

The State Decoded Documentation

XML Format

The State Decoded parser can either be modified to use bespoke PHP to parse a legal code in its native format, or it can parse a prescribed XML format. This is a totally invented XML format. There is no legal code that is provided in this format. It is provided as an option because, via [XSLT](#), it can be substantially easier to convert an existing XML to this format than to write a parser for it in PHP.

The State Decoded includes a [sample XSLT file](#) on which you can base your own XSLT. It is well commented and verbose, so even if you're not familiar with XSLT, you've got a fighting chance of transforming your XML to The State Decoded's format. If the legal code that you are ingesting is in Lexis-Nexis' XML format (which sometimes travels under the name of "legislativedoc v1"), then use [lexis-nexis.xml](#), which is included with The State Decoded.

Note that this is *not* an effort to create a standard interchange or storage format for legal codes. This exists solely as a convenience for importing data into The State Decoded.

Terminology

Legal codes throughout the U.S. all use different terminology (e.g., "statutes" vs "code" vs "laws"). We need to adopt some terminology here, and while it can't be the right terminology everywhere, we hope to avoid any disputes about what exactly is the right terminology. The terminology we've adopted was the simplest that also reflected the reality of the data. A lot of people implementing The State Decoded are programmers, not attorneys, and the terminology that the project uses needs to reflect the reality of that user base. Below, the word "law" refers to what some jurisdictions would call a section, and it's usually prefixed with the \$ symbol. It corresponds to a single page on a State Decoded website. Within a "law", we use "section" elements to represent the hierarchical structure of the paragraphs within it. In some jurisdictions this would be called a subsection. A "section" corresponds to a paragraph plus any logically nested (typically indented) content within it.

Descriptive Specifications

- There must be one file per law.
- All files must be in the same directory.
- The names of the files are irrelevant.
- Every file contains both structural data about its container—which chapter/title/part/article/etc. that it's within—and data about the law itself.
- Text must be broken up by subsection, rather than provided as a single block of text, unless the text of the law does not have subsections.
- Subsections may be nested.

This being open source software, some relatively simple tweaks to `Parser::iterate` and `Parser::parse` make all of these specifications malleable.

Fields

- law: the container for all other fields—required
 - structure: the container for structural units (chapters, titles, articles, parts, etc.)—required
 - unit: a single structural unit that contains this law—required
 - label: the type of structural unit that this is (chapter, title, article, part, etc.)—required
 - identifier: the unique identifier for this structural unit (but not necessarily globally unique!), almost always a number (e.g., 18.2, 6)—required
 - orderby: when listing all structural units of this label, in what ordinal position this structural unit should appear—optional
 - level: a whole number, starting at one, indicating the depth of this structural unit—required
 - sectionnumber: the unique identifier for this law—required
 - catchline: the title of this law—required
 - orderby: when listing all laws within this structural unit, in what ordinal position this law should appear—optional
 - text: the container for the text of the law itself—required
 - section: a single labelled section of a law—optional
 - prefix: the identifying label for this section (e.g., A, 6, vi)—required
 - type: the type of material within this section; default is text, other options are "table" and "image"—optional
 - history: the textual description of the history of amendments to this law—optional
 - metadata: a container for any additional fields, which will be stored as key:value in the `laws_meta` table—optional; values of `true` and `false` will be stored in the database as `y` and `n`, respectively, but converted back to `true` and `false` within the internal and external APIs

Sample XML

Structure

With All Fields

```
<?xml version="1.0" encoding="utf-8"?>
<law>
  <structure>
    <unit label="" identifier="" order_by="" level=""></unit>
  </structure>
  <section_number></section_number>
  <catch_line></catch_line>
  <order_by></order_by>
  <text>
    <section prefix=""></section>
  </text>
  <history></history>
  <metadata></metadata>
  <tags>
    <tag></tag>
  </tags>
</law>
```

Bare Minimum

```
<?xml version="1.0" encoding="utf-8"?>
<law>
  <structure>
    <unit label="" identifier="" level=""></unit>
  </structure>
  <section_number></section_number>
  <catch_line></catch_line>
  <text></text>
</law>
```

Populated

Here is a complete, populated sample XML file that includes every available parameter.

```
<?xml version="1.0" encoding="utf-8"?>
<law>
  <structure>
    <unit label="title" identifier="18.2" order_by="18.2" level="1">Crimes and Offenses Generally</unit>
    <unit label="chapter" identifier="6" order_by="6" level="2">Crimes Involving Fraud</unit>
  </structure>
  <section_number>18.2-186</section_number>
  <catch_line>False statements to obtain property or credit.</catch_line>
  <order_by>00000000009104318.2</order_by>
  <text>
    <section prefix="1">
      Notwithstanding § 18.2-186:

      <section prefix="A">A person shall be guilty of a Class 1 misdemeanor if he makes,
      causes to be made or conspires to make directly, indirectly or through an agency, any
      materially false statement in writing, knowing it to be false and intending that it be
      relied upon, concerning the financial condition or means or ability to pay of himself,
      or of any other person for whom he is acting, or any firm or corporation in which he is
      interested or for which he is acting, for the purpose of procuring, for his own benefit
      or for the benefit of such person, firm or corporation, the delivery of personal
      property, the payment of cash, the making of a loan or credit, the extension of a
      credit, the discount of an account receivable, or the making, acceptance, discount, sale
      or endorsement of a bill of exchange or promissory note.</section>

      <section prefix="B">Any person who knows that a false statement has been made in writing
      concerning the financial condition or ability to pay of himself or of any person for
      whom he is acting, or any firm or corporation in which he is interested or for which he
      is acting and who, with intent to defraud, procures, upon the faith thereof.

      <section prefix="i" type="table">
        +-----+-----+
        | PERSON          | STATEMENT          |
        +-----+-----+
        | any              | false              |
        | no               | true               |
        +-----+-----+
      </section>

      For his own benefit, or for the benefit of the person, firm or corporation in which he
      is interested or for which he is acting, any such delivery, payment, loan, credit,
      extension, discount making, acceptance, sale or endorsement, shall, if the value of the
      thing or the amount of the loan, credit or benefit obtained is $200 or more, be guilty
      of grand larceny or, if the value is less than $200, be guilty of petit
      larceny.</section>

      <section prefix="C">Venue for the trial of any person charged with an offense under this
      section may be in the county or city in which (i) any act was performed in furtherance
      of the offense, or (ii) the person charged with the offense resided at the time of the
      offense.</section>
```

```
</text>
</section>
<section prefix="D">As used in this section, "in writing" shall include information
transmitted by computer, facsimile, e-mail, Internet, or any other electronic medium,
and shall not include information transmitted by any such medium by voice
transmission.</section>
</section>
<section prefix="2">
  This law shall expire on July 1, 2014.
</section>
</text>
<history>Code 1950, § 18.1-125; 1960, c. 358; 1975, cc. 14, 15.</history>
<metadata>
  <repealed>y</repealed>
  <expiration>2014-07-01</expiration>
</metadata>
<tags>
  <tag>lying</tag>
  <tag>fraud</tag>
  <tag>theft</tag>
</tags>
</law>
```

[Edit this on GitHub](#)

This project is maintained by [statedecoded](#)

Hosted on GitHub Pages — Theme by [orderedlist](#)