



Title	人文学研究者必見! テキストデータとTEIで描く新たな研究ビジョン
Author(s)	吉賀, 夏子; 田畑, 智司; 甲斐, 尚人 他
Citation	
Version Type	VoR
URL	https://doi.org/10.18910/101978
rights	This content 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ハンズオン

基礎編

sample_poem.txtの準備

エクスプローラー

開いているエディター

- × sample_poem.txt U

DEMO

- > .git
- ◆ .gitignore
- JS CETEL.js A
- # CETELcean.css A
- <> index.html A
- 👤 LICENSE A
- 🔗 meros_unformatted.xml A
- 🔗 meros.xml A
- 🔗 myFirstTeiFile.xml A
- # mystyle.css A
- 📄 README.md A
- ☰ sample_poem.txt U
- 📄 TEI_hands-on.ipynb M
- 🔗 template.rng A

sample_poem.txt U ×

sample_poem.txt

- 1 自然の美
- 2 阪大出版
- 3 2024
- 4
- 5 序章
- 6 美しい自然が目の前に広がっていた。
- 7
- 8 草の香りが風に乗る
- 9 木々がささやく
- 10 山が静かに見守る
- 11
- 12 第1章 朝の光
- 13 朝日が東の空を染める
- 14 一筋の光が森を照らす
- 15 鳥たちがさえずりを交わし
- 16 小川が静かにさざめく
- 17
- 18 その光の中で
- 19 草花は目を覚まし
- 20 命の息吹が
- 21 静かに、しかし確かに満ちていく

テキストの選定と基本構造の作成

自然の美
山田花子
阪大出版
2024

序章
美しい自然が目の前に広がっていた。

草の香りが風に乗
木々がささやく
山が静かに見守る

タグ:

<xxxxxx> 文章や文字の塊 </xxxxxx>

タグ付けの実践1

<TEI>

このドキュメントはTEI形式に構造化されていることを示す

<teiHeader>

テキスト本文についての書誌情報であることを示す

<text>

実際のテキスト本文を記述していることを示す

```
<TEI>
  <teiHeader>
    自然の美
    山田花子
    阪大出版
    2024
  </teiHeader>
  <text>
    序章
    美しい自然が目の前に広がっていた。
    草の香りが風に乗る
    木々がささやく
    山が静かに見守る
  </text>
</TEI>
```

タグ付けの実践2

TEIガイドライン

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

TEI

TEI: TEIガイドライン

P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a

[English] [Deutsch] [Español] [Italiano] [Français] [日本語] [한국어] [中文]



前付

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

後付

本文

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

10 Dictionaries

11 Manuscript Description

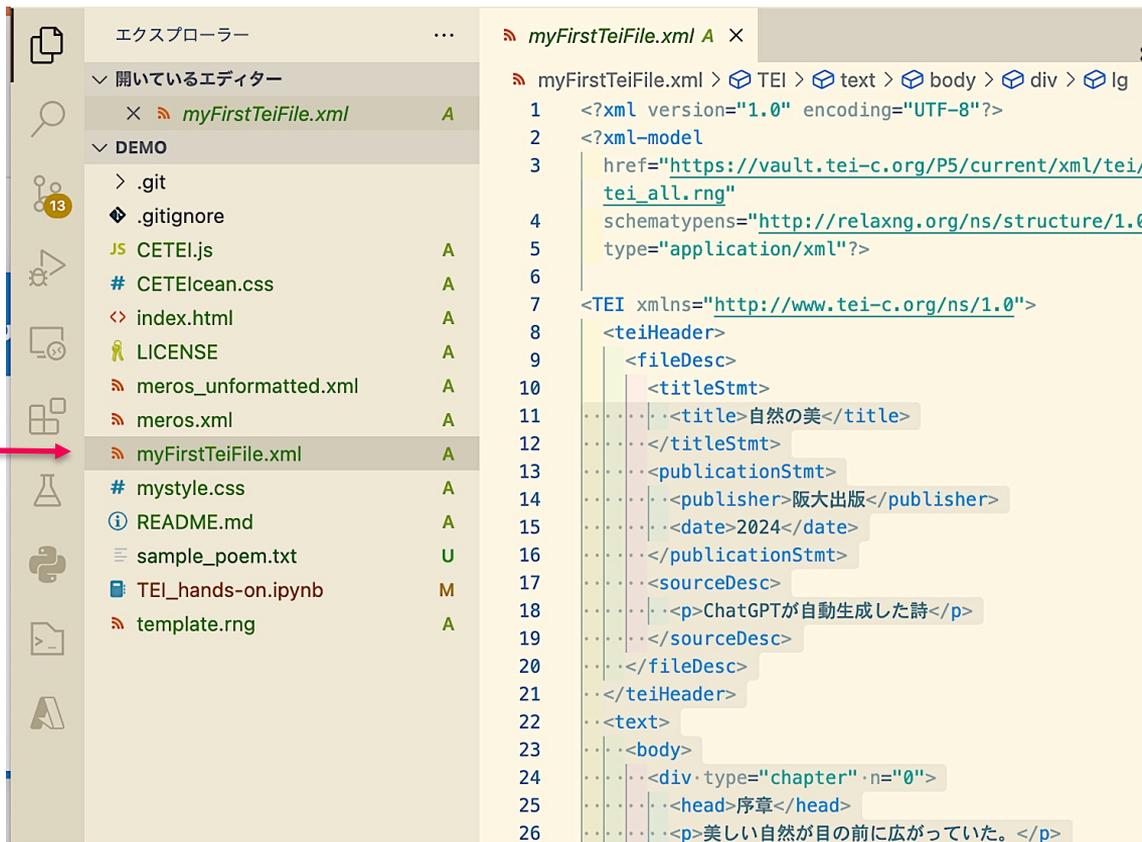
12 Representation of Primary Sources

13 Critical Apparatus

14 Names, Dates, People, and Places

お手本も開こう

お手本:
MyfirstTeiFile.xml



The screenshot shows a code editor interface. On the left is a file explorer with a tree view containing a project named 'DEMO'. The file 'myFirstTeiFile.xml' is selected and highlighted. A red arrow points from the text box on the left to this file. The main editor area displays the XML content of 'myFirstTeiFile.xml'. The XML is a TEI document with a file description and a body containing a chapter with a head and a paragraph.

```
myFirstTeiFile.xml A X
myFirstTeiFile.xml > TEI > text > body > div > lg
1 <?xml version="1.0" encoding="UTF-8"?>
2 <?xml-model
3 href="https://vault.tei-c.org/P5/current/xml/tei/
  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>自然の美</title>
12      </titleStmt>
13      <publicationStmt>
14        <publisher>阪大出版</publisher>
15        <date>2024</date>
16      </publicationStmt>
17      <sourceDesc>
18        <p>ChatGPTが自動生成した詩</p>
19      </sourceDesc>
20    </fileDesc>
21  </teiHeader>
22  <text>
23    <body>
24      <div type="chapter" n="0">
25        <head>序章</head>
26        <p>美しい自然が目の前に広がっていた。</p>
```

タグ付けの実践3

章やセクションのタイトル -----

`<head>`

パラグラフ ----- `<p>`

詩の各行 (line) ----- `<l>`

詩の各連 (stanza) - `<lg>`

```
<TEI>
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>自然の美</title>
      <author>山田花子</author>
    </titleStmt>
    <publicationStmt>
      <publisher>阪大出版</publisher>
      <date>2024</date>
    </publicationStmt>
    <sourceDesc>
      <p>ChatGPTが自動生成した詩</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
<text>
  <body>
    <head>序章</head>
    <p>美しい自然が目の前に広がっていた。</p>
    <lg>
      <l>草の香りが風に乗り</l>
      <l>木々がささやく</l>
      <l>山が静かに見守る</l>
    </lg>
  </body>
</text>
</TEI>
```

タグ付けの実践4

第1章

朝の光

朝日が東の空を染める
一筋の光が森を照らす
鳥たちがさえずりを交わし
小川が静かにさざめく

その光の中で
草花は目を覚まし
命の息吹が
静かに、しかし確かに満ちていく

属性 (attribute)

1. type

- 要素の種類や分類を明示的に指定する

2. n

- 要素に番号や識別子を付与して要素の順序や参照を簡単にする

```
<text>
<body>
  <div type="chapter" n="0">
    <head>序章</head>
    <p>美しい自然が目の前に広がっていた。</p>
    <lg>
      <l>草の香りが風に乗り</l>
      <l>木々がささやく</l>
      <l>山が静かに見守る</l>
    </lg>
  </div>
  <div type="chapter" n="1">
    <head>第1章</head>
    <title>朝の光</title>
    <lg>
      <l>朝日が東の空を染める</l>
      <l>一筋の光が森を照らす</l>
      <l>鳥たちがさえずりを交わし</l>
      <l>小川が静かにさざめく</l>
    </lg>
    <lg>
      <l>その光の中で</l>
      <l>草花は目を覚まし</l>
      <l>命の息吹が</l>
      <l>静かに、しかし確かに満ちていく</l>
    </lg>
  </div>
</body>
</text>
```

ひとまず全て手入力してみた結果

```
<TEI>
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>自然の美</title>
      <author>山田花子</author>
    </titleStmt>
    <publicationStmt>
      <publisher>阪大出版</publisher>
      <date>2024</date>
    </publicationStmt>
    <sourceDesc>
      <p>ChatGPTが自動生成した詩</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
<text>
  <body>
    <div type="chapter" n="0">
      <head>序章</head>
      <p>美しい自然が目の前に広がっていた。</p>
```

```
</lg>
  <l>草の香りが風に乗り</l>
  <l>木々がささやく</l>
  <l>山が静かに見守る</l>
</lg>
</div>
<div type="chapter" n="1">
  <head>第1章</head>
  <title>朝の光</title>
  <lg>
    <l>朝日が東の空を染める</l>
    <l>一筋の光が森を照らす</l>
    <l>鳥たちがさえずりを交わし</l>
    <l>小川が静かにさざめく</l>
  </lg>
  <lg>
    <l>その光の中で</l>
    <l>草花は目を覚まし</l>
    <l>命の息吹が</l>
    <l>静かに、しかし確かに満ちていく</l>
  </lg>
</div>
</body>
</text>
</TEI>
```

より楽&正確にTEIを 作成しよう1

“template.rng”の内容:

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

(右の<TEI>…に続く)

できるだけtemplate.rngをコピーしてください。

```
<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>
```

より楽&正確にTEIを作成しよう2

<TEI>の前に<?xml … … xml"?">の部分を追加する

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

より楽&正確にTEIを作成しよう3

エラーが消えるまで修正する

```
7 <TEI xmlns="http://www.tei-c.org/ns/1.0">
8   ..<teiHeader>
9     ..<fileDesc>
10      ..<titleStmt>
11       ..<title>自然の美</title>
12      ..</titleStm>
13     ..<publicationStmt>
14      ..<publisher>阪大出版</publisher>
15      ..<date>2024</date>
16     ..</publicationStmt>
```

TEIファイルの見た目を変えよう

右のファイルをエクスプローラにドラッグ&ドロップ

CETEIcean

TEI形式のファイルをブラウザで表示する便利なツール

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

demoフォルダ/TEI-tools

- CETEI.js
- CETEIcean.css
- Index.html
- myFirstTeiFile.xml
- Mystyle.css
- ...etc.

Live Server 機能拡張

Live Server

ウェブブラウザを使って表示するファイルを保存するたびにリアルタイムで更新する。



Live Server v5.7.9

Ritwick Dey | 56,120,524 | ★★★★★ (492)

Launch a development local Server with live reload fea...

無効にする アンインストール | 自動更新

詳細 機能 変更ログ

[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]

カテゴリ

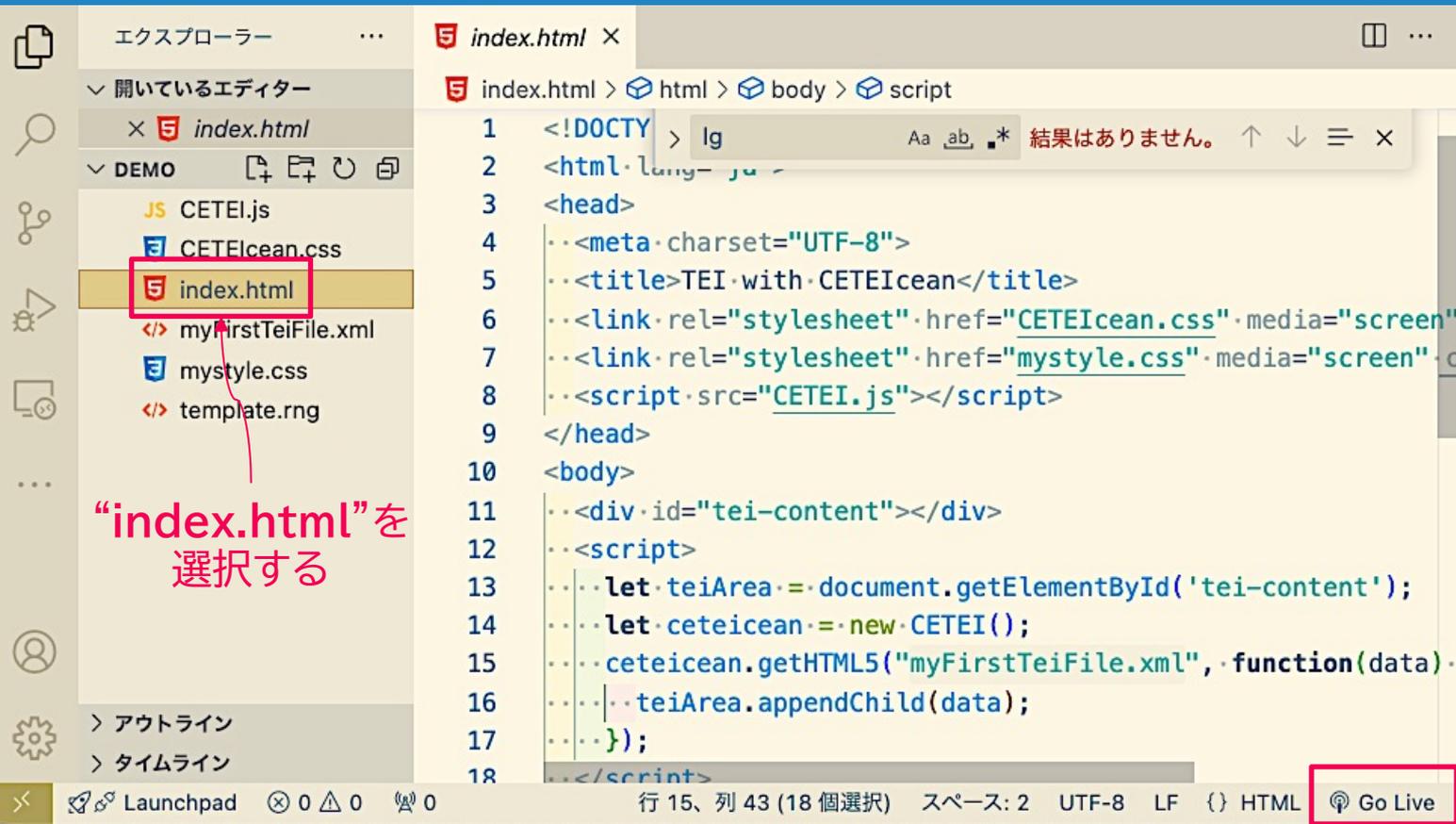
Other

リソース

マーケットプレイス
問題
リポジトリ
ライセンス
Ritwick Dey

詳細情報

ブラウザを起動する



“index.html”を
選択する

“Go Live”を
クリックする

Go Live

詩を表示する1

序章

美しい自然が目の前に広がっていた。

草の香りが風に乗り

木々がささやく

山が静かに見守る

第1章 朝の光

朝日が東の空を染める

一筋の光が森を照らす

鳥たちがさえずりを交わし

小川が静かにさざめく

その光の中で

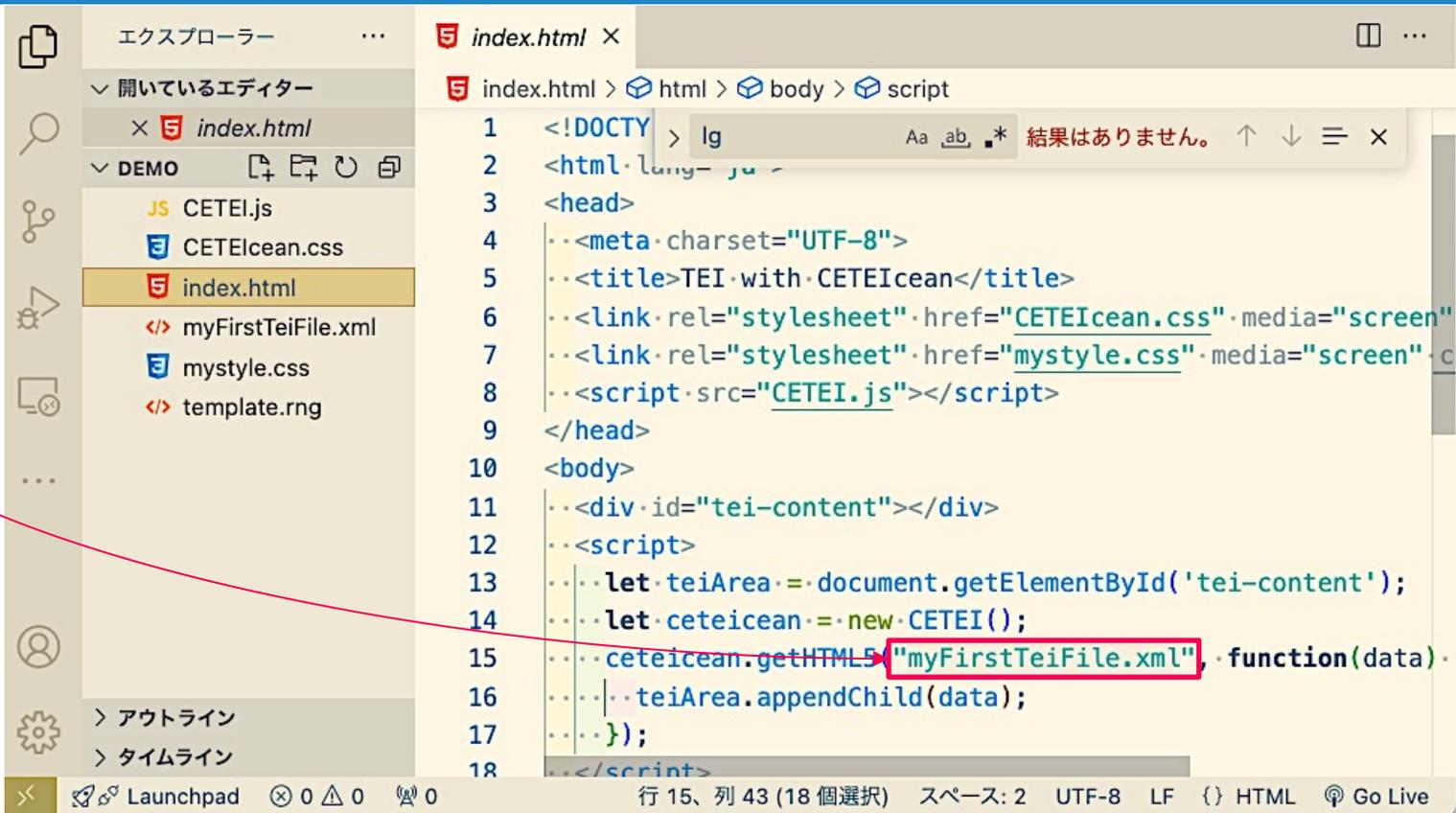
草花は目を覚まし

命の息吹が

静かに、しかし確かに満ちていく

詩を表示する2

TEI形式の
ファイル



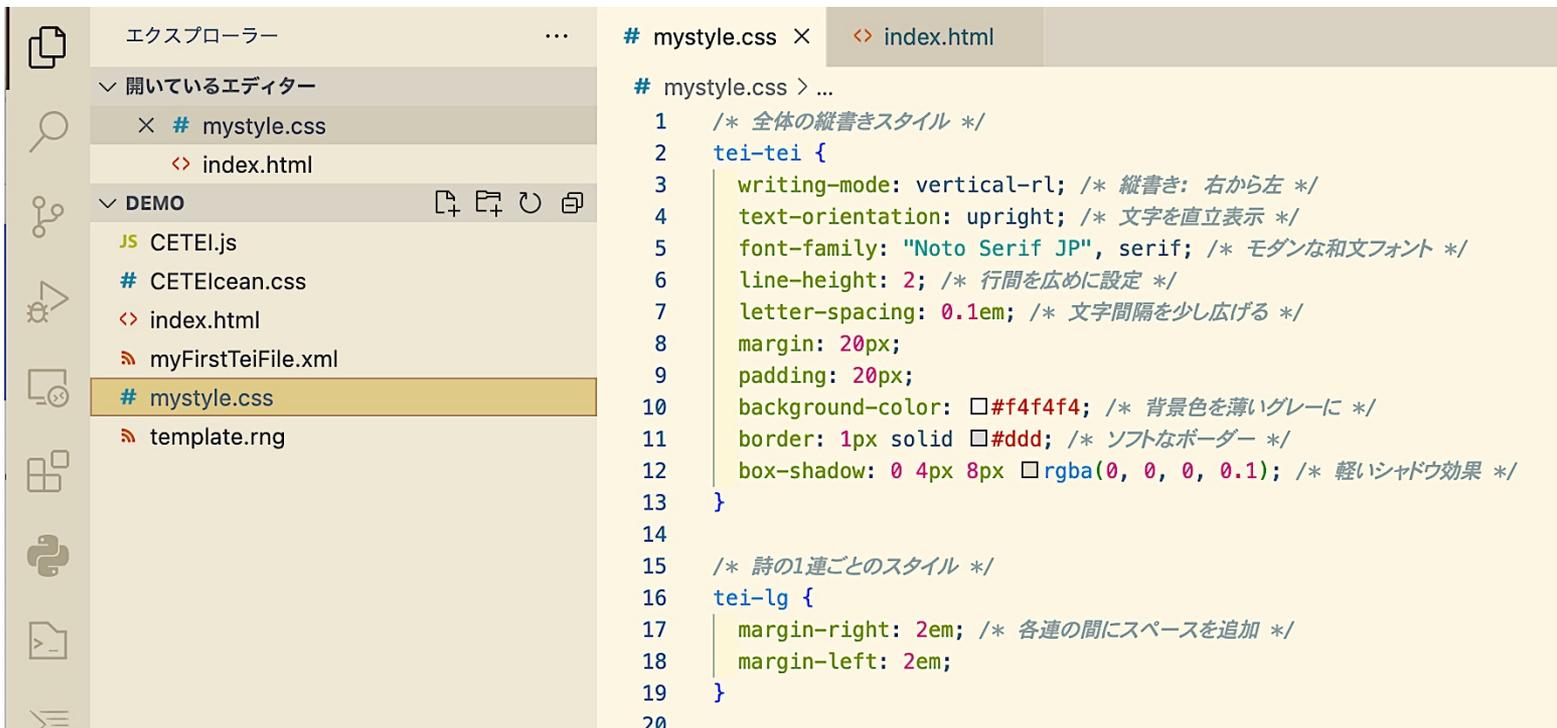
The screenshot shows a code editor with a file explorer on the left and a code editor on the right. The file explorer shows a project named 'DEMO' with several files: 'CETEI.js', 'CETEIcean.css', 'index.html', 'myFirstTeiFile.xml', 'mystyle.css', and 'template.rng'. The 'index.html' file is selected. The code editor shows the following HTML code:

```
1 <!DOCTYPE html>
2 <html lang="ja" >
3 <head>
4   <meta charset="UTF-8">
5   <title>TEI with CETEIcean</title>
6   <link rel="stylesheet" href="CETEIcean.css" media="screen">
7   <link rel="stylesheet" href="mystyle.css" media="screen">
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 CETEIcean();
15     ceteicean.getHTML5("myFirstTeiFile.xml", function(data) {
16       teiArea.appendChild(data);
17     });
18   </script>
```

A red box highlights the string "myFirstTeiFile.xml" in the code. A red arrow points from the text "TEI形式のファイル" to this box. The status bar at the bottom shows "行 15、列 43 (18 個選択) スペース: 2 UTF-8 LF {} HTML Go Live".

詩を表示する3

CSS(Cascade Style Sheet)



The image shows a code editor interface with a file explorer on the left and a code editor on the right. The file explorer shows a project named 'DEMO' with files like 'CETEL.js', 'CETELcean.css', 'index.html', 'myFirstTeiFile.xml', 'mystyle.css', and 'template.rng'. The code editor shows the content of 'mystyle.css'.

```
# mystyle.css > ...
1  /* 全体の縦書きスタイル */
2  tei-tei {
3      writing-mode: vertical-rl; /* 縦書き: 右から左 */
4      text-orientation: upright; /* 文字を直立表示 */
5      font-family: "Noto Serif JP", serif; /* モダンな和文フォント */
6      line-height: 2; /* 行間を広めに設定 */
7      letter-spacing: 0.1em; /* 文字間隔を少し広げる */
8      margin: 20px;
9      padding: 20px;
10     background-color: #f4f4f4; /* 背景色を薄いグレーに */
11     border: 1px solid #ddd; /* ソフトなボーダー */
12     box-shadow: 0 4px 8px rgba(0, 0, 0, 0.1); /* 軽いシャドウ効果 */
13 }
14
15 /* 詩の1連ごとのスタイル */
16 tei-lg {
17     margin-right: 2em; /* 各連の間にスペースを追加 */
18     margin-left: 2em;
19 }
20
```

さいごに

VSCoodeで任意のテキストをTEI形式の構造化データに変換すると…

- TEIデータを利用した柔軟なテキスト表示
- ある特定の単語の出現頻度を可視化

