/tei_all.rng"
  type="application/xml"
  schematypens="http://relaxng.org/ns/
  structure/1.0"?>
<?xml-stylesheet type="text/css" href="https://
  cwrc.ca/templates/css/tei.css"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmnt>
        <title>Sample Document Title</
          title>
        </titleStmnt>
        <publicationStmnt>
          <p/>
        </publicationStmnt>
        <sourceDesc sameAs="http://www.
          cwrc.ca">
          <p>Created from original
            research by members of
            WRC/CSEC unless otherwise
            noted.</p>
        </sourceDesc>
      </fileDesc>
    </teiHeader>
    <text>
      <body>
        <div type="letter">
          <head>
            <title>Bertrand Russell to
              Patricia Spence -
              October 21, 1935</
                title>
            </head>
            <opener>
              <note>
                <p>Bad writing due to
                  shaky train</
                    p><p>In train</
                      p><p>
                        Oslo to Bergen</p>

```

EDIT

Bugs / Requests LEAF-Writer 3.6.0 Powered by Tiny



Bertrand Russell to Patricia Spence - October 21, 1935

NOTE:

Bad writing due to shaky train

In train

Oslo to Bergen

21.10.35

Dearest -

I have had no letter from you since I left **Stockholm**, but I had a nice one from John in an envelope you had sent him. I had sent him one addressed to Copenhagen but he hadn't used it.

When I reached Oslo yesterday evening, Brynjulf Bull should have been there to meet me, but wasn't. He is not on the telephone, so I took a taxi to his address, which turned out to be a students' club with no one about on Sundays, so I went to a hotel feeling rather non-plussed. But presently he turned up. He had got the

2

time of my arrival wrong, and **when** **[sic]** **when** he had found he had missed me he phoned to every hotel in Oslo till he hit on the right one. He left me at 10, and then I had to do a Sunday Referee article. Today my journey lasts from 9 till 9 - fortunately one of the most beautiful railway journeys in the world. Tomorrow I lecture at Bergen to the Anglo-Norwegian Society. Next day I go back to Oslo, lecture there Fri. and Sat. and then start for home via Bergen.

3

Scrape Candidate Entities

Bertrand Russell

EDIT



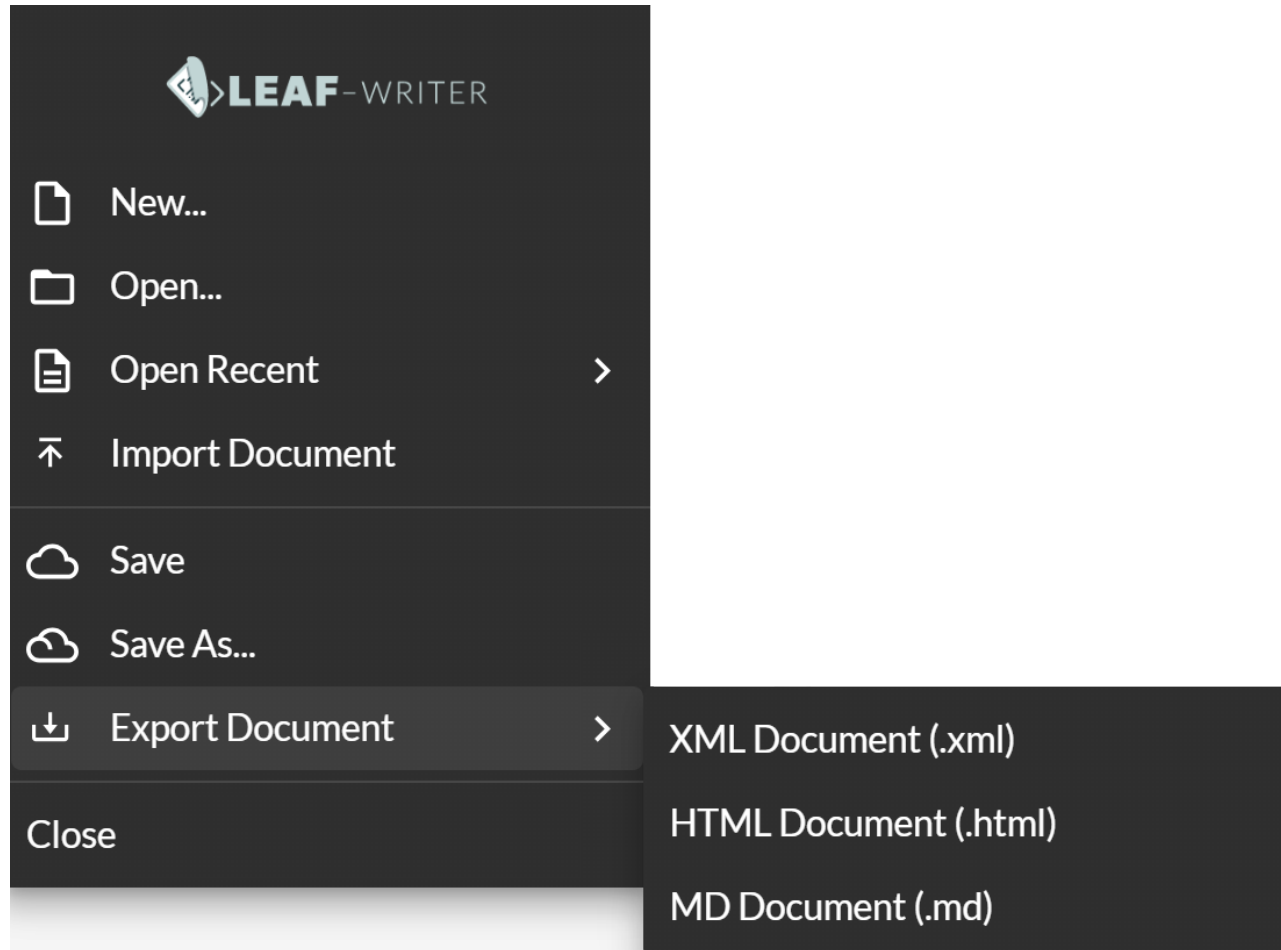


```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="https://tei-c.org/release/
xml/tei/custom/schema/relaxng/tei_all.rng"
type="application/xml"
schematypens="http://relaxng.org/ns/
structure/1.0"?>
<?xml-stylesheet type="text/css" href="https://
cwrc.ca/templates/css/tei.css"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Sample Document Title</
          title>
        </titleStmt>
        <publicationStmt>
          <p>
        </publicationStmt>
        <sourceDesc sameAs="http://www.
          cwrc.ca">
          <p>Created from original
            research by members of
            CWRC/CS C unless otherwise
            noted.</p>
        </sourceDesc>
      </fileDesc>
    </teiHeader>
    <text>
      <body>
        <div type="letter">
          <head>
            <title><persName
              key="Bertrand Russell"
              ref="http://www.
                wikidata.org/entity/
                Q33760" cert="high"
              type="real">Bertrand
                Russell</persName> to
                Patricia Spence -
                October 21, 1935</
              title>
            </head>
          <opener>
```

EDIT

Bugs / Requests LEAF-Writer 3.6.0 Powered by Tiny

Exporting from LEAF-Writer



XML: eXtensible
Markup Language
HTML: Hypertext
Markup Language
MD: Markdown

To Learn More

<https://www.leaf-vre.org/>

<https://tei-c.org/>



DTToC
Dynamic Table of Contexts



TALOS ERA CHAIR IN ARTIFICIAL INTELLIGENCE FOR HUMANITIES AND SOCIAL SCIENCES



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE



Thank you!

Rachel Milio

PhD Candidate, RT1

TALOS-AI4SSH

rachel.milio@outlook.com



<https://talos-ai4ssh.uoc.gr/>



Funded by the
European Union

Horizon ERA Chair TALOS AI4SSH Project funded by the European Commission
Grant Agreement n° 101087269, <https://cordis.europa.eu/project/id/101087269>

