Akoma Ntoso to HTML5 transform

The HTML5 versions of UK legislation were designed to be (as near as possible) one-to-one representations of the Akoma Ntoso XML, while at the same time using the full semantics of HTML5. They also allow for correct presentation in a browser using CSS alone.

When appropriate, Akoma Ntoso elements are mapped to an HTML element of the same name. When an Akoma Ntoso element is mapped to an HTML element of another name, then the name of the Akoma Ntoso element will be the first value in the class attribute of the HTML element. If the Akoma Ntoso element is a generic element, which will always have a name attribute, then the value of the name attribute will be the second value of the class attribute of the HTML element. For example:

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <p> | <p> |
| <part> | <section class="part"> |
| <hcontainer name="regulation"> | <section class="hcontainer regulation"> |

There are only a few exceptions to a one-to-one element mapping:

* Footnotes are moved to the place at which they should appear, generally at the end of the document but sometimes immediately following a table.
* Punctuation surrounding and following quoted structures appears inline at the correct place in the document.
* TNA commentaries to revised legislation appear within the HTML5 document, generally at the end of each section, following the practice established on legislation.gov.uk.

A complete list of element mappings follows.

# Document level elements

The root element in any Akoma Ntoso document is the <akomaNtoso> element, which only ever has one child. In UK legislation, this child element is always an <act> element, with a “name” attribute containing the short-form document type.[[1]](#footnote-1) This Akoma Ntoso <act> element is mapped to the HTML5 <article> element with a “class” attribute containing three values: the name of the Akoma Ntoso element (“act”); “primary”, “secondary” or “euretained”, for each class of legislation; and the short document type (the values of the Akoma Ntoso “name” attribute).

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <akomaNtoso> | <html> |
| <act name="ukpga"> | <article class="act primary ukpga"> |
| <act name="uksi"> | <article class="act secondary uksi"> |

# Metadata

Akoma Ntoso metadata is represented in the body of HTML5 documents, using RDFa Lite. The hierarchy of the Akoma Ntoso metadata structure is represented by nested HTML <div> elements. Akoma Ntoso *container* elements are given a @resource attribute; “leaf” elements with attributes are given a @property attribute. The Akoma Ntoso metadata attribute values are represented by HTML <meta> elements.

For example:

|  |  |
| --- | --- |
| Akoma Ntoso | <meta>  <identification>  <FRBRWork>  <FRBRthis value="ukpga/2020/1"/>  <FRBRWork>  </identification>  </meata> |
| HTML5 | <div class="meta" vocab="…" style="display:none">  <div resource="#identification" typeof="identification">  <div resource="#FRBRWork" property="FRBRWork" typeof="FRBRWork">  <div property="FRBRthis" typeof="FRBRthis">  <meta property="value" content=" ukpga/2021/1">  </div>  </div>  </div>  </div> |

All values are contained in attributes, and there are no text nodes, so browsers should render no content. To avoid the appearance of blank space cause by nested <div> elements, a style attribute with the value "display:none" is placed on the root metadata element.

# Front matter (preamble, etc.)

The introductory material is represented by a simple structure of nested <div> elements. The main document title appears within an <h1> element, the only time that element is used.

## Primary

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <coverPage> | <div class="coverPage"> |
| <preface> | <div class="preface"> |
| <block name="title"> | <h1 class="block title"> |
| <block name="number"> | <div class="block number"> |
| <longTitle> | <div class="longTitle"> |
| <block name="dateOfEnactment"> | <div class="block dateOfEnactment"> |
| <preamble> | <div class="preamble"> |
| <formula name="enactingText"> | <div class="formula enactingText"> |

## Secondary

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <preface> | <div class="preface"> |
| <block name="banner"> | <div class="block banner"> |
| <block name="number"> | <div class="block number"> |
| <container name="subjects"> | <div class="container subjects"> |
| <container name="subject"> | <div class="container subject"> |
| <block name="subject"> | <div class="block subject"> |
| <block name="subsubject"> | <div class="block subsubject"> |
| <block name="title"> | <h1 class="block title"> |
| <container name="dates"> | <div class="container dates"> |
| <block name="laidInDraft"> | <div class="block laidInDraft"> |
| <block name="madeDate"> | <div class="block madeDate"> |
| <block name="laidDate"> | <div class="block laidDate"> |
| <block name="commenceDate"> | <div class="block commenceDate"> |
| <preamble> | <div class="preamble"> |
| <formula name="enactingText"> | <div class="formula enactingText"> |

# Main structural elements (core document hierarchy)

The “body” of an Akoma Ntoso document, is represented in HTML by a <div class="body">, because the HTML <body> element has special significance and must be used as a container for the entire document contents.

The HTML5 <section> element is used to represent the core divisions of an act, above the level of a Subsection, including Schedules but not including the Preamble or other introductory content. Numbered paragraphs below the level of a Subsection are represented with HTML5 <div> elements.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <body> | <div class="body"> |
| <hcontainer name="groupOfParts"> | <section class="hcontainer groupOfParts"> |
| <title>[[2]](#footnote-2) | <section class="title"> |
| <part> | <section class="part"> |
| <chapter> | <section class="chapter"> |
| <hcontainer name="crossheading"> | <section class="hcontainer crossheading"> |
| <hcontainer name="subheading"> | <section class="hcontainer subheading"> |
| <section> | <section class="section"> |
| <article> | <section class="article"> |
| <hcontainer name="regulation"> | <section class="hcontainer regulation"> |
| <rule> | <section class="rule"> |
| <subsection> | <section class="subsection"> |
| <level class="something"> | <div class="level something"> |
| <hcontainer name="schedules"> | <div class="hcontainer schedules"> |
| <hcontainer name="schedule"> | <section class="schedule"> |
| <paragraph>[[3]](#footnote-3) | <section class="paragraph"> |
| <subparagraph> | <section class="subparagrah"> |
| <hcontainer name="division"> | <div class="hcontainer division"> |
| <hcontainer name="definition"> | <div class="hcontainer definition"> |

## Numbers and headings of structural elements

The number and heading of each section are represented with HTML5 <span> elements. The headings of <section> elements, at and above the level of a Subsection, are placed within an <h2> element, a direct child of the <section>. Those of numbered paragraphs below the level of a Subsection are placed within “h” elements of increasing number: that is, within an <h3> at one level below a Subsection, within an <h4> two levels below, etc.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <num> | <span class="num"> |
| <heading> | <span class="heading"> |
| <subheading> | <span class="subheading"> |
| <authorialNote class="referenceNote">[[4]](#footnote-4) | <span class="authorialNote referenceNote"> |

Note that this HTML structure does not reflect the appearance of these structures. Sometimes a section number and its associated heading will appear on separate lines. And sometimes a number will appear on the same line as the first bit of content text, or even on the same line as the number of its first child level. These affects are achieved through CSS alone.

## Content / intro / wrapUp

When hierarchical structures in Akoma Ntoso do not have other hierarchical structures are children, they have a <content> element that contains their content. When they do have other hierarchical structures as children, the children may be preceded by or followed by content, in which case it is wrapped in an <intro> or <wrapUp> element. These <content>, <intro> and <wrapUp> elements do not influence the appearance of a document, but they are preserved in HTML for fidelity.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <content> | <div class="content"> |
| <intro> | <div class="intro"> |
| <wrapUp> | <div class="wrapUp"> |

## An example Chapter:

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <chapter>  <num>Chapter 1</num>  <heading>The Chapter Title</heading>  <section>  <num>1</num>  <heading>The first section</heading>  <subsection>  <num>(1)</num>  <level>  <num>(a)</num>  <content>  <p>…</p>  </content>  </level>  </subsection>  </section>  </chapter> | <section class="chapter">  <h2>  <span class="num">Chapter 1</span>  <span class="heading">The Chapter Title</span>  </h2>  <section class="section">  <h2>  <span class="num">1</span>  <span class="heading">The first section</span>  </h2>  <section class="subsection">  <h2>  <span class="num">(1)</span>  </h2>  <div class="level">  <h3>  <span class="num">(a)</span>  </h3>  <div class="content">  <p>…</p>  <div>  </section>  </section>  </section> |

# Lines of text

The main text of an Akoma Ntoso document occurs within <p> elements, which are mapped to the HTML <p> element.[[5]](#footnote-5)

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <p> | <p> |

# Inline elements

## Semantic inline

Akoma Ntoso contains many inline semantic elements, some of which have no effect on formatting. But all are retained for completeness.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <docTitle> | <span class="docTitle"> |
| <docNumber> | <span class="docNumber"> |
| <shortTitle> | <span class="shortTitle"> |
| <docDate date="1968-04-05"> | <time class="docDate" datetime="1968-04-05"> |
| <date date="1968-04-05"> | <time class="date" datetime="1968-04-05"> |
| <abbr> | <abbr> |
| <ref> | <cite><a>[[6]](#footnote-6) |
| <rref> | <cite><a> |
| <a> | <a> |
| <ins> | <ins> |
| <del> | <del> |
| <quotedText> | <q class="quotedText"> |
| <embeddedText> | <q class="embeddedText"> |
| <def> | <span class="def"> |
| <term> | <span class="term"> |
| <abbr> | <span class="abbr"> |

## Presentational inline

The common inline formatting elements in Akoma Ntoso were modeling on HTML elements and share the same name and meaning. There are a few additional generic <inline> elements with straightforward names.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <i> | <i> |
| <b> | <b> |
| <u> | <u> |
| <inline name="smallCaps"> | <span class="inline smallCaps"> |
| <inline name="uppercase"> | <span class="inline uppercase"> |
| <inline name="dropCap"> | <span class="inline dropCap"> |
| <sup> | <sup> |
| <sub> | <sub> |
| <span> | <span> |

# Lists

In UK Legislation, lists are represented with the <blockList> element, which can have complex structures within it. Nevertheless, it’s mapped to the simpler <ol> and <ul> elements in HTML5.

Numbered lists in Akoma Ntoso generally have the number specified for each item, within the <num> element. When these numbers are present, they are preserved in the HTML5, and the default numbering is overridden.

Akoma Ntoso lists occasionally have introduction and wrapUp components, like other AKN structures, and these need to be mapped to simple HTML <li> elements, with class attributes indicating their role.

There is a very rare “key list” structure in some older legislation, which gets mapped to HTML definition lists, using the <dl>, <dt> and <dd> elements.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <blockList> | <ol> or <ul> |
| <item> | <li> |
| <listIntroduction> | <li class="listIntroduction"> |
| <listWrapUp> | <li class="listWrapUp"> |
| <blockList class="key"> | <dl> |
| <heading> (child of key list) | <dt> |
| <p> or <blockList> (child of key list) | <dd> |

# Tables

The UK legislative model for Akoma Ntoso uses HTML table elements to represent tables[[7]](#footnote-7), which makes their conversion to HTML trivial. Tables in Legislation are often given a number or a heading, which are represented by native Akoma Ntoso elements, instead of HTML <caption> elements.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <tblock class="table"> | <div class="table"> |
| <num> | <div class="num"> |
| <heading> | <div class="heading"> |
| <foreign> | (nothing)[[8]](#footnote-8) |
| <html:table> | <table> |
| <html:thead> | <thead> |
| <html:tbody> | <tbody> |
| <html:tfoot> | <tfoot> |
| <html:tr> | <tr> |
| <html:th> | <th> |
| <html:td> | <td> |
| <authorialNote class="tablenote"> | <a class="fnRef"> |
| <tblock class="table">  <num>Table 1</num>  <heading>A Table Heading</heading>  <foreign>  <table xmlns="http://www.w3.org/1999/xhtml">  …  </table>  </foreign>  </tblock> | <div class="table">  <div class="num">Table 1</div>  <div class="heading">A Table Heading</div>  <table>  …  </table>  </div> |

# Quoted structures

Quoted structures are represented with the HTML <blockquote> element, which will be given a class attribute containing a space-separated list of values, representing the context for the foreign structure. The first value will be “primary,” “secondary”, “euretained” or “unknown”, corresponding to the class of legislation from which the structure is taken. The second value will be specific type of document from which the structure is taken, if known. The document type will be one of the short forms used on the root <article> element, such as “ukpga”. See above. Sometimes the class with contain a third value, “context-main” or “context-schedule”, indicating whether the structure appears in the main body or in a schedule.

Quoted structures appear within mixed content elements in Akoma Ntoso, usually a <p> element. Because an HTML <p> element may contain only inline content, text blocks that contain quoted structures are represented in HTML with <div class="p">.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <p> (containing a <mod>) | <div class="p"> |
| <mod> | <div class="mod"> |
| <quotedStrucutre> | <blockquote class="…"> |
| <embeddedStructure> | <blockquote class="…"> |
| <inline name="appendText">[[9]](#footnote-9) | <span class="inline appendText"> |
| <p>after Subsection 2 insert  <mod>  <quotedStructure uk:docName="ukpga">  <subsection>  <num>(2a)</num>  …  </subsection>  </quotedStructure>  </mod>  </p> | <div class="p">after Subsection 2 insert  <div class="mode">  <blockquote class="primary ukpga">  <section class="subsection">  <h2>  <span class="num">(2a)</span>  </h2>  …  <section>  </blockquote>  </div>  </div> |

# Images

Images in Legislation often occur alone within a block, sometimes preceded by a number or heading. These block images are presented in HTML with the <figure> element, and their heading by a <figcaption>.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <tblock class="figure"> | <figure> |
| <num> | <div class="num"> |
| <heading> | <div class="heading"> |
| <img> | <img> |
| <tblock class="figure">  <num>Figure 1</num>  <p><img src="image.jpg"/></p>  </tblock> | <figure>  <figcaption>  <div class="num">Figure 1</div>  </figcaption>  <p><img src="image.jpg"></p>  </figure> |

# Mathematical formulae

UK Legislation in Akoma Ntoso uses MathML to express mathematical formulae are represented using in MathML, embedded within Akoma Ntoso. MathML elements are valid , which makes their

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <tblock class="formula"> | <div class="tblock formula"> |
| <num> | <div class="num"> |
| <heading> | <div class="heading"> |
| <foreign> | (nothing) |
| <math:math>[[10]](#footnote-10) | <math> |
| <tblock class="formula">  <num>Formula 1</num>  <foreign>  <math:math>  …  </math:math>  </foreign>  </tblock> | <div class="tblock formula">  <div class="num">Formula 1</div>  <math>  …  </math>  </div> |

# Signatures

The signature block is an <hcontainer> that holds a series of <block> elements, each with a name attribute indicating its function: “seal”, “signature” (for a person’s name), “role”, “address”, “location” or “organization”. Within each, there is a semantic inline element indicating the same thing, often with the same name. That is to say that the Akoma Ntoso model contains

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <hcontainer name="signatures"> | <div class="hcontainer signatures"> |
| <hcontainer name="signatureBlock"> | <div class="hcontainer signatureBlock"> |
| <block name="seal"> | <div class="seal"> |
| <block name="signature"> | <div class="block signature"> |
| <signature> | <span class="signature"> |
| <block name="role"> | <div class="block role"> |
| <role> | <span class="role"> |
| <container class="address"> | <div class="container address"> |
| <block name="location"> | <div class="block location"> |
| <location> | <span class="location"> |
| <block name="date"> | <div class="block date"> |
| <date date="2021-05-25"> | <time datetime="2021-05-25"> |
| <block name="organization"> | <div class="block organization"> |
| <organization> | <span class="organization"> |
| <hcontainer name="signatureBlock">  <content>  <p>Signed by…</p>  <block name="signature"><signature>Jim Mangiafico</signature></block>  <block name=" role"><role>Secretary of…</role></block>  <block name="date"><date date="2021-05-25">25th May 2021</date></block>  <block name="organization"><organization>Ministry of…</organization></block>  </content>  </hcontainer> | <div class="hcontainer signatureBlock">  <div class="content">  <p>Signed by…</p>  <div class="block signature"><inline name="signature">Jim Mangiafico</inline></div>  <div class="block role"><inline name="role">Secretary of…</inline></div>  <div class="block date"><time class="date" datetime="2021-05-25">25th May 2021</time></div>  <div class="block organization"><inline class="organization">Ministry of…</inline></div>  </div>  </div> |

# Explanatory notes

The Explanatory Note in a piece of secondary legislation usually contains simple text block or numbered paragraphs. Because of the requirements of the Akoma Ntoso schema, numbered paragraphs in an Explanatory Note are represented with a <blockContainer> instead of an <hcontainer> that would appear in the body, but otherwise the structures are very similar.

|  |  |
| --- | --- |
| Akoma Ntoso | HTML5 |
| <conclusions> | <div class="conclusions"> |
| <blockContainer name="explanatoryNote"> | <div class=" blockContainer explanatoryNote"> |
| <blockContainer name="explanatoryNote">  <p>This is a line of text.</p>  <blockContainer class="para1">  <num>(a)</num>  <p>This is a numbered paragraph.</p>  </blockContainer>  </blockContainer> | <div class=" blockContainer explanatoryNote">  <p>This is a line of text.</p>  <div class="para1">  <span class="num">(a)</span>  <p>This is a numbered paragraph.</p>  </div>  </div> |

# Footnotes

Footnote markers appear in the text as links. The footnotes appear either at the end of the document or immediately following a table, in the case of some footnotes referenced within the table. A footnote’s number or “marker” is represented with a link to the footnote reference.

|  |  |
| --- | --- |
|  | HTML5 |
| footnote marker in text | <a class="fnRef" id="ref-x" href="#x"> |
| footnote at end of document | <div class="authorialNote footnote" id="x">  <a class="marker" href="#ref-x">1</a>  <p>…</p>  </div> |
| “table note” beneath a table | <div class="authorialNote tablenote" id="y">  <a class="marker" href="#ref-y">\*</a>  <p>…</p>  </div> |

1. Such as “ukpga” or “uksi”. [↑](#footnote-ref-1)
2. for EU Titles [↑](#footnote-ref-2)
3. The <paragraph> element is used to represent both the largest subdivision of Articles, Rules and Regulations in secondary legislation and the main provisions of schedules. [↑](#footnote-ref-3)
4. inside a <num> or <heading> [↑](#footnote-ref-4)
5. Except when they contain quoted structures. See below. [↑](#footnote-ref-5)
6. Akoma Ntoso references are represented by an HTML <cite> element containing an <a> element. [↑](#footnote-ref-6)
7. Akoma Ntoso allows for the inclusion of markup from other namespaces, provided it is wrapped in the Akoma Ntoso <foreign> element. [↑](#footnote-ref-7)
8. The Akoma Ntoso <foreign> element signifies that the markup is moving from the Akoma Ntoso namespace into another namespace. It represents nothing structural about the document. In UK legislation, the only other “foreign” namespaces are HTML and MathML. [↑](#footnote-ref-8)
9. <inline name="appendText"> elements represent punctuation, typically a period, which appears after the closing quotation marks of a quoted structure. In Akoma Ntoso, <inline name="appendText"> elements appear immediately *following* a <quotedStructure> or <embeddedStruture>, whereas in HTML5 they appear *within* the <blockquote>, immediately following the last non-empty text node, which may be many layers deep. [↑](#footnote-ref-9)
10. The MathML namespace URI is http://www.w3.org/1998/Math/MathML [↑](#footnote-ref-10)