Omnilex Code Assignment: Brazilian Civil Code

To goal is to scrape the page: https://www.planalto.gov.br/ccivil_03/Leis/2002/L10406.htm which is the official Brazilian civil code.

The output should be a .json file with an array of object with the following schema:

```
articleCitation : string;
articleContent : string;
sections : string[]
}
```

where articleCitation is the citation of the article like "art. 41"., articleContent its content like: "São pessoas jurídicas de direito público interno: I - a União; II - os Estados, o Distrito Federal e os Territórios; III - os Municípios; IV - as autarquias; IV - as autarquias, inclusive as associações públicas; (Redação dada pela Lei nº 11.107, de 2005) V - as demais entidades de caráter público criadas por lei." and sections the list of section this article is in: ["PARTE GERAL", "LIVRO I DAS PESSOAS", "TÍTULO II DAS PESSOAS JURÍDICAS", "CAPÍTULO I Disposições Gerais"].

The main challenge here is in keeping the right nested section path for each article, given the large size of the webpage. as well as avoiding revoked articles.

Constraints:

- Typescript with Playwright

Delivery,

- a link to a github repo with both the code and .json file



Presidência da República Casa Civil Subchefia para Assuntos Jurídicos

LEI Nº 10.406, DE 10 DE JANEIRO DE 2002

Texto compilado

INDIC

Vigência

Institui o Código Civil

Lei de Introdução às normas do Direito Brasileiro

(Vide Lei nº 14.195, de 2021)

Vide Lei nº 14.451, de 2022 Vigência

O PRESIDENTE DA REPÚBLICA Faço saber que o Congresso Nacional decreta e eu sanciono a seguinte Lei:

LIVEOL

DAS PESSOAS TÍTULO I DAS PESSOAS NATURAIS

CAPÍTULO I

Da Personalidade e da Capacidade

- Art. 1º Toda pessoa é capaz de direitos e deveres na ordem civil.
- Art. 2º A personalidade civil da pessoa começa do nascimento com vida; mas a lei põe a salvo, desde a concepção, os direitos do nascituro.
- Art. 3 ** São absolutamente incapazes de exercer pessoalmente os atos da vida civil:
- os menores de dezesseis anos:
- os que, por enfermidade ou deficiência mental, não tiverem o necessário discernimento para a prática desses atos
- ll os que, mesmo por causa transitória, não puderem exprimir sua vontade
- Art. 3º São absolutamente incapazes de exercer pessoalmente os atos da vida civil os menores de 16 (dezesseis) anos. (Redação dada pela Lei nº 13.146, de 2015) (Vigência)
 - I (Revogado); (Redação dada pela Lei nº 13.146, de 2015) (Vigência)
 - II (Revogado); (Redação dada pela Lei nº 13.146, de 2015) (Vigência)
- III (Revogado). (Redação dada pela Lei nº 13.146, de 2015) (Vigência
- Art. 4-9-São incapazes, relativamente a certos atos, ou à maneira de os exercer
- Art. 4º São incapazes, relativamente a certos atos ou à maneira de os exercer: (Redação dada pela Lei nº 13.146, de 2015) (Vigência)
 - I os maiores de dezesseis e menores de dezoito anos

TÍTULO II DAS PESSOAS JURÍDICAS

CAPÍTULO I Disposições Gerais

- Art. 40. As pessoas jurídicas são de direito público, interno ou externo, e de direito privado.
- Art. 41. São pessoas jurídicas de direito público interno:
- I a União;
- II os Estados, o Distrito Federal e os Territórios;
- III os Municípios:
- IV as autorquias
- IV as autarquias, inclusive as associações públicas; (Redação dada pela Lei nº 11.107, de 2005)
- V as demais entidades de caráter público criadas por lei.
- Parágrafo único. Salvo disposição em contrário, as pessoas jurídicas de direito público, a que se tenha dado estrutura de direito rivado, regem-se, no que couber, quanto ao seu funcionamento, pelas normas deste Código.
- Art. 42. São pessoas jurídicas de direito público externo os Estados estrangeiros e todas as pessoas que forem regidas pelo direito internacional público.
- Art. 43. As pessoas jurídicas de direito público interno são civilmente responsáveis por atos dos seus agentes que nessa qualidade causem danos a terceiros, ressalvado direito regressivo contra os causadores do dano, se houver, por parte destes, culpa ou dolo.
 - Art. 44. São pessoas jurídicas de direito privado
- I as associações;
- II as sociedades;