Citation Elements:
Bindings for RDFa

11 September 2017

FHISO's suite of **Citation Elements** standard provides an extensible framework and vocabulary for encoding all the data about a genealogical *source* that might reasonably be included in a *formatted citation* to that *source*.

This information is represented as a sequence of *citation elements*, logically self-contained pieces of information about a *source*. This document defines a means by which *citation elements* may be identified and tagged within an XML or HTML *formatted citation*, allowing a computer to extracted them in a systematic manner. The tagging of *citation elements* is done using a standard set of HTML attributes known as RDFa attributes, which can also be used in XML languages besides HTML.

Other documents in the suite of Citation Elements standards are as follows:

— **Citation Elements: General Concepts**. This standard defines the general concepts used in FHISO's suite of Citation Elements standards, and the basic framework and data model underpinning them.

— **Citation Elements: Vocabulary**. This standard defines a collection of *citation elements* allowing the representation of information normally found in *formatted citations* to diverse types of source.

— **Citation Elements: Bindings for GEDCOM X**. This standard defines extensions to the GEDCOM X data model and its JSON and XML serialisations to allow *citation elements* to be represented in GEDCOM X.

— **Citation Elements: Bindings for ELF**. This standard defines how *citation elements* should be represented in FHISO's Extensible Legacy Format (ELF), a format based on and compatible with GEDCOM 5.5.1, but with the addition of a new extensibility mechanism.

> *Editorial note* — Not all of these documents are yet at the stage of having a first public draft.

1 Introduction

1.1 Conventions used

Where this standard gives a specific technical meaning to a word or phrase, that word or phrase is formatted in bold text in its initial definition, and in italics when used elsewhere. The key words MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, NOT RECOMMENDED, MAY and OPTIONAL in this standard are to be interpreted as described in [RFC 2119].

An application is **conformant** with this standard if and only if it follows all the requirements and prohibitions contained in this document, as indicated by use of the words MUST, MUST NOT, REQUIRED, SHALL and SHALL NOT, and the relevant parts of its normative references. Standards referencing this standard MUST NOT loosen any of the requirements and prohibitions made by this standard, nor place additional requirements or prohibitions on the constructs defined herein.

> *Note* — Adding requirements or prohibitions is disallowed so as to preserve interoperability between applications: data generated by one *conformant* application must always be acceptable to another *conformant* application, regardless of what additional standards each may conform to.

This standard depends on the *Citation Elements: General Concepts* standard [CEV Concepts]. To be *conformant* with this standard, an application MUST also be conformant with [CEV Concepts]. Some words and phrases defined in that standard are used here without further definition.

> *Note* — Readers are advised to read at least the introduction to [CEV Concepts] before reading this standard.

Indented text in coloured boxes, such as preceding paragraph, does not form a normative part of this standard, and is labelled as either an example or a note.

> *Editorial note* — Editorial notes, such as this, are used to record outstanding issues, or points where there is not yet consensus; they will be resolved and removed for the final standard. Examples and notes will be retained in the standard.

1.2   RDFa attributes

The tagging of *citation elements* in *formatted citations* is done using a standard set of HTML attributes known as RDFa attributes which are defined in [RDFa Core]. Compliance with this FHISO standard does not require full RDFa compliance: support for the full [RDFa Core] is OPTIONAL, and RDFa features other than those for which support is REQUIRED by this standard SHOULD NOT be used when compatibility between implementations is desirable.

> *Note —* The specification of [RDFa Core] assumes a detailed working knowledge of the RDF graph model. A more accessible introduction to RDFa can be found in the [RDFa Primer], but FHISO's use of RDFa attributes here is limited, and this standard is designed to be used without any knowledge of RDFa or RDF. An application parsing RDFa attributes according to this specification does not need a full RDFa parser, far less to support the full RDF graph model.

These attributes MAY be used in HTML or any XML-based markup language, but for the purpose of tagging *citation elements* in *formatted citations* it is RECOMMENDED that they be used in XHTML. The language they are used in is referred to here as the **host language**.

> *Note —* Applications wishing to implement a fully-compliant RDFa parser for HTML will find the formal specification on the use of RDFa in HTML in two standards, [HTML+RDFa] and [XHTML+RDFa].

In the simplest case, the *citation element name* (which is an IRI) can be put in a `property` attribute on an XML or HTML element, and the *citation element value* is the text contents of the element. The particular type of element on which the attributes are placed is not relevant.

> *Example —* A simplified *formatted citation* to Settipani's book *Les ancêtres de Charlemagne* might be marked up as the following HTML fragment:
>
> ```
> <p>Settipani, Christian.  <i>Les ancêtres de Charlemagne</i>.</p>
> ```
>
> The title of the book can be tagged by adding a `property` attribute to the existing `<i>` element. As written above, no element contains just the author's name, as the `<p>` element also encloses the title; however author's name can be wrapped in an `<span>` element and a `property` attributed added to that. HTML's `<span>` element has no defined meaning of its own, but exists to provide a place for attributes such as this.
>
> ```
> <p><span property="https://terms.fhiso.org/sources/authorName"
>   >Settipani, Christian</span>.
>   <i property="https://terms.fhiso.org/sources/title">Les ancêtres
>     de Charlemagne</i>.</p>
> ```
>
> An HTML renderer will correctly format this while ignoring the two `property` attributes, but an application that conforms to this standard will extract these two *citation elements* from this HTML:

```
authorName:  "Settipani, Christian"
title:       "Les ancêtres de Charlemagne"
```

Note the *citation element value* of `title` *citation element* contains no line break, despite the HTML being split across two lines. This is because [CEV Concepts] says applications SHOULD *whitespace-normalise citation element values*.

Note — In many examples in this standard, including the previous one, the list of *citation elements* is given as a list of name, value pairs with both presented as a *string*. In practice the *citation element value* is a *localisation set* containing one *string* which is additionally tagged with a *datatype* and possibly a *language tag*. This detail is frequently omitted from examples where it is not germane to the point being illustrated.

1.2.1 Index of attributes used

This standard makes use of the following attributes:

— The `vocab` and `prefix` attributes are used to allow the creation of shorthand IRIs per §2. Full support for their RDFa semantics is REQUIRED by this standard, except that the use of an *initial context* to provide defaults is OPTIONAL.

— The `typeof` attribute is used to locate *formatted citations* per §3.1. Support for any other use of this attribute is OPTIONAL; any unsupported use of it SHALL be marked as a *source-exclusion element* per §3.2 and is not further processed by this standard.

— The `property` attribute contains a *citation element name* as per §4. Full support for its RDFa semantics is REQUIRED, other than when it is used in constructs that define *source-exclusion elements*, and except for the special behaviour RDFa gives to an `rdfa:copy` property for which support is OPTIONAL.

— The `content` attribute can be used to represent a *citation element value* as per §4.2. Full support for its RDFa semantics is REQUIRED.

— The `href` and `src` attribute can be used to represent a *citation element value* as per §4.2. They are not formally considered RDFa attributes but are part of the *host language*. Full support for their RDFa semantics is REQUIRED if the *host language* permit their use, as HTML does.

— The `datetime` attribute can also be used to represent a *citation element value* as per §4.2 if the *host language* is HTML.

— The `xml:lang` and `lang` attributes are used to represent a *language tag* as per §4.4. Full support for their RDFa semantics is REQUIRED.

— The `datatype` attribute is used to identify the *datatype* of *strings* in a *citation element value*. Full support for its RDFa semantics is REQUIRED.

— The `rel` and `rev` attributes are used to denote *layer derivation links* per §5.3. Support for any other use of this attribute is OPTIONAL; any unsupported use of them SHALL be marked as a *source-exclusion element* per §3.2 and is not further processed by this standard.

— The `about`, `inlist` and `resource` attributes are not used by this standard. Support for their RDFa semantics is OPTIONAL. Any unsupported use of them SHALL be marked as a *source-exclusion element* per §3.2 and is not be processed by this standard, except when the presence of one of these attributes (but not its particular value) prevents the recognition of *nested source-type element* per §5.1.

In addition, when the *host language* is HTML, special meaning is attached to the `<time>` element.

1.3 Motivation and limitations

In this standard, unless otherwise stated, the term HTML refers to any backwards-compatible version of HTML, and XHTML refers to any version of HTML that is also well-formed XML.

> *Note —* This definition of HTML includes HTML 4.01, XHTML 1.0, XHTML 1.1, HTML5 and HTML 5.1. For the last two, it includes both their XML and non-XML forms. It will include future editions of HTML5 too, assuming they retain backwards compatibility. This definition of XHTML includes not just the standards that are named XHTML, but also the XML forms of HTML5 and later.

The use of HTML, or a subset of HTML, is often permitted in genealogy applications to allow users to add formatting to text in various contexts. It is RECOMMENDED that applications which allow users to edit or manually lay out *formatted citations* SHOULD permit the use of some HTML elements in them.

> *Note —* [CEV Concepts] recommends that if high quality *formatted citations* are required, users SHOULD be allowed to fine-tune the presentation by hand because it is not anticipated that an application will always do a perfect job. Many citation styles use italics and some use bold, underlining or other text-level formatting when formatting certain *citation elements*. In order to allow the user to fine-tune the use of such formatting, the user SHOULD be allowed to edit the *formatted citation* as HTML.

If an application automatically generates an HTML *formatted citation* from a *citation element set*, it SHOULD add RDFa attributes in such a manner that will another application *conformant* with this standard will be able to extract the *citation elements* again. This SHOULD NOT be an application's principal means of serialising a *citation element set*: applications SHOULD prefer a format that serialises the *citation element set* directly rather than after converting it to a *formatted citation*.

> *Note —* RDFa attributes are not the recommended way of serialising *citation element sets* primarily because it requires creating a *formatted citation*. Doing this to a reasonable standard is non-trivial, and results in particular language and style being favoured. This standard is provided for situations when a *formatted citation* is desired or required anyway. For example, an enormous amount of genealogical research has been published online and includes *formatted citations*. If they are tagged according to this standard, these *formatted citations* can be copied and pasted into a genealogy application which can convert them back to a *citation element set*.

2 Shorthand IRIs

The [CEV Concepts] standard makes heavy use of IRIs as identifiers, as does RDFa. In particular, the `datatype`, `property`, `rel`, `rev` and `typeof` attributes contain IRIs.

The `datatype` attribute SHALL contain a single IRI. The `property`, `rel`, `rev` and `typeof` attributes SHALL contain a list of IRIs separated by *whitespace.* Leading and trailing *whitespace* is discarded.

> *Example* — A common reason why multiple IRIs might be present is when two IRIs exist with similar meanings and the creator of the citation wishes to use both for compatibility.
>
> ```
>     <i property="https://terms.fhiso.org/sources/title
>                   http://purl.org/dc/terms/title">Les ancêtres de
>       Charlemagne</i>
> ```
>
> Here two alternative IRIs are used to tag the title, presumably because the citation's creator anticipated it being processed by applications that support [Dublin Core] metadata as well as FHISO's Citation Elements standards. A parser conforming to this standard will treat both IRIs as valid and create two *citation elements*, both with the same *citation element value*, however if the Dublin Core IRI is not known to the application, it will likely be ignored.

In the uses described by this standard the `property` attribute will always contain a *citation element term,* and the `datatype` attribute will always contain a *datatype name.* The `typeof` attribute will contain an IRI that allows this standard's use of RDFa to be distinguished from any other uses also present in the document. The `rev` and `rel` attributes will contain a *source derivation type* to denote *citation layer links.*

RDFa provides two separate mechanisms for abbreviating the IRIs in these attributes: by setting a *local default vocabulary*, and by using prefixes to create compact URIs expressions (CURIEs) as a form of *prefix notation.* Applications processing *formatted citations* in accordance with this standard MUST support both of these mechanisms. Expansion of *terms* using the *local default vocabulary* SHALL be done before the expansion of CURIEs. An application MUST behave as if all `datatype`, `property`, `rel`, `rev` and `typeof` attributes have been expanded before continuing to process the data.

> *Note* — Applications MAY opt to expand these attributes on demand, provided the effect is the same. The `typeof` attribute is the only one whose value invariably needs expanding.

2.1 Default vocabularies

A **term** in RDFa is an XML NCName that also permits slash (U+002F) as a non-leading character. It matches the `term` production given in §7.4.3 of [RDFa Core].

> *Note* — This production is as follows:
>
> ```
> term      ::=  NCNameStartChar termChar*
> termChar ::=  ( NameChar - ':' ) | '/'
> ```
>
> The definitions of `NameChar` and `NCNameStartChar` are found in [XML] and [XML Names] respectively.

> *Note* — The [CEV Concepts] standard also uses the word "term", and defines it to mean a vocabulary item identified by an IRI. To minimise confusion, this standard never uses the word "term" in that sense, and only uses it in the RDFa sense given above.

When a `datatype`, `property`, `rel`, `rev` or `typeof` attribute contains a *term*, it SHALL be converted to an IRI by prepending the *local default vocabulary* if one exists. The **local default vocabulary** is an IRI which is specified using a `vocab` attribute. It applies to the element where it is specified and to all elements in its content unless overridden with another `vocab` attribute.

> *Note* — Terms look similar to relative IRIs and this process is similar to resolving relative IRIs against a base IRI, but the process of applying a *local default vocabulary* is simpler as the two strings are simply concatenated without understanding the structure of the IRI.

Markup generators SHOULD ensure that a `vocab` attribute is present if *terms* are being used when compatibility between implementations is desirable. When these attributes are used in a *host language* other than HTML, the definition of the *host language* MAY provide a *default vocabulary* that applies in the event that no `vocab` attribute is found; HTML provides no such default.

If no *local default vocabulary* was found, a parser MAY use an *initial context* as described in §9 of [RDFa Core] to resolve the *term* to an IRI; if not, or if it was not found in the *initial context*, the term SHALL be ignored. When an *initial context* is used, it MUST be the standard one for the *host language*: implementations MUST NOT define their own *initial context*.

> *Example* —
>
> ```
> <p><span property="authorName">Settipani, Christian</span>.
>   <i vocab="https://terms.fhiso.org/sources/"
>      property="title">Les ancêtres de Charlemagne</i>.</p>
> ```
>
> In this fragment, both `property` attributes contain a *term*. The `title` *term* is converted to the IRI of FHISO's `title` *citation element*:
>
> ```
> https://terms.fhiso.org/sources/title
> ```
>
> In considering the `authorName` *term*, a parser looks for a `vocab` attribute on the `<span>` or the enclosing `<p>` element. No such attribute exists, and the RDFa attributes are being used in HTML which provides no *default vocabulary*.

> The parser MAY consider the standard *initial context* too, and if it is a full RDFa parser it MUST. As the host language is HTML, the *initial context* is defined in [HTML5+RDFa Context]. At the present time this only includes mappings for describedBy, license and role. These are to be matched case-sensitively, or failing that case-insensitively, but the title *term* used in this example clearly does not match.
>
> Regardless of whether the application considered the *initial context*, the title *term* cannot be resolved to an IRI and is therefore ignored.

> *Editorial note —* If use of the *initial context* is changed to be REQUIRED for CURIEs, below, it should be changed here too.

2.2   Compact URI Expressions (CURIEs)

A **CURIE** comprises two components, a *prefix* and a *reference*, separated by a colon (U+003A). It matches the curie production given in §6 of [RDFa Core].

> *Note —* This production is defined as follows:
>
> ```
> curie       ::=   ( prefix? ':' )? reference
> prefix      ::=   NCName
> reference   ::=   ( ipath-absolute | ipath-rootless | ipath-empty )
>                         ( '?' iquery )?  ( '#' ifragment )?
> ```
>
> The definitions of NCName is found in [XML Names]. The various productions referenced in the definition of reference are defined in [RFC 3987]. None of these ipath productions match a string beginning "//", therefore IRIs of the form http://... never match the curie syntax production. There is a conflict with certain other, less-used IRI schemes, and mailto:user@example.com does match the syntax. However this only results in this IRI being treated as a CURIE if mailto is defined as a CURIE prefix. The RDFa working group considered the risk of this to be minimal.
>
> Although this syntax definition allows the omission of both prefix and the colon, in practice there is no situation in RDFa where both can be omitted and the result still parsed as a CURIE. A parser conforming to this standard MAY safely treat the colon as mandatory.

When a datatype, property, rel, rev or typeof attribute contains a *whitespace* separated token that is syntactically a CURIE, the parser should look up its *prefix* to see whether a *prefix mapping* (which is an IRI) has been defined. This look-up is done case-insensitively.

If the *prefix* has been omitted and the CURIE begins with a colon, parsers MAY ignore the CURIE and MUST NOT fall back to treating it as an IRI; if is is not ignored, the *prefix mapping* MUST be

    http://www.w3.org/1999/xhtml/vocab#

> *Note* — This vocabulary contains little of use in marking up *formatted citations*.

When the *prefix* is present, a parser MUST try to look it up in the *local prefix mappings*. These are set using `prefix` attributes. This attribute MUST contain an even number of whitespace separated tokens: the first and every subsequent odd token must be an `NCName` followed by a colon; the second and every subsequent even token must be an IRI. The `NCName` is the *prefix* and the IRI is its *prefix mapping*. The mapping applies to the element where it is specified and to all elements in its content unless overridden.

> *Example* — The following is an example of a well-formed `prefix` attribute.
>
> ```
> <div prefix="cev: https://terms.fhiso.org/sources/
>               dc:  http://purl.org/dc/elements/1.1/">
>   <i prefix="dc:  http://purl.org/dc/terms/"
>      property="cev:title dc:title">Les ancêtres de Charlemagne</i>
> </div>
> ```
>
> The `prefix` attribute on the `<div>` defines two *local prefix mappings*, one for the `cev` *prefix*, the other for the dc *prefix*. The dc *local prefix mapping* is overridden by the `prefix` attribute on the `<i>` element; the `cev` *local prefix mapping* has not been overridden and remains in operation.

The *prefix* consisting of a single underscore character (U+005F) has special meaning in §7.4.5 [RDFa Core] for referencing *blank nodes*. It MUST NOT be used in CURIEs other than for that purpose. Support for *blank nodes* is NOT RECOMMENDED in this standard. Applications that do not support *blank nodes* MUST ignore CURIEs with a *prefix* consisting of a single underscore.

In determining the *local prefix mappings*, a parser MAY also use XML namespace declarations as defined in §7.5, item 3 of [RDFa Core]. This is not required even in full RDFa parsers and is deprecated; it is NOT RECOMMENDED by this standard.

If the *prefix* was not found in the *local prefix mappings*, a parser MAY use an *initial context* as described in §9 of [RDFa Core] to determine the *prefix mapping*. When an *initial context* is used, it MUST be the standard one for the language on which the RDFa tags are used: implementations MUST NOT define their own *initial context*.

> *Editorial note* — It may be worth making this REQUIRED rather than OPTIONAL as the *initial context* for HTML contains *prefix mappings* for several potentially useful vocabularies including Dublin Core and PROV. It is unlikely to add much complexity to the parser or this specification.

If a *prefix mapping* is found, the CURIE is converted to an IRI by prepending the *prefix mapping* to the *reference* part of the CURIE.

> *Example —* The two CURIEs in the previous example expand to these IRIs:
>
>   https://terms.fhiso.org/sources/title
>   http://purl.org/dc/terms/title

If no *prefix mapping* is found, the CURIE SHALL be treated as an IRI if it is syntactically valid as one or ignored otherwise. If this results in an IRI with an unknown scheme, the parser MAY ignore it; parsers MUST NOT ignore the `http`, `https` or `urn` schemes.

> *Note —* Virtually all CURIEs are syntactically valid IRIs since `prefix:reference` is a valid IRI, despite having an unknown scheme. The option of ignoring unknown IRI with unknown schemes is introduced because this standard makes the use of an *initial context* OPTIONAL. CURIEs with prefixes that would be resolved via the *initial context* in a full RDFa parser may therefore be left unresolved by a parser conforming to this standard. Almost invariably they will have an unknown scheme when reinterpreted as an IRI and can therefore be dropped. Full RDFa parsers MUST use *initial contexts* and therefore MUST NOT ignore IRIs with unknown schemes.

> *Editorial note —* If support for *initial contexts* becomes REQUIRED, the ability to ignore unknown schemes should probably be dropped.

3 Locating citation elements

In general a document will contain more than just a single *citation element set*, and other parts of the document may also contain RDFa attributes for entirely different purposes; even if the only use of RDFa is for tagging *citation elements* it is important not to confuse the *citation elements* from one *formatted citation* or *citation layer* with those of another.

Citation elements are identified using `property` attributes. However a `property` attribute SHALL only be interpreted as representing a *citation element* if:

— it is on an element contained inside a *source-type element* (but is not directly on the *source-type element* itself); and

— it is not located on a *source-exclusion element* within the *source-type element*, nor is it on an element contained within a *source-exclusion element*.

Any `property` attributes matching the above criteria SHALL be considered in the order they appear in the document and used to generate *citation elements* as described in §4.

> *Note —* The detailed specification in §7.5 of [RDFa Core] requires that `property` attributes are processed and used to generate RDF triples in document order. However the [RDFa Core] processing model requires these triples be added to an RDF graph, and RDFa graphs are not required to preserve the order of triples; nevertheless, most current RDFa proces-

> sors do output properties in document order. Implementations using an RDFa parser to implement this specification should verify that the document order of properties can be determined.

The *citation elements* contained within a *source-type element* SHALL form a *citation element set* which represents a *citation layer* (or a *single-layered citation*) as described in §5.

3.1 Source-type elements

A **source-type element** is any element that has a `typeof` attribute whose value, once shorthand IRIs have been expanded, includes either of the following IRIs:

```
https://terms.fhiso.org/sources/Source
https://terms.fhiso.org/sources/CitedSource
```

HTML or XML content is only considered to be part of a *formatted citation* if it is a *source-type element* or is contained within one.

> *Example —* The following example contains two entirely unrelated uses of RDFa attributes:
>
> ```
> <p vocab="https://terms.fhiso.org/sources/" typeof="Source">
>   <span property="authorName">Settipani</span>, <i>Ibid.</i></p>
> <div vocab="http://creativecommons.org/ns#">Released under a
>   <a href="http://creativecommons.org/licenses/by/3.0/"
>       property="license">Creative Commons License</a>.</div>
> ```
>
> The `typeof` attribute of the `<p>` element has a value that expands to the required IRI. This marks the `<p>` element as a *source-type element*, and its contents as a *formatted citation*. This contains just one `property` attribute, so a parser will find just one *citation element*: an `authorName` one with value "Settipani".
>
> The `license` property is not contained in a *source-type element* and therefore does not denote a *citation element*. It is a use of RDFa that is outside the scope of this standard. This is as well: Settipani's book is not licensed under a Creative Commons License, though a page discussing it may well be.

An external mechanism may be used to designate the entirety of an HTML document or fragment a *source-type element*.

> *Example —* A non-HTML syntax might embed fragments of HTML to represent individual *formatted citations*. It would likely designate each fragment to be a *source-type element*, in which case the `typeof` attribute is OPTIONAL.

Editorial note — There has been some discussion about the possibility of using the `resource` attribute on *source-type elements* to generate certain "meta" *citation elements* such as a UUID or a "citation authority IRI".

3.2 Source-exclusion elements

Note — The concept of a *source-exclusion element* is necessary to prevent a parser from misinterpreting `property` attributes that are part of more complex RDFa constructs which this standard does not require to be supported. Future FHISO standards may make use of some of these RDFa constructs and this restriction also allows for forwards compatibility.

An application that supports only those RDFa features for which support is REQUIRED by this standard MUST consider an element to be a **source-exclusion element** of a given *source-type element* if it is contained within the *source-type element* (but is not the *source-type element* itself) and has an attribute named `about`, `inlist`, `rel`, `resource`, `rev`, or `typeof`.

Editorial note — The circumstances in which the *source-type element* is itself excluded needs further consideration giving particular attention to the processing sequence in §7.5 of [RDFa Core].

Example — The following example includes a more complex use of RDFa attributes, beyond what this standard requires to be understood.

```
<p prefix="foaf: http://xmlns.com/foaf/0.1/"
   vocab="https://terms.fhiso.org/sources/" typeof="CitedSource">
  <span rel="foaf:maker">
    <span property="foaf:name">Settipani</span></span>,
  <i property="title">Les ancêtres de Charlemagne</i>.
</p>
```

The <p> element is a *source-type element* due to the `typeof="CitedSource"` attribute, and the *formatted citation* is the string "Settipani, *Les ancêtres de Charlemagne*."

The <p> element has one *source-exclusion element*: the outer <span> element due to its `rel` attribute. Parsers are not expected to understand the meaning of this `rel` attribute, just to note its presence. As the inner <span> element is contained within this *source-exclusion element*, the `property="foaf:name"` attribute *must not* be treated as a *citation element*.

The `property` attribute on the <i> element is not located within a *source-exclusion element*, and therefore it does denote a *citation element*. This is the only *citation element* in this example.

> *Note* — These rules allow *source-type elements* to nest, with the inner *source-type element* being a *source-exclusion element* of the outer *source-type element*. This behaviour is used in the representation of *layered citations*, as discussed in §5.

Applications which support a larger part of RDFa than this standard requires MAY treat fewer elements as *source-exclusion elements*. If so, they MUST ensure that RDFa constructs are only treated as *citation elements* when they produce an RDF triples whose subject has the following RDF types, or a subtype thereof:

    https://terms.fhiso.org/sources/Source

In addition, applications supporting a larger part of RDFa MUST discard triples where the object is an RDF blank node.

> *Note* — A future FHISO standard might extend this data model to include support for blank nodes, likely using them to represent objects with properties of their own.

> *Note* — This standard is designed to allow implementers to parse those RDFa constructs used without having to consider how they map to RDF. The preceding text is only of relevant if an implementer wishes to make greater use of the RDF features underlying RDFa.

4   Parsing citation elements

As defined in the [CEV Concepts] standard, a *citation element* consists of two components:

— a *citation element name*, which SHALL be an IRI defined as a *citation element term*; and
— a *citation element value*, which SHALL be a *localisation set*.

Once a parser has identified the `property` attributes that are tagging *citation element* it SHALL determine each component of each *citation element* as described in the following sub-sections.

For the purpose of this section, the **current element** refers to the XML or HTML element that has the `property` attribute which tags the current *citation element*.

The *citation element name* SHALL be the value of the `property` attribute, once shorthand IRIs have been expanded. If the `property` attribute contains more than one IRI, each SHALL be used as the *citation element name* of a separate *citation element* with a copy of the same *citation element value*.

To construct the *citation element value*, an application SHALL determine its *current property value*, as defined in §4.2 below. This is a *string* and is used to construct a new *localisation set* to be the *citation element value*. The application SHALL then determine the *datatype* of the *string* per §4.3, and if the result is a *language-tagged datatype*, SHALL also determine its *language tag* per §4.4. Alternatively, applications that opt to parse RDFa to RDF triples, as a full RDFa parser does, MAY determine the *current property value*, *datatype* and *language tag* per §4.5.

> *Note —* These rules are illustrated by example in the sections below.

4.1 List flattening

RDFa, as used in this standard, is a *list-flattening format*. This means it does not naturally provide a means of keeping the *localisation sets* of each *citation element* separate because it has no means of distinguishing *multi-valued citation elements* from translated or localised versions of the same *citation element*. Applications must therefore assume every `property` attribute identifies a separate *citation element*.

> *Editorial note —* It would have been possible for this standard to have defined a usage of RDFa that was not a *list-flattening format*. This was not done because it would make most straightforward uses unidiomatic, and likely compromise the uptake of this standard.

> *Example —* The following RDFa markup is well-formed but will be misinterpreted by a parser conforming to this specification.
>
> ```
> <p lang="en-GB" typeof="Source">
>   <span property="authorName"
>         content="Lansdowne, Marquess of">Lord Lansdowne</span> and
>   <span property="authorName" lang="jp-Latn">Hayashi Tadasu</span>
>   (<span property="authorName" lang="jp">林 董</span>),
>   <i property="title">The Anglo-Japanese Treaty</i>,
>   <span property="publicationDate">1902</span>.
> </p>
> ```
>
> The Anglo-Japanese Treaty was (at least nominally) authored by two people: the Marquess of Lansdowne and Count Hayashi Tadasu whose name is written in kanji as 林 董. A *conformant* application will see three `authorName`s and make each into a separate *citation element*, when in fact the desired behaviour is for "林 董" to be part of the same *localisation set* as "Hayashi Tadasu".

Applications are REQUIRED to use the `localisedElement` mechanism defined in §3.4.1 of [CEV Concepts] when multiple translations or localisations of a single *citation element value* are needed.

> *Example —* The RDFa markup from the previous example can be fixed by using a `localisedElement` to encode the second for of Hayashi's name. At its simplest, this alters the two <span> elements referring to Hayashi to read:
>
> ```
> <span property="authorName" lang="jp-Latn">Hayashi Tadasu</span>
> (<span property="localisedElement" lang="jp">林 董</span>)
> ```
>
> However, [CEV Concepts] recommends that the first *string* in the *localisation set* SHOULD be the untranslated, and ideally untransliterated form of the *citation element*. Undoubtedly it

is the Latin form that is the transliteration, and therefore these elements are the wrong way round. While this is only a recommendation, applications SHOULD try to follow it; this can be achieved as follows:

```
<span property="authorName" lang="jp" content="林 董" />
<span property="localisedElement"
      lang="jp-Latn">Hayashi Tadasu</span> (林 董)
```

This use of the `content` attribute is discussed below. It provides a value for the *citation element* while hiding the value from an HTML renderer.

4.2 Current property value

Editorial note — This section, together with the following section defining the *datatype*, derive from step 11 in the processing sequence given in §7.5 of [RDFa Core], as amended by §3.1 of [HTML+RDFa].

The **current property value** is a *string* which will be used to create the *citation element value*. It is determined based on the RDFa attributes present on the *current element* as follows.

Note — The use of the term *current property value* in this standard coincides with its definition in [RDFa Core].

If *current element* has a `content` attribute, and either has no `datatype` attribute, or its `datatype` attribute is empty or has a value (after expanding shorthand IRIs) other than either of the following IRIs, then the *current property value* SHALL be the value of the `content` attribute.

```
http://www.w3.org/1999/02/22-rdf-syntax-ns#XMLLiteral
http://www.w3.org/1999/02/22-rdf-syntax-ns#HTML
```

Note — These two IRIs have special treatment in RDFa. This standard excludes them for complete compatibility with a full RDFa parser, but it is not anticipated that they will arise in practice.

Example — The purpose of the `content` attribute is to allow the *citation element value* to be something that is not rendered or otherwise used in HTML. This is particularly important when the *citation element* is required to have a value in a format that is different to how the element is formatted.

```
<span property="https://terms.fhiso.org/sources/publicationDate"
      content="2017-05-22">May 22nd, 2017</span>
```

In this case, the use of a `content` attribute is necessary because the `publicationDate` *citation element value* MUST be a date in the prescribed date format based on [ISO 8601]: it MUST NOT be a date like "May 22nd, 2017".

> *Editorial note —* It would be desirable to add support for the `resource` attribute here. Before adding it, it is necessary to establish how safe it is to remove `resource` from the list of attributes that make a *source-exclusion element*.

Otherwise, if the *host language* is HTML, if the *current element* has a `datetime` attribute, the *current property value* SHALL be the value of the `datetime` attribute.

Otherwise, in HTML or in other XML languages that support an `href` attribute, if the *current element* has an `href` attribute and no `datatype` attribute, the *current property value* SHALL be the value of the `href` attribute, which SHALL be an IRI.

Otherwise, in HTML or in other XML languages that support a `src` attribute, if the *current element* has a `src` attribute and no `datatype` attribute, the *current property value* SHALL be the value of the `src` attribute, which SHALL be an IRI.

> *Note —* The [HTML+RDFa] standard does not change which HTML elements can have a `datetime`, `href` or `src` attribute. At present, the `datetime` attribute is only permitted on a `<time>` element; most `href` attributes in HTML are found on `<a>` elements; most `src` attributes are on elements that display some form of media, particularly `<img>` and in HTML5, `<video>` and `<audio>`.

> *Example —* When an `href` or `src` attribute links to an online *source*, it can be tagged as a *citation element*.
>
> ```
> <div vocab="https://terms.fhiso.org/sources/" typeof="Source">
>   <a href="http://discovery.nationalarchives.gov.uk/"
>       property="accessURL"><span property="title">Discovery</span></a>
>   (online catalogue)
> </div>
> ```
>
> This example has two *citation elements*:
>
> | | |
> |---|---|
> | accessURL: | http://discovery.nationalarchives.gov.uk/ |
> | title: | "Discovery" |
>
> The fact that the second `property` attribute is on a child element of the element containing the first `property` attribute is irrelevant and does not signify any additional connection between the `title` and the `accessURL` over and above their usual relationship.

Otherwise, the *current property value* SHALL be formed by concatenating the text contained in each of the descendant text nodes of the *current element* in document order.

> *Example —* This definition allows citation elements to nest which can be useful when tagging full titles and short versions of them.

```
<p vocab="https://terms.fhiso.org/sources/" typeof="Source">
  <i property="title"><span property="shortTitle">The visitations
  of Kent</span>, taken in the years 1530–1 by Thomas Benolte,
  Clarenceux, and 1574 by Robert Cooke, Clarenceux.</i>
</p>
```

The `shortTitle` property takes the value "The visitations of Kent", while the `title` property takes the value "The visitations of Kent, taken in the years ..." by concatenating the text in the nested `<span>` element with the text directly in the `<i>` element.

4.3 Datatypes

A *conformant* parser MUST determine the *datatype* which tags the *string* in the *citation element value* as follows.

If the *current element* has a non-empty `datatype` attribute, then the *datatype* SHALL be the value of `datatype` attribute once shorthand IRIs have been expanded. The `datatype` attribute MUST NOT contain the name of a *language-tagged datatype* or the built-in `rdfs:Resource` *datatype*. The use of a `datatype` attribute is RECOMMENDED for *citation elements* that are not well-known if the *datatype* is known not to be one that is prohibited in a `datatype` attribute.

Example — Suppose a vendor defines a *citation element* called `reviewDate` which contains an [ISO 8601] date. This third-party element may not be well known, so an RDFa author SHOULD mark up its use with a `datatype` attribute:

```
<span prefix="vendor: http://example.com/sources/
              xsd:    http://www.w3.org/2001/XMLSchema#"
      property="vendor:reviewDate" datatype="xsd:date"
      content="2000-10-08" />
```

By using a `datatype` attribute, the RDFa author is ensuring the application processing the data knows the *citation element* is a date and will display it to the user appropriately, even if it does not know exactly what the date signifies.

Note — Language-tagged datatypes MUST NOT be placed in a `datatype` attribute because the RDFa parsing rules mean the *language tag* is discarded if a `datatype` attribute is found.

Otherwise, if the *host language* is HTML, if the *current property value* was found in a `datetime` attribute or was the contents of a `<time>` element, an application MAY examine the *current property value*, and if it is syntactically valid as the following *structured non-language-tagged datatypes* defined in [XSD Pt2], it MAY determine that to the *datatype*:

```
http://www.w3.org/2001/XMLSchema#date
http://www.w3.org/2001/XMLSchema#time
```

17

```
http://www.w3.org/2001/XMLSchema#dateTime
http://www.w3.org/2001/XMLSchema#duration
http://www.w3.org/2001/XMLSchema#gYear
http://www.w3.org/2001/XMLSchema#gYearMonth
```

> *Note —* This rule exists for compatibility with a full HTML+RDFa parser where this behaviour is REQUIRED; implementation of this rule is otherwise NOT RECOMMENDED. Document authors SHOULD NOT rely on this behaviour, and SHOULD instead add a `datatype` attribute.

> *Example —* An application that implements this rule would read the markup below and generate a *citation element value* whose single *string* "2000-10-08" would be tagged with the `xsd:date` *datatype*.
>
> ```
> <time property="vendor:reviewDate">2000-10-08</time>
> ```
>
> Had a different HTML element been used, say a `<span>`, or if the parser does not support this rule, the *datatype* would fall back to `rdf:langString`. If this third-party *citation element* were unfamiliar to the application, it would not undergo *datatype correction* per §4.4 of [CEV Concepts], and would remain with the wrong *datatype*. For this reason, an explicit `datatype` attribute is recommended:
>
> ```
> <time property="vendor:reviewDate"
>       datatype="xsd:date">2000-10-08</time>
> ```

Otherwise, if the *current property value* was found in a `src` or `href` attribute, then the *datatype* SHALL be:

```
http://www.w3.org/2000/01/rdf-schema#Resource
```

This *datatype* IRI MUST NOT be given explicitly in a `datatype` attribute.

> *Note —* RDF's notion of a datatype is narrower than the definition in these Citation Elements standard and `rdfs:Resources` is not a datatype in the RDF sense which is why it MUST NOT be given in an RDFa `datatype` attribute.

> *Editorial note —* The handling of `src` and `href` attributes should be revisited as the [CEV Vocabulary] progresses. If there are no obvious use cases, support for them could be made OPTIONAL, with them behaving as *source-exclusion elements* if not supported.

Otherwise, the application SHALL attempt to determine whether a *language tag* is in scope per §4.4; if a *language tag* can be determined, the *datatype* SHALL be the `rdf:langString` type:

```
http://www.w3.org/1999/02/22-rdf-syntax-ns#langString
```

> *Note* — This is so that the current *language tag* is not lost, as it would be if the default were a *string*.

Otherwise, the application SHALL determine the *datatype* to be:

```
http://www.w3.org/2001/XMLSchema#string
```

> *Note* — Applications wishing not to handle the `xsd:string` *datatype* are allowed by §2.4.2 of [CEV Concepts] to change this *datatype* to
>
> ```
> http://www.w3.org/1999/02/22-rdf-syntax-ns#langString
> ```
>
> and tag the *string* with a *language tag* of und.

> *Note* — The *datatypes* selected in the last three cases are the three *datatypes* which are defined to participate in the *datatype correction* mechanism defined in §4.4 of [CEV Concepts]:
>
> ```
> http://www.w3.org/2000/01/rdf-schema#Resource
> http://www.w3.org/1999/02/22-rdf-syntax-ns#langString
> http://www.w3.org/2001/XMLSchema#string
> ```
>
> Applications MAY opt to apply *datatype correction* while parsing RDFa for *citation elements*; if so, these *datatypes* will often be replaced by the *default datatype* of the *citation element term*.

4.4 Language tags

The *language tag* of the *citation element* SHALL be the value of `xml:lang` or `lang` attribute on the *current element*, failing which on the nearest ancestor element of the *current element*. If both attributes are present on the same element, the `xml:lang` attribute takes precedence.

> *Note* — This standard does not change when the `xml:lang` and `lang` attributes may be used on an HTML element. In particular, the `xml:lang` attribute is only allowed in XHTML documents.

> *Example* —
>
> ```
> <p vocab="https://terms.fhiso.org/sources/" typeof="Source" lang="en">
>   <span property="authorName"
>         content="Settipani, Christian">Christian Settipani</span>,
>   <i property="title" lang="fr">Les ancêtres de Charlemagne</i>,
>   <span property="edition" content="2">2nd ed.</span>
> </p>
> ```
>
> This *formatted citation* is correctly tagged with the *language tag* en denoting English. This is because, even though the book's title is French, the citation as a whole is in English. Had the

> citation been written in French, the edition would have been written "2ᵉ éd" rather than "2nd ed".
>
> This example contains three *citation elements*. The authorName and edition *citation elements* both inherit the en *language tag*. In the case of authorName this may or may not be what was intended: the author is French but his name would not normally be altered in translation to English. The explicit *language tag* is necessary on the title *citation element*, as the title is clearly French.

If no applicable xml:lang or lang, an external mechanism may be used to supply the *language tag*.

> *Example —* In a document fetched via HTTP, a Content-Language header may provide the default *language tag* for the whole document.

> *Example —* If the *formatted citation* is a fragment of XHTML in a different XML language, the value of any xml:lang attributes in the host XML will be inherited by the XHTML as defined in §2.12 of [XML].

When these attributes are used in *host languages* other than HTML, the definition of the *host language* MAY provide a default *language tag* that applies in the event that no such attribute is found.

> *Note —* FHISO does *not recommend* the use of a default *language tag* when it gives privileged status to one language. If technical considerations require a default *language tag*, a neutral *language tag* such as und (defined in [ISO 639-2] to represent an undetermined language) SHOULD be used.

If no applicable xml:lang or lang attribute was found, no value was supplied through an external mechanism and no default applies, or if provided *language tag* is an empty string, the *citation element* has no *language tag*.

4.5 Parsing RDF triples

> *Note —* This section is only relevant if an implementation wishes to make greater use of the RDF features that underlie RDFa. Support for everything in this section is therefore OPTIONAL.

Applications supporting more RDFa features than this standard requires MAY determine the *current element value*, its *datatype* and, where applicable, its *language tag* from the object of an RDF triple that was identified as representing a *citation element* per §3.2 of this standard.

If the object of the RDF triple is a literal, then the *current element value* SHALL be the lexical form of the literal, as defined in §3.3 of [RDF Concepts]. Its *datatype* SHALL be the datatype IRI of the literal, and its *language tag* SHALL be the language tag of the literal if that is present exists.

Otherwise, if the object of the RDF triple is an IRI, then the *current element value* SHALL be that IRI, and its *datatype* SHALL be:

20

```
http://www.w3.org/2000/01/rdf-schema#Resource
```

> *Note* — The object of the RDF cannot be a blank node as RDF triples whose objects are blank nodes are discarded in §3.2.

5 Layered citations

Once the *citation elements* in a document have been located, parsed and grouped into *citation element sets*, the application SHALL interpret each *citation element set* as a *citation layer*.

In [CEV Concepts], a *citation* is represented with three parts:

— a list of *citation layers*;
— a marker identifying the *head citation layer*; and
— a set of *layer derivation links*.

In these RDFa bindings, *citation layers* are represented by a *source-type element* which are nested in *layered citations*.

5.1 Nested source-type elements

A **nested source-type element** is a *source-type element* that:

— is a *source-exclusion element* of some other *source-type element*, known as its **outer source-type element**, but is not contained in another *source-exclusion element* within the *outer source-type element*; and
— has an attribute named `rev` or `rel` (or has both), but does not also have an attribute named `about`, `href`, `inlist`, `resource` or `src`.

The *citation layer* represented by a *nested source-type element* SHALL be part of the same *layered citation* as the *citation layer* represented by its *outer source-type element*. *Source-type elements* MAY be nested arbitrarily deep, and multiple *nested source-type elements* MAY be present within the same *outer source-type element*: they all represent *citation layers* which are part of the same *layered citation*.

> *Example* — The following fragment of HTML represents a *layered citation* with three *citation layers*.
>
> ```
> <p vocab="https://terms.fhiso.org/sources/" typeof="CitedSource">
>   <span property="authorName">Settipani</span>, citing
>   <span rel="cites" typeof="Source"><i property="title">Vita
>     Sancti Arnulfi</i></span> and
>   <span rel="cites" typeof="Source"><i property="title">Testamentum
>     Bertichramni</i></span>.</p>
> ```

> The second <span> element is a *source-type element* by virtue of its typeof attribute, which also makes it a *source-exclusion element* of the <p> element. It has a rel attribute, and together these facts make it a *nested source-type element*. The <p> element is its *outer source-type element*. Exactly the same applies to the third <span> element, and as both are part of the same *layered citation* as their shared *outer source-type element*, both must be in the same *layered citation* as each other.
>
> As the second and third <span> elements are *source-exclusion elements* of the *outer source-type element*, their title property is only a *citation element* of the *nested source-type elements*, and not also of the *outer source-type element*. The *outer source-type element* therefore only has one *citation element*: the authorName.

All but one of the *source-type elements* in a *layered citation* will be *nested source-type elements*. The one that is not is known as the **outermost source-type element**.

The collection of *citation layers* in a *layered citation* is an ordered list, and the *citation layers* SHOULD be include given in document order.

5.2   The head citation layer

The *head citation layer* MAY be indicated by *source-type element* with a typeof attribute whose value, once shorthand IRIs have been expanded, includes the following IRI:

    https://terms.fhiso.org/sources/CitedSource

If precisely one such element exists in the *layered citation*, the *head citation layer* SHALL be the *citation element* represented by that element; otherwise the *head citation layer* SHALL be the *citation element* represented by the *outermost source-type element*. There SHALL NOT be more than one *source-type elements* in a *layered citation* with a typeof attribute whose value includes this IRIs.

> *Note* — The *head citation layer* is defined in [CEV Concepts] as the *citation layer* representing the *source* that was actually consulted, but this need not be presented first in a *formatted citation*. More generally, this suite of standard makes no recommendation on how *citation layers* should be ordered within a *formatted citation*. Different style guides make different recommendations, and the decision may depend on the precise circumstances and what the author wishes to emphasise. The CitedSource type is provided to facilitate the correct identification of the *head citation layer*, regardless of where it is placed.

> *Example* — Individual *citation elements* have not been tagged in this example for reasons of brevity.
>
> ```
> <p vocab="https://terms.fhiso.org/sources/" typeof="Source">
>   1810 U.S. census, York County, Maine, town of York,
>   p. 435 (penned), line 9, Jabez Young;
> ```

```
      <span rev="facsimileOf" typeof="CitedSource">NARA microfilm
        publication M252, roll 12</span>.</p>
```

This *formatted citation*, based on an example in [Evidence Explained], places the *head citation layer* (the microfilm) at the end of the *formatted citation*, and marks it with a `Cited-Source` type. In this case, the same effect could have been achieved by nesting the HTML elements differently:

```
  <p vocab="https://terms.fhiso.org/sources/" typeof="Source">
    <span rel="facsimileOf" typeof="Source">1810 U.S. census,
      York County, Maine, town of York, p. 435 (penned),
      line 9, Jabez Young</span>;
    NARA microfilm publication M252, roll 12.</p>
```

In this second version, there is no need to use the `CitedSource` type as it defaults to the *outermost source-type element*.

5.3 Layer derivation links

In the [CEV Concepts] data model, *layer derivation links* have components:

— the *derived reference* to the *citation layer* representing the *derived source*;
— the *base reference* to the *citation layer* representing the *base source*; and
— the *source derivation type*.

In this standard, *layer derivation links* are represented by `rel` and `rev` attributes on *nested source-type elements*.

Once shorthand IRIs have been expanded, each IRI in the `rel` and `rev` attributes SHALL be used as the *source derivation type* of a new *layer derivation link*. If the IRI was in a `rel` attribute, the *derived source* SHALL be the *source* represented by the *outer source-type element*, and the *base source* SHALL be the *source* represented by the *nested source-type element*. If the IRI was in a `rev` attribute, the *derived source* SHALL be the *source* represented by the *nested source-type element*, and the *base source* SHALL be the *source* represented by the *outer source-type element*.

> *Note* — The `rel` and `rev` attributes provide forwards and reverse versions of the same functionality: the difference being that the `rel` attribute is placed on the *base source*, while the `rev` attribute is placed on the *derived source*.

> *Example* — In the previous example, the microfilm is derived from the 1810 census returns. The first version needs to use a `rev` attribute because the *nested source-type element* is the *derived source*, while the second version uses a `rel` attribute because the *nested source-type element* is the *base source*.

> *Note —* This representation of *layer derivation links* does not allow an arbitrary set of *layer derivation links* to be encoded as there is no way to reference a *citation layer* that is encoded elsewhere, but it does cope with any tree of derivations which is the case that is anticipated to arise in practice. Applications supporting more RDFa functionality than this standard requires can express arbitrary collections of *layer derivation links*, and an example of this is given in §5.4.

5.4 Full RDFa considerations

> *Note —* This section is only relevant if an implementation wishes to make greater use of the RDF features that underlie RDFa. Support for everything in this section is therefore OPTIONAL.

Documents that use more RDFa features than this standard requires to be supported MUST NOT include any *source-type elements*, other than the *head citation layer* as determined by the above rules, whose RDF type can be inferred to be:

```
https://terms.fhiso.org/sources/CitedSource
```

> *Note —* The above restriction is to prevent a full RDFa parser from disagreeing with an application just implementing this standard over the identity of the *head citation layer*. The term "inferred" is meant broadly, and includes inferences made through entailment regimes, as defined in [RDF Semantics].

Applications MAY utilise the fact that `https://terms.fhiso.org/sources/CitedSource` is an RDF subclass of `https://terms.fhiso.org/sources/Source`.

Applications which support a larger part of RDFa MAY find additional *layer derivation links*. If so, they MUST ensure that RDFa constructs are only treated as *layer derivation links* when they produce an RDF triple whose subject and object both have the following RDF types, or a subtype thereof:

```
https://terms.fhiso.org/sources/Source
```

In addition, the predicate of the RDF triple MUST be the following, or an RDF subproperty thereof:

```
https://terms.fhiso.org/sources/derivedFrom
```

The subject of the RDF triple corresponds to *derived source* and its object is the *base source*; the predicate is the *source derivation type*. Such triples SHOULD NOT also be used to generate a *citation element* as would otherwise be permitted by §3.2.

> *Example —* In the following example, the layers have been shorted to just contain placeholder text for brevity.
>
> ```
> <p vocab="https://terms.fhiso.org/sources/" typeof="Source">
>   Source A; derived from
>   <i resource="#B" rel="derivedFrom" typeof="Source">B</i> &amp;
> ```

```
      <i rel="derivedFrom" typeof="Source">C
        <span rel="derivedFrom" resource="#B"/>
      </i>.
    </p>
```

An application conforming only to this standard will parse this and find three *citation layers,* and two *layer derivation links* saying that A is derived from both B and C. The re-source attribute on the first <i> element will be ignored, and the <span> element is a *source-exclusion element* and so will also be ignored.

However a full RDFa parser will find three derivedFrom triples. In addition to the triples saying A is derived from B and C, there is a third triple saying that C is derived from B. An application MAY use this information to generate a third *layer derivation link*.

This arrangement of three *layer derivation links* is an example that cannot be represented in the subset of RDFa that this standard requires to be supported.

6 Synchronising citation elements

When an application has both a *formatted citation* tagged with RDFa attributes per this standard and a *citation element set* for the same *citation,* the two will typically have much content in common. This introduces the possibility that the data in the two places becomes unsynchronised. This section discusses ways of avoiding this.

In general, applications SHOULD consider information from the *citation element set* to have precedence over information extracted from a *formatted citation.*

If an application allows the manual editing of *formatted citations* tagged with RDFa attributes per this standard, it SHOULD take steps to prevent this from changing the *citation element values* that a *conformant* application would extract from the *formatted citation* to be different from the *citation element values* in the *citation element set.*

Note — This document does not prescribe a particular mechanism for ensuring this, but most strategies will involve parse the RDFa attributes before and after the edit and identify any *citation elements* whose values have changed. An application might ask the user whether the change should be propagated back to the original *citation element set.* If the change is not to be propagated back to the *citation element set,* the application might delete the property attribute so the changed data is no longer recognised as a *citation element,* or insert a content attribute containing the correct data per §4.2.

Example — Suppose an application generates the following *formatted citation.*

```
    <p><span property="https://terms.fhiso.org/sources/authorName"
      >Settipani, Christian</span>.
```

```
      <i property="https://terms.fhiso.org/sources/title">Les ancêtres
        de Charlemagne</i>.</p>
```

If a user edits this HTML to replace *Les ancêtres de Charlemagne* with *Ibid.*, the application SHOULD then take steps to ensure a future parser does not believe the source literally has the title *Ibid.* In this case, clearly the change should not be propagated back to the *citation element set* as the source isn't titled *Ibid.*, and the user would presumably decline if offered this option. An application might delete the `property` attribute so *Ibid.* is not understood to be a title, or insert a `content` attribute containing real title as follows:

```
    <p><span property="https://terms.fhiso.org/sources/authorName"
        >Settipani, Christian</span>.
        <i property="https://terms.fhiso.org/sources/title"
          content="Les ancêtres de Charlemagne">Ibid.</i></p>
```

If an application stores *formatted citations* tagged with RDFa attributes as per this standard, it SHOULD take steps to ensure that changes to the underlying *citation element set* propagate to the *formatted citation*.

> *Note* — An application doing this would parse the *formatted citation* per this standard, locate the part of the HTML or XML that contains the old *citation element value* and overwrite it with the new value. For *citation elements* that are *multi-valued* elements, the application needs to know both the old and the new *citation element value* so that it knows which value is being updated; for other elements it is not necessary to know the old value.

7   Longer example

> *Example* — This example gives a full HTML document of the sort a genealogist might publish online. In a paragraph of narrative text it gives some brief details of King Edward II's birth and parents. Although brief, this information is properly sourced to three published books with the citations formatted according to the *Chicago Manual of Style*. Each of these *formatted citations* has been marked up with RDFa attributes as described in this standard. The document includes several other instances of RDFa attributes that will not be detected as *citation elements* by a compliant parser.
>
> ```
> <!DOCTYPE html>
> <html lang="en">
>   <head>
>     <meta charset="UTF-8" />
>     <title property="dc:title">Edward II</title>
>     <meta property="dc:creator" content="FHISO, Inc." />
>     <style>
>       p { max-width: 720px; }
>       .notes p, .note { font-size: smaller; }
> ```

```
    .fnref { vertical-align: super; font-size: smaller; }
    .fnref::before { content: '['; }
    .fnref::after { content: ']'; }
  </style>
</head>

<body>
  <h1>Edward II</h1>

  <p>
    Edward II was the fourth son of Edward I and his wife, Eleanor
    of Castile.<a id="fnref1" class="fnref" href="#fn1">1</a>
    He was born in Caernarfon Castle in North Wales on
    25 April 1284, less than a year after Edward I had conquered
    the region, and as a result is sometimes called Edward of
    Caernarfon.<a id="fnref2" class="fnref" href="#fn2">2</a>
    His father was the King of England, and had also inherited
    Gascony in south-western France, which he held as the
    feudal vassal of the King of France, and the Lordship
    of Ireland.<a id="fnref3" class="fnref" href="#fn3">3</a>
  </p>

  <div vocab="http://terms.fhiso.org/sources/" class="notes">
    <h2>References</h2>

    <p typeof="Source" id="fn1"><a href="#fnref1">1</a>.
      <span property="authorName">Roy Martin Haines</span>,
      <i property="title">King Edward II: His Life, his Reign and
        its Aftermath, 1284–1330</i>
      (<span property="publicationPlace">Montreal, Canada
        &amp; Kingston, Canada</span>:
       <span property="publisher">McGill-Queen's
         University Press</span>,
       <span property="publicationDate">2003</span>),
      <span property="page" content="3">3</span>.
    </p>

    <p typeof="Source" id="fn2"><a href="#fnref2">2</a>.
      <span property="authorName">Seymour Phillips</span>,
      <i property="title">Edward II</i>
      (<span property="publicationPlace">New Haven, US
          &amp; London, UK</span>:
```

```
      <span property="publisher">Yale University Press</span>,
      <span property="publicationDate">2011</span>),
    <span property="page" content="33, 36">33 &amp; 36</span>.
  </p>

  <p typeof="Source" id="fn3"><a href="#fnref3">3</a>.
    <span property="authorName">Michael Prestwich</span>,
    <i property="title">Edward I</i>
    (<span property="publicationPlace">Berkeley, US
      &amp; Los Angeles, US</span>:
     <span property="publisher">University of California
       Press</span>,
     <span property="publicationDate">1988</span>),
    <span property="page" content="13-14">13–14</span>.
  </p>
</div>

<hr/>
<p class="note">This file is an example of an HTML document
  containing formatted citations marked up with RDFa attributes
  per the FHISO draft standard
  <a href="http://tech.fhiso.org/drafts/cev-rdfa-bindings"
    >Citation Elements: Bindings for RDFa</a>.</p>

<p vocab="http://creativecommons.org/ns#"
    class="note">Content copied from
  <a href="https://en.wikipedia.org/wiki/Edward_II_of_England"
     property="dc:source">Wikipedia</a> and released under a
  <a href="http://creativecommons.org/licenses/by-sa/3.0/"
     property="license">Creative Commons License</a>.</p>
  </body>
</html>
```

8 References

8.1 Normative references

[CEV Concepts]

FHISO (Family History Information Standards Organisation). *Citation Elements: General Concepts". Exploratory draft of standard. See http://tech.fhiso.org/drafts/cev-concepts.

[RDF Concepts]

W3C (World Wide Web Consortium). *RDF 1.1 Concepts and Abstract Syntax*. W3C Recommendation, 2014. See http://www.w3.org/TR/rdf11-concepts.

[RDFa Core]

W3C (World Wide Web Consortium). *RDFa Core 1.1*. W3C Recommendation, 3rd ed., 2015. See http://www.w3.org/TR/rdfa-core.

[RFC 2119]

IETF (Internet Engineering Task Force). *RFC 2119: Key words for use in RFCs to Indicate Requirement Levels.* BCP 14. Scott Bradner, 1997. See http://tools.ietf.org/html/rfc2119.

[XML]

W3C (World Wide Web Consortium). *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. W3C Recommendation, 26 Nov 2008. See https://www.w3.org/TR/REC-xml/.

8.2 Other references

[CEV ELF]

FHISO (Family History Information Standards Organisation). *Citation Elements: Bindings for ELF". Early draft of standard.

[CEV GEDCOM X]

FHISO (Family History Information Standards Organisation). *Citation Elements: Bindings for GEDCOM X". Early draft of standard.

[Dublin Core]

Dublin Core Metadata Initiative. *Dublin Core metadata element set*. Dublin Core recommendation, version 1.1, 1999. See http://dublincore.org/documents/dcmi-terms/.

[Evidence Explained]

Elizabeth Shown Mills. *Evidence Explained*, 2nd ed. Baltimore: Genealogical Publishing Company, 2009.

[HMTL+RDFa]

W3C (World Wide Web Consortium). *HTML+RDFa 1.1*. W3C Recommendation, 2nd ed., 2015. See http://www.w3.org/TR/html-rdfa.

[HTML5+RDFa Context]

W3C (World Wide Web Consortium). *HTML5+RDFa Initial Context*. Last updated 9 Dec 2011. See http://www.w3.org/2011/rdfa-context/html-rdfa-1.1.

[ISO 639-2]

ISO (International Organization for Standardization). *ISO 639-2:1998. Codes for the representation of names of languages — Part 2: Alpha-3 code.* 1998. (See http://www.loc.gov/standards/iso639-2/.)

[ISO 8601]

ISO (International Organization for Standardization). *ISO 8601:2004. Data elements and interchange formats — Information interchange — Representation of dates and times.* 2004.

[RDF Schema]

W3C (World Wide Web Consortium). *RDF Schema 1.1.* W3C Recommendation, 2014. See http://www.w3.org/TR/rdf-schema.

[RDF Semantics]

W3C (World Wide Web Consortium). *RDF 1.1 Semantics.* W3C Recommendation, 2014. See http://www.w3.org/TR/rdf11-mt.

[RDFa Primer]

W3C (World Wide Web Consortium). *RDFa 1.1 Primer.* W3C Recommendation, 3rd ed., 2015. See http://www.w3.org/TR/rdfa-primer.

[XHMTL+RDFa]

W3C (World Wide Web Consortium). *XHTML+RDFa 1.1.* W3C Recommendation, 3rd ed., 2015. See http://www.w3.org/TR/xhtml-rdfa.

[XML Names]

W3 (World Wide Web Consortium). *Namespaces in XML 1.0 (Third Edition).* W3C Recommendation, 8 Dec 2009. See https://www.w3.org/TR/REC-xml-names/.