



Title	Essential Material for Humanities Researchers! Envisioning a New Path in Humanities Research with Text Data and TEI
Author(s)	Vogatza, Dimitra; Voulgaris, Nikolaos; Yoshiga, Natsuko et al.
Citation	
Version Type	VoR
URL	https://doi.org/10.18910/104315
rights	This article is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

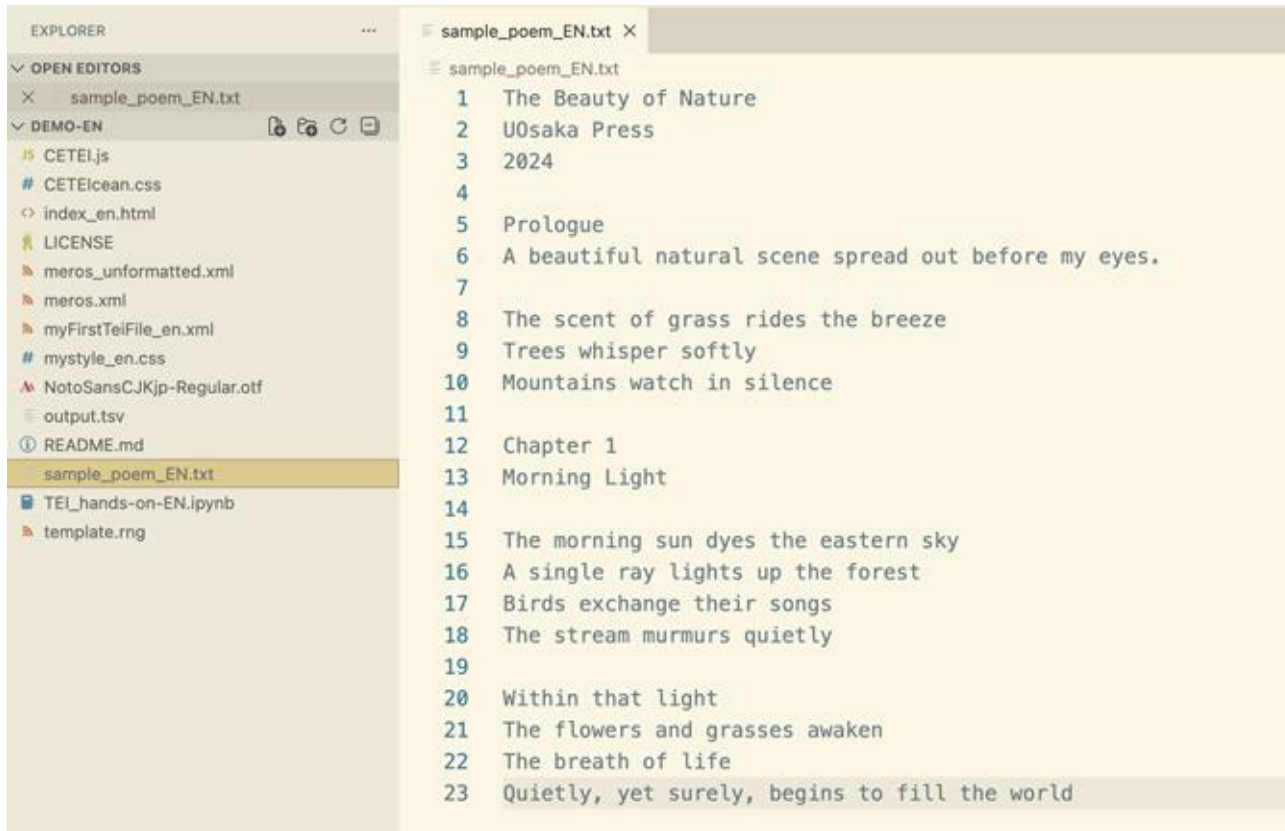
<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

TEI Hands-on

The Basics

Preparing sample_poem_EN.txt



```
EXPLORER
...
OPEN EDITORS
X sample_poem_EN.txt
DEMO-EN
  CETEI.js
  CETEIclean.css
  index_en.html
  LICENSE
  meros_unformatted.xml
  meros.xml
  myFirstTeiFile_en.xml
  mystyle_en.css
  NotoSansCJKjp-Regular.otf
  output.tsv
  README.md
  sample_poem_EN.txt
  TEI_hands-on-EN.ipynb
  template.rng

sample_poem_EN.txt X
sample_poem_EN.txt
1 The Beauty of Nature
2 UOsaka Press
3 2024
4
5 Prologue
6 A beautiful natural scene spread out before my eyes.
7
8 The scent of grass rides the breeze
9 Trees whisper softly
10 Mountains watch in silence
11
12 Chapter 1
13 Morning Light
14
15 The morning sun dyes the eastern sky
16 A single ray lights up the forest
17 Birds exchange their songs
18 The stream murmurs quietly
19
20 Within that light
21 The flowers and grasses awaken
22 The breath of life
23 Quietly, yet surely, begins to fill the world
```

Selecting the Text and Creating the Basic Structure

The Beauty of Nature
Hanako Yamada
UOsaka Press
2024

Prologue

A beautiful natural scene spread out
before my eyes.

The scent of grass rides the breeze
Trees whisper softly
Mountains watch in silence

Tag:

`<xxxxxx>`A block of text`</xxxxxx>`

TEI Tagging in Practice 1

<TEI>

Indicates that the document is structured in TEI format.

<teiHeader>

Indicates that the content concerns bibliographic information about the text

<text>

Indicates that the actual content of the text is written here

```
<TEI>
  <teiHeader>
    The Beauty of Nature
    Hanako Yamada
    UOsaka Press
    2024
  </teiHeader>
  <text>
    Prologue
    A beautiful natural scene spread out
    before my eyes.
    The scent of grass rides the breeze
    Trees whisper softly
    Mountains watch in silence
  </text>
</TEI>
```

TEI Tagging in Practice 2

TEI Guidelines: <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

 Guidelines ▾ Activities ▾ Tools ▾ Membership ▾ Support ▾ About ▾ News

TEI: Guidelines for Electronic Text Encoding and Interchange

P5 Version 4.9.0. Last updated on 24th January 2025, revision f73186978

[\[English\]](#) [\[Deutsch\]](#) [\[Español\]](#) [\[Italiano\]](#) [\[Français\]](#) [\[日本語\]](#) [\[한국어\]](#) [\[中文\]](#)

Front Matter

- [Title](#)
- i. [Releases of the TEI Guidelines](#)
- ii. [Dedication](#)
- iii. [Preface and Acknowledgments](#)
- iv. [About These Guidelines](#)
- v. [A Gentle Introduction to XML](#)
- vi. [Languages and Character Sets](#)

Text Body

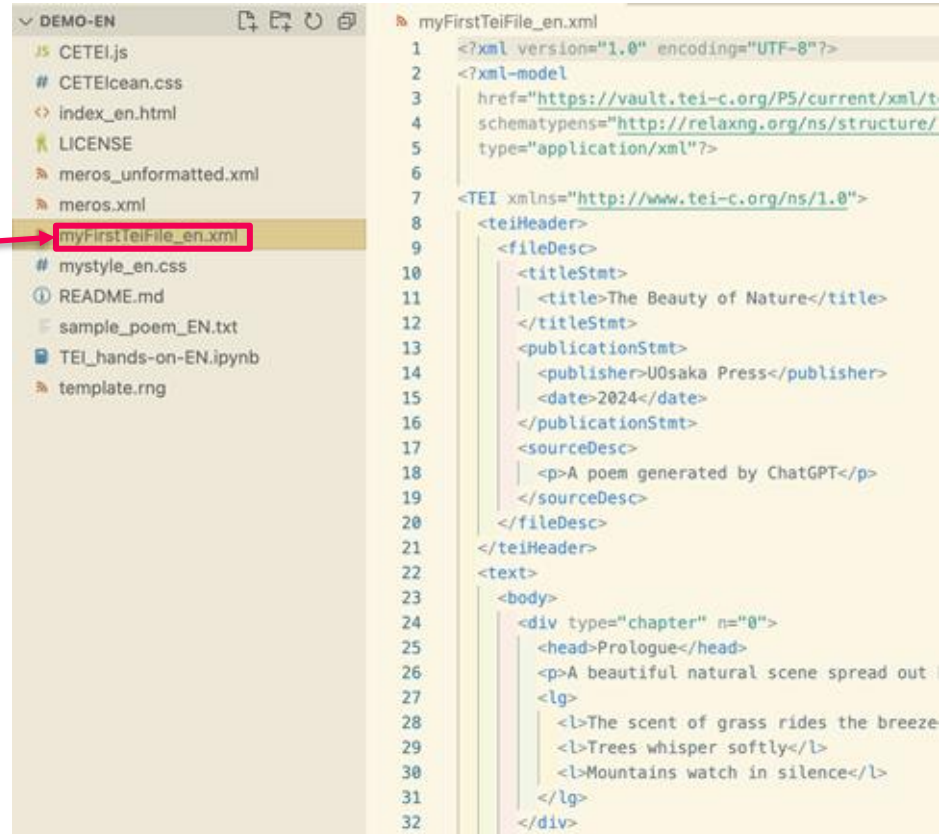
- 1 [The TEI Infrastructure](#)
- 2 [The TEI Header](#)
- 3 [Elements Available in All TEI Documents](#)
- 4 [Default Text Structure](#)
- 5 [Characters, Glyphs, and Writing Modes](#)
- 6 [Verse](#)
- 7 [Performance Texts](#)
- 8 [Transcriptions of Speech](#)
- 9 [Computer-mediated Communication](#)

TEI sourcecode

- [Getting and Using the TEI Sources](#)
- [TEI GitHub Repository](#)
- [Bug Reports, Feature Requests, etc.](#)

Opening a Sample File

Example file:
MyFirstTeiFile_en.xml



```
DEMO-EN
├── CETEI.js
├── CETEIclean.css
├── index_en.html
├── LICENSE
├── meros_unformatted.xml
├── meros.xml
├── myFirstTeiFile_en.xml
├── mystyle_en.css
├── README.md
├── sample_poem_EN.txt
├── TEI_hands-on-EN.ipynb
└── template.rng
```

```
myFirstTeiFile_en.xml
1  <?xml version="1.0" encoding="UTF-8"?>
2  <?xml-model
3    href="https://vault.tei-c.org/P5/current/xml/t
4    schematypens="http://relaxng.org/ns/structure/
5    type="application/xml"?>
6
7  <TEI xmlns="http://www.tei-c.org/ns/1.0">
8    <teiHeader>
9      <fileDesc>
10        <titleStmt>
11          <title>The Beauty of Nature</title>
12        </titleStmt>
13        <publicationStmt>
14          <publisher>UOsaka Press</publisher>
15          <date>2024</date>
16        </publicationStmt>
17        <sourceDesc>
18          <p>A poem generated by ChatGPT</p>
19        </sourceDesc>
20      </fileDesc>
21    </teiHeader>
22    <text>
23      <body>
24        <div type="chapter" n="0">
25          <head>Prologue</head>
26          <p>A beautiful natural scene spread out
27          <lg>
28            <l>The scent of grass rides the breeze
29            <l>Trees whisper softly</l>
30            <l>Mountains watch in silence</l>
31          </lg>
32        </div>
```

TEI Tagging in Practice 3

Chapter/section title--- **<head>**

Paragraph ----- **<p>**

Line of poetry----- **<l>**

Stanza (group of lines)- **<lg>**

```
<TEI>
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>The Beauty of Nature</title>
        <author>Hanako Yamada</author>
      </titleStmt>
      <publicationStmt>
        <publisher>UOsaka Press</publisher>
        <date>2024</date>
      </publicationStmt>
      <sourceDesc>
        <p>A poem automatically generated by
ChatGPT</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <head>Prologue</head>
      <p>A beautiful natural scene spread out before
my eyes.</p>
      <lg>
        <l>The scent of grass rides the breeze</l>
        <l>Trees whisper softly</l>
        <l>Mountains watch in silence</l>
      </lg>
    </body>
  </text>
</TEI>
```

TEI Tagging in Practice 4

Chapter 1

Morning Light

The morning sun dyes the eastern sky

A single ray lights up the forest
Birds exchange their songs
The stream murmurs quietly

Within that light
The flowers and grasses awaken
The breath of life
Quietly, yet surely, begins to fill the world

Attribute

1. type

- Specifies the type or category of the element.

2. n

- Assigns an identifier to the element to simplify ordering or referencing.

```
<text>
<body>
  <div type="chapter" n="0">
    <head>Prologue</head>
    <p>A beautiful natural scene spread out
before my eyes.</p>
    <lg>
      <l>The scent of grass rides the breeze</l>
      <l>Trees whisper softly</l>
      <l>Mountains watch in silence</l>
    </lg>
  </div>
  <div type="chapter" n="1">
    <head>Chapter 1</head>
    <title>Morning Light</title>
    <lg>
      <l>The morning sun dyes the eastern sky</l>
      <l>A single ray lights up the forest</l>
      <l>Birds exchange their songs</l>
      <l>The stream murmurs quietly</l>
    </lg>
    <lg>
      <l>Within that light</l>
      <l>The flowers and grasses awaken</l>
      <l>The breath of life</l>
      <l>Quietly, yet surely, begins to fill the
world</l>
    </lg>
  </div>
</body>
</text>
```

Manual Tagging Result

```
<TEI>
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>The Beauty of Nature</title>
      <author>Hanako Yamada</author>
    </titleStmt>
    <publicationStmt>
      <publisher>Osaka University Press</publisher>
      <date>2024</date>
    </publicationStmt>
    <sourceDesc>
      <p>A poem generated by ChatGPT</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
<text>
  <body>
    <div type="chapter" n="0">
      <head>Prologue</head>
      <p>A beautiful natural scene spread out before
my eyes.</p>
```

```
<lg>
  <l>The scent of grass rides the breeze</l>
  <l>Trees whisper softly</l>
  <l>Mountains watch in silence</l>
</lg>
</div>
<div type="chapter" n="1">
  <head>Chapter 1</head>
  <title>Morning Light</title>
  <lg>
    <l>The morning sun dyes the eastern sky</l>
    <l>A single ray lights up the forest</l>
    <l>Birds exchange their songs</l>
    <l>The stream murmurs quietly</l>
  </lg>
  <lg>
    <l>Within that light</l>
    <l>The flowers and grasses awaken</l>
    <l>The breath of life</l>
    <l>Quietly, yet surely, begins to fill the
world</l>
  </lg>
</div>
</body>
</text>
</TEI>
```

Creating TEI Files Easily & Accurately 1

Contents of “template.rng”

```
<?xml version="1.0" encoding="UTF-8"?>  
<?xml-model href="https://vault.tei-  
c.org/P5/current/xml/tei/custom/schema/relaxng/te  
i_all.rng"  
schematypens="http://relaxng.org/ns/structure/1.0"  
type="application/xml"?>
```

(Continues with <TEI> ... on the right)

Please copy from template.rng as much as possible

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">  
  <teiHeader>  
    <fileDesc>  
      <titleStmt>  
        <title></title>  
      </titleStmt>  
      <publicationStmt>  
        <publisher></publisher>  
      </publicationStmt>  
      <sourceDesc>  
        <p></p>  
      </sourceDesc>  
    </fileDesc>  
  </teiHeader>  
  <text>  
    <body>  
      <div></div>  
    </body>  
  </text>  
</TEI>
```

Creating TEI Files Easily & Accurately 2

Add the `<?xml ... ?>` declaration at the very top, before the `<TEI>` tag



```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <?xml-model
3   href="https://vault.tei-c.org/P5/current/xml/tei/custom/schema/relaxng/tei_all.rng"
4   schematypens="http://relaxng.org/ns/structure/1.0"
5   type="application/xml"?>
6
7 <TEI xmlns="http://www.tei-c.org/ns/1.0">
8   <teiHeader>
9     <fileDesc>
10      <titleStmt>
11        <title>The Beauty of Nature</title>
```

Creating TEI Files Easily & Accurately 3

Keep correcting it until the error disappears

```
7 <TEI xmlns="http://www.tei-c.org/ns/1.0">
8   <teiHeader>
9     <fileDesc>
10      <titleStmt>
11        <title>The Beauty of Nature</title>
12      </titleStm>
13      <publicationStmt>
14        <publisher>UOsaka Press</publisher>
15        <date>2024</date>
16      </publicationStmt>
```

Changing the appearance of TEI files

Drag and drop the file on the right into your file explorer

CETELcean

A convenient tool for displaying TEI-format files in a browser

<https://github.com/TEIC/CETELcean>

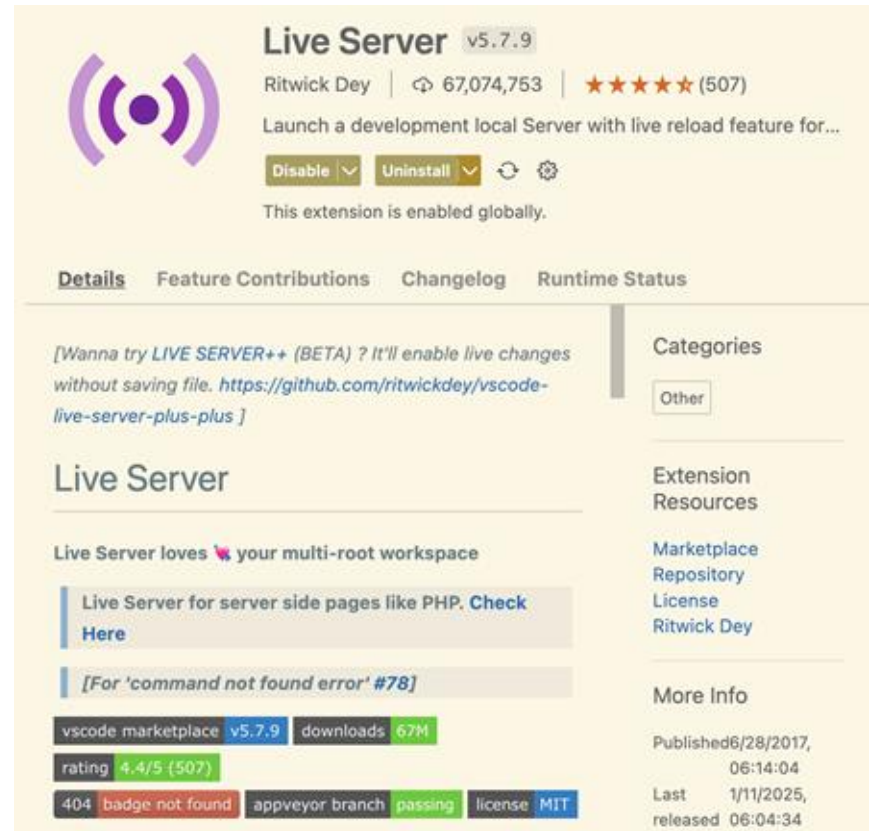
demo folder/TEI-tools

- CETEL.js
- CETELcean.css
- Index.html
- myFirstTeiFile_en.xml
- mystyle_en.css
- ...etc.

Live Server Extension

Live Server

Each time you save a file to be displayed in a web browser, it updates in real time



The screenshot shows the 'Live Server' extension page in the VS Code marketplace. The extension is by Ritwick Dey, version 5.7.9, with 67,074,753 downloads and a 4.4/5 rating from 507 reviews. It is currently installed and enabled globally. The page includes tabs for Details, Feature Contributions, Changelog, and Runtime Status. The Details tab is active, showing a description of the extension's functionality, a link to the GitHub repository, and a section for 'Live Server loves' your multi-root workspace. The 'Live Server loves' section includes a link to 'Live Server for server side pages like PHP' and a note about a 'command not found error'. The bottom of the page displays the extension's status in the marketplace, including the version, downloads, rating, and license (MIT).

Live Server v5.7.9
Ritwick Dey | 67,074,753 | ★★★★★ (507)
Launch a development local Server with live reload feature for...

Disable **Uninstall**  

This extension is enabled globally.

Details Feature Contributions Changelog Runtime Status

[Wanna try LIVE SERVER++ (BETA) ? It'll enable live changes without saving file. <https://github.com/ritwickdey/vscode-live-server-plus-plus>]

Live Server

Live Server loves 🍷 your multi-root workspace

Live Server for server side pages like PHP. [Check Here](#)

[For 'command not found error' #78]

vscode marketplace v5.7.9 downloads 67M
rating 4.4/5 (507)
404 badge not found appveyor branch passing license MIT

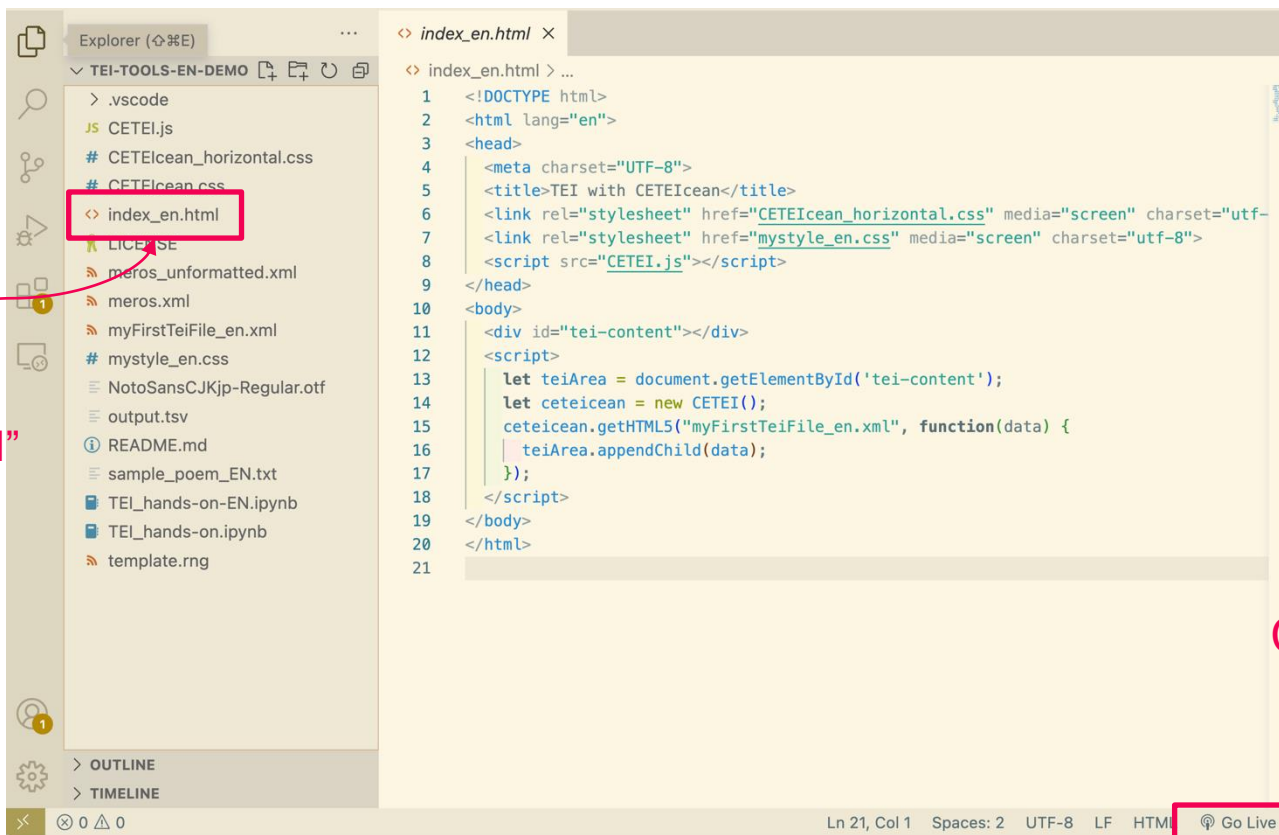
Categories
Other

Extension Resources
Marketplace
Repository
License
Ritwick Dey

More Info
Published 6/28/2017, 06:14:04
Last released 1/11/2025, 06:04:34

Launching the web browser

Choose
"index_en.html"



Click: "Go Live"

Displaying the Poem 1

Prologue

A beautiful natural scene spread out before my eyes.

The scent of grass rides the breeze

Trees whisper softly

Mountains watch in silence

Chapter 1

Morning Light

The morning sun dyes the eastern sky

A single ray lights up the forest

Birds exchange their songs

The stream murmurs quietly

Within that light

The flowers and grasses awaken

The breath of life

Quietly, yet surely, begins to fill the world

Displaying the Poem 2

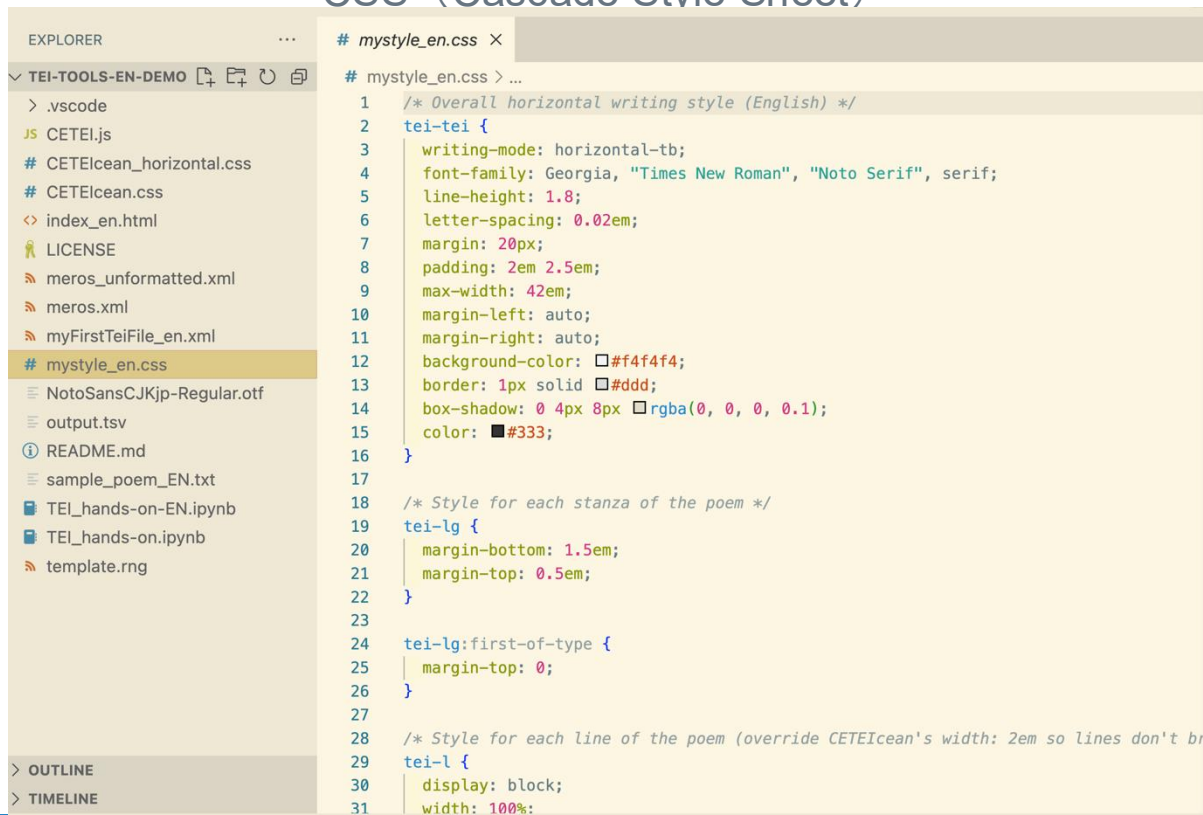
File in TEI
format



```
index_en.html > ...
1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4    <meta charset="UTF-8">
5    <title>TEI with CETEIcean</title>
6    <link rel="stylesheet" href="CETEIcean_horizontal.css" media="screen" charset="utf-8">
7    <link rel="stylesheet" href="mystyle_en.css" media="screen" charset="utf-8">
8    <script src="CETEI.js"></script>
9  </head>
10 <body>
11   <div id="tei-content"></div>
12   <script>
13     let teiArea = document.getElementById('tei-content');
14     let ceteicean = new CETEI();
15     ceteicean.getHTML5("myFirstTeiFile_en.xml", function(data) {
16       teiArea.appendChild(data);
17     });
18   </script>
19 </body>
20 </html>
```

Displaying the Poem 3

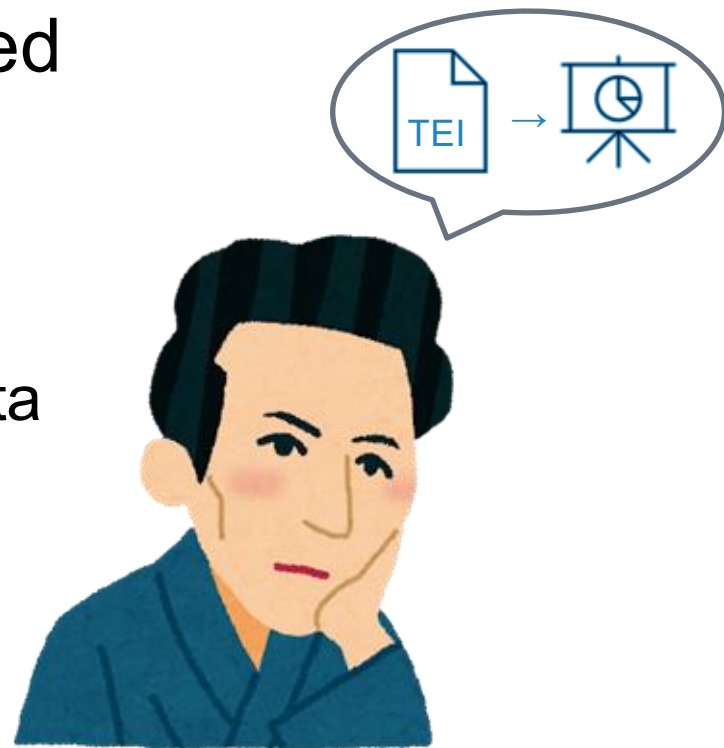
CSS (Cascade Style Sheet)



In Conclusion

Converting any text into structured TEI format using VSCode allows for...

- Flexible text display utilizing TEI data
- Visualization of the appearance frequency of specific words



Note: Vertical & Horizontal Text

The same procedure can be applied for vertical text.

Please refer to the respective files in the GitHub repository that do not contain the suffix EN/en for Japanese vertical writing.