# Review T: Utilities domain and sorting

# Table of contents

# 1 Sorting

Processors can be configured to sort elements. Typical processing includes sorting glossary entries, index entries, list items, and tables based on the contents of cells in specific columns or rows.

Each element to be sorted must

have some inherent text on which it will be sorted. This text is the *base sort phrase* for the element. For elements that have titles, the base sort phrase usually is the content of the `<title>` element. For elements that do not have titles, the base sort phrase might be literal content in the DITA source, or it might be generated or constructed based on the semantics of the element involved; for example, it could be constructed from various attribute or metadata values.

001 (19)    Processors that perform sorting **SHOULD** explicitly document how the base sort phrase is determined for a given element.

The `<sort-as>` element can be used to specify an effective sort phrase when the base sort phrase is not appropriate for sorting. For index terms, the `<sort-as>` element specifies the effective sort phrase for an index entry.

| Disposition: Closed |
| --- |

The details of sorting and grouping are implementation specific. Processors might provide different mechanisms for defining or configuring collation and grouping details. Even where the `<sort-as>` element is specified, two processors might produce different sorted and grouped results because they might use different collation and grouping rules. For example, one processor might be configured to sort English terms before non-English terms, while another might be configured to sort them after.

002 (19)

When a `<sort-as>` element is specified, processors that sort the containing element **MUST** construct the effective sort phrase by prepending the content of the `<sort-as>` element to the base sort phrase. This ensures that two items with the same `<sort-as>` element but different base sort phrases will sort in the appropriate order.

For example, if a processor uses the content of the `<title>` element as the base sort phrase, and the title of a topic is "24 Hour Support Hotline" and the value of the `<sort-as>` element is "twenty-four hour", then the effective sort phrase would be "twenty-four hour24 Hour Support Hotline".

**Related reference**

sort-as (15)

For elements that are sorted, the `<sort-as>` element provides text that is combined with the base sort phrase to construct the effective sort phrase.

---

**Comment by dstevens73**

This description is difficult to understand without reading the entire topic; in fact, I would say I had absolutely no idea what this meant until I read the whole topic. However, I don't think I still understand it enough to suggest what to do about my comment. :(

This sentence, from the Sorting topic, might help some. Perhaps something like this:

"The `<sort-as>` element can be used to specify an effective sort phrase when the base sort phrase is not appropriate for sorting; for example, Japanese characters or letters with diacritics. This element prepends text with the base content to construct an effective sort phrase."

Also it helped me to read other sites about language sorting to have these basics: "Each language has a set or sets of rules for how language strings should be sorted. The <sort-as> element transforms those rules into a standard Latin alphabetical order (actually I'm not entirely sure about this, but the example is giving a Latin word for the Chinese character, so I think that's what it's doing?)

I'm babbling, but I guess it just validates that I, for one, didn't get it.

---

Kris Eberlein, 31 January 2023

The `<sort-as>` element is a difficult, complex element. Sort is inherently complicated, especially since details of how it is handled are left up to the processor! And it is very difficult to write a solid short description for very complex elements or concepts ...

Part of the difficulty, I think, is that the short description includes phases ("base sort phrase" and "effective sort phrase" that are defined elsewhere.

---

Nancy Harrison 3 February 2023

I find the term 'combined with' to be a bit confusing; my understanding of sorting, which I once needed to know a fair amount about and which seems to be confirmed by later content, is that the

combination is done by pre-pending the sort-as value to the base sort content, so it would make more sense to me to replace 'combined with' above by 'prepended-to'. My sense is that the only folks who will read this are those who need to understand sorting, and they will need to learn the basics somewhere else.

Kris Eberlein, 04 February 2023

Good point about replacing "combined with the base sort phrase" to "preprended to the base sort phrase"

Kris Eberlein, 14 February 2023

Recast the short description as "The <sort-as> element can be used to specify an effective sort phrase when the base sort phrase is not appropriate for sorting, for example, Japanese characters. This element prepends text to the base sort phrase in order to construct an effective sort phrase."

**Disposition: Accepted**

**Comment by Zoë Lawson on 5 Feb 2023**
Double-checking that the title of the topic was mangled by accident and isn't in the actual repo.

Kris Eberlein, 06 February 2023

As you can see from the review PDF, this is an error introduced by someone adding review comments.

**Disposition: Closed**

# 2 Utilities domain elements

The utilities domain elements describe image maps and enable sorting.

---

**Comment by Zoë Lawson on 5 Feb 2023**
This short description is describing exactly what the domain contains, but would we rather describe the purpose of the domain?

"The utilities domain includes elements to enhance existing content such as image maps and sorting."

---

Kris Eberlein, 05 February 2023

I'll have to look and see what pattern we currently use, and then we can standardize on it.

---

Kris Eberlein, 07 February 2023

Here's what we have for the container topics for the domains that we have reviewed:

- The hardware domain elements are used to document physical devices.
- The markup domain contains an element that can be used for the mention of named components in markup languages.
- The programming domain elements are used to define the syntax for programming languages. They can also be used to provide examples.
- The software domain elements are used to describe the presentation and operation of a software program.
- The user-interface domain elements are used to describe the graphical user interface of a software program.
- The XML mention domain is designed to describe and document XML document types and applications. It also can enable typographic styling, search and retrieval, and automatic indexing for XML constructs.

---

Kris Eberlein, 14 February 2022

Changed the short description to "The utilities domain elements are used to describe image maps and enable sorting."

**Disposition: Completed**

---

## 2.1 <area>
The `<area>` element defines a linkable region in an image map.

### Usage information

Each area in an image map specifies the shape and coordinates of the linkable region, as well as the link target and its link text.

---

**Comment by Kristen J Eberlein on 26 January 2023**

Do we want to use natural language in the short description? I have not done so because "area" (I think we"d need to use "An area in an image map is ..." phrasing) is such a general term. Also, it probably would really require modifying this "Usage information" section.

---

Overall, I used natural language for `<imagemap>`, but left the other elements as-is.

Bill Burns, 31 January 2023: I agree. Best to leave it as is. It's much less ambiguous that way.

**Disposition: Closed**

---

**Comment by Nancy Harrison on 03 February 2023**

I'd recommend adding a comment right about here about matching the HTML5 definitions for sub-elements, similar to the comment under coords. Alternatively, it could go under shape, which is the section whose oddity is directly the result of the matchup with HTML5.

---

Kris Eberlein, 04 February 2023

@Nancy, I moved your comment into a separate draft comment, since it does not pertain to the short description.

---

Kris Eberlein, 07 February 2023

I don't think that this topic needs any pointer to the HTML5 spec. The DITA `<area>` element is just a container. Any information about the HTML5 spec should be directly in the `<coords>` and `<shape>` topics – and that necessary info is in those topics.

**Disposition: Closed**

---

**Comment by Zoë Lawson**

I think <area> needs some extra description because it's an element that is a container and without the bits in the container, it's a little confusing.

Although, as I went back to other elements that are more container like such as ul, ol, table, dl, etc., now with the natural language, I'm not sure it's as obvious that you have to have other elements as children. However, I think that's a design decision because we're going to assume folks use an XML editor that validates?

---

Kris Eberlein, 06 February 2023

For folks who are authoring DITA content and seeing the content of the short description rendered in their authoring interface – Yes, we can assume that they are using an XML editor that inserts necessary child elements.

Marking this comment as "Closed."

**Disposition: Closed**

## Specialization hierarchy

The `<area>` element is specialized from `<div>`. It is defined in the utilities domain module.

---

**Comment by tammy-PCAS**
The <area> element is specialized from the <div> element.

---

Kris Eberlein, 07 February 2023.

No, the pattern used in this topic is consistent throughout the DITA for Technical Content spec. And it is what I specified in the editorial guidelines.

## Attributes

The following attributes are available on this element: universal attributes.

## Example

The following code sample shows how the `<area>` element defines the linkable region in an image map. It defines the shape and coordinates of the linkable region, and it specifies the link target and its link text.

```
<area>
  <shape>rect</shape>
  <coords>276, 60, 425, 460</coords>
  <xref keyref="ruytenburch">Willem van Ruytenburch</xref>
```

> **Comment by**
> **dstevens73**
> This is nitty, because it's not wrong, but I note that all other examples in using Night Watch, the xref uses an href to a wikipedia article (including this specific area in the imagemap topic.
>
> Kris Eberlein, 31 January 2023
>
> Thanks for noting the inconsistency. I suspect that I changed the markup to use @keyref for reasons of brevity. I've changed the markup in this topic to be consistent.
>
> **Disposition: Completed**

```
</area>
```

> **Comment by Zoë Lawson on 5 Feb 2023**
> Do we want to add a link to the imagemap example? "For complete context see the imagemap example."?
>
> ---
>
> Kris Eberlein, 07 February 2023
>
> We have not done that in other similar topics. Accordingly, I think this will need to be something that we consider doing consistently in the spec if we end up with time to fine tune things.
>
> **Disposition: Deferred**

## 2.2 <coords>

The `<coords>` element specifies the coordinates of a linkable region in an `<imagemap>`.

## Usage information

The values for use in the `<coords>` element are defined by the HTML5 specification. The unit used for coordinates is CSS pixels. The following table provides a summary of the syntax, which depends on the shape of the linkable region:

> **Comment by Zoë Lawson on 5 Feb 2023**
> Can we provide information that the Shape is defined by the <shape> in the same <area>?
>
> ---
>
> Kris Eberlein, 07 February 2023
>
> I don't think that additional piece of information is really needed here.

**Disposition: Rejected**

| Shape | Syntax |
|-------|--------|
| **circle** | Three integers, separated by commas: <br> 1. The distance from the left edge of the image to the center of the circle <br> 2. The distance from the top edge of the image to the center of the circle <br> 3. The radius of the circle |
| **default** | The default shape is the entire image. The `<coords>` element should be empty. |
| **poly** | An even number of at least six integers, separated by commas: <br> • Each pair of integers represents a coordinate given as the distances from the left and the top of the image respectively. <br> • The coordinates represent the points of the polygon, in order. |
| **rect** | Four integers, separated by commas: <br> 1. The distance from the left edge of the image to the left side of the rectangle <br> 2. The distance from the top edge of the image to the top side of the rectangle <br> 3. The distance from the left edge of the image to the right side of the rectangle <br> 4. The distance from the top edge of the image to the bottom side of the rectangle |

**Comment by nancylph**

Nancy Harrison 3 February 2023

So just to confirm, if there is no shape element inside the imagemap (or an empty shape element? see my comment for <shape>), I would need 4 comma-separated digits in coords, but if the value of shape is "default," there would be no coords element at all?

Kris Eberlein, 07 February 2023

OK, here is what we stated in this DITA 2.0 draft (`<coords>`) topic:

(If the shape is "default") The default shape is the entire image. The `<coords>` element should be empty.

Here's what the DITA 1.3 spec stated: Nothing. We did not list "default" in the `<coords>` topic. But we did state "The syntax of the coordinate data depends on the shape described by the coordinates, and is based on the image map definition in HTML." See http://docs.oasis-open.org/dita/dita/v1.3/ errata02/os/complete/part3-all-inclusive/langRef/base/coords.html#coords`<coords>`

I kind of suspect that the only time the shape would be default would be the following scenario: An single <area> element with an empty <coords> element. Or maybe processors would simply disregard what was specified in the <coords> element.

Kris Eberlein, 07 February 2023

We discussed this at today's TC call. To respond to Nancy's original set of questions:

• The `<shape>` element is required.
• If the `<shape>` element is empty, processors will assume the rectangle shape, and so the `<coords>` element will need to specify four comma-separated digits.

> - If the `<shape>`element contains `default`, then processors assume that the entire image is the linkable region. The `<coords>` element can be empty. I suspect, but can't verify it, that if the `<coords>` is NOT empty, that processors will ignore any coordinates that are specified.
>
> **Disposition: Closed**

### Specialization hierarchy

The `<coords>` element is specialized from `<ph>`. It is defined in the utilities-domain module.

### Attributes

The following attributes are available on this element: universal attributes and `@keyref`.

For this element, the `@translate` attribute has a default value of "no".

### Example

The following code sample shows how the `<coords>` element specifies the coordinates of a polygonal region in an image map:

```
<area>
  <shape>poly</shape>
  <coords>119, 4, 90, 7, 87, 20, 53, 36, 45, 51, 7, 188, 12, 467,
          223, 464, 240, 315, 223, 254, 210, 168, 193, 146, 173, 121, 167,
          87, 169, 70, 181, 57, 189, 35, 164, 24, 140, 4
  </coords>
  <xref format="html" scope="external"
        href="https://en.wikipedia.org/wiki/Frans_Banninck_Cocq">
        Willem van Ruytenburch</xref>
</area>
```

## 2.3 <imagemap>

A image map is an image with areas that are linked to DITA topics, web pages, PDFs, or any other resources that can be targeted by the `@href` attribute.

> **Comment by Stan Doherty**
> Hmm . . . "The DITA imagemap element serves the same purpose as the HTML <map> element. The <imagemap> element references an image file and then defines shapes (region coordinates) that can be clicked to execute a JavaScript script or to navigate to a referenced topic or page."
>
> ---
>
> Kris Eberlein, 07 January 2023
>
> I'm going to leave the short description as it is, since it is a clear, natural-language definition of an image map.
>
> The open question for me is whether we want to discuss – in the "Usage information" section – whether the DITA `<imagemap>` element is related to the HTML `<map>` element.
>
> ---
>
> Kris Eberlein, 14 February 2023
>
> Talked about this with Robert Anderson at our spec editors' call yesterday. We decided *against* adding this information. HTML is now a very fluid standard, and we also did not want to go back and look at all the mappings between HTML5 <map> and DITA <imagemap>.

## Usage information

A DITA image map references the image, followed by a series of `<area>` elements that define the linkable regions of the image. Each `<area>` element specifies the shape and coordinates of the region, as well as the link target and link text.

**Comment by Zoë Lawson on 5 Feb 2023**
Can we please repeat that second sentence in the <area> element? I think that will solve my confusion.

(Huh. This information is in the example for <area>, but not the usage area. )

Kris Eberlein, 07 January 2023

@Zoe, I'm not sure I'm clear about what you are requesting. The "Usage information" section for the `<area>` topic clearly states: "Each area in an image map specifies the shape and coordinates of the linkable region, as well as the link target and its link text."

**Disposition: Closed**

## Rendering expectations

The rendering expectations for the `<imagemap>` element depends on the output format. For HTML-based formats, it can be rendered as standard HTML images or as an alternate form of navigation, such as table-based image maps. For print-based formats, it can be rendered as an image and a list of the specified link targets.

## Specialization hierarchy

The `<imagemap>` element is specialized from `<div>`. It is defined in the utilities domain module.

## Attributes

The following attributes are available on this element: display attributes and universal attributes.

## Example

The following code sample shows an image map for a detail of Rembrandt's "The Night Watch":

```
<imagemap id="the-night-watch">
  <image href="Detail_from_The_Night_Watch.jpg" id="night_watch">
    <alt>Detail from Rembrandt's The Night Watch</alt>
  </image>
  <!-- Area #1: Frans Banninck Cocq -->
  <area>
    <shape>poly</shape>
    <coords>119, 4, 90, 7, 87, 20, 53, 36, 45, 51, 7, 188, 12, 467,
        223, 464, 240, 315, 223, 254, 210, 168, 193, 146, 173, 121, 167,
        87, 169, 70, 181, 57, 189, 35, 164, 24, 140, 4</coords>
    <xref format="html" scope="external"
        href="https://en.wikipedia.org/wiki/Frans_Banninck_Cocq">
        Frans Banninck Cocq</xref>
  </area>
```

```
  <!-- Area #2: A member of the schutterij (the night watch) -->
  <area>
    <shape>circle</shape>
    <coords>223, 98, 48</coords>
    <xref format="html" scope="external"
          href="https://en.wikipedia.org/wiki/Schutterij">A member of the
          schutterij (the night watch)</xref>
  </area>
  <!-- Area #3: Willem_van_Ruytenburch -->
  <area>
    <shape>rect</shape>
    <coords>276, 60, 425, 460</coords>
    <xref format="html" scope="external"
          href="https://en.wikipedia.org/wiki/Willem_van_Ruytenburch">
          Willem van Ruytenburch</xref>
  </area>
</imagemap>
```

The following image shows the areas that are defined by the image map. Each of the three supported shapes are used.
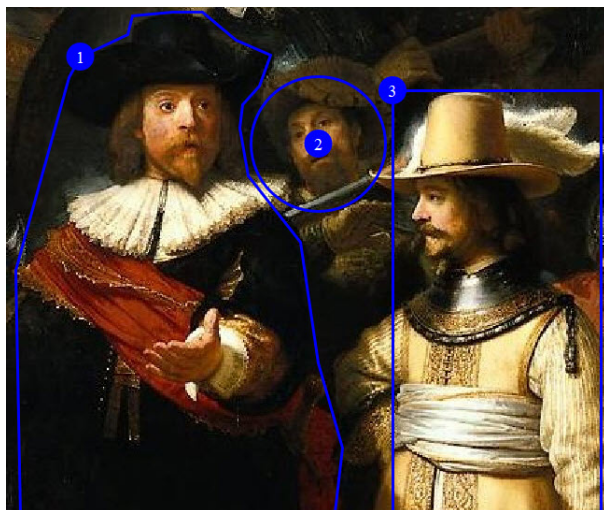


The following table lists the defined areas, the shape used, alternate text, and link targets:

| Area | Shape | Alternate text | Link target |
|------|-------|----------------|-------------|
| 1 | Polygon | Frans Banninck Cocq | Wikipedia entry for Frans Banninck Cocq |
| 2 | Circle | A member of the schutterij (the night watch) | Wikipedia entry for Schutterij |
| 3 | Rectangle | Willem van Ruytenburch | Wikipedia entry for Willem van Ruytenburch |

## 2.4 <shape>

The <shape> element defines the shape of a linkable region in an image map.

> **Comment by Zoë Lawson on 5 Feb 2023**
> So we're not repetitive, "...defines the outline of a linkable region..."?
>
> ---
>
> Kris Eberlein, 07 January 2023
>
> I think the current wording is more precise, so I am leaving it as-is.

## Usage information

The following values are supported as the content of the `<shape>` element:

> **Comment by Zoë Lawson on 5 Feb 2023**
> Should each of the dd have a period at the end since they're sentences?
>
> ---
>
> Kris Eberlein, 05 February 2023
>
> No. We standardize on the *IBM Style Guide*, which does not use periods if the text is not a complete sentence.
>
> **Disposition: Closed**

**circle**
  Indicates that the linkable region is a circle

**default**
  Indicates the linkable region is the entire image

**poly**
  Indicates that the linkable region is a polygon

**rect**
  Indicates that the linkable region is a rectangle

If the `<shape>` element is left empty, a rectangular shape is assumed.

> **Comment by dstevens73**
> I found this somewhat confusing. If the shape is blank, wouldn't default be the value, so the region is the entire image, which is assumed to be rectangular?
>
> ---
>
> Kris Eberlein, 31 January 2023
>
> This is messy in the HTML5 spec, alas. Here's what is says:
>
> - "The [shape] attribute may be omitted. The *missing value default* is the rectangle state."
> - "In the default state, area elements must not have a @coords attribute. The area is the whole image."
>
> Note that in HTML5, many of the properties for which DITA uses elements are represented by attributes.
>
> ---
>
> I've made the following changes:
>
> Changed the definition of "default" to read "Indicates the linkable region is a rectangle that exactly covers the entire image". (There are two sentences that define "default" in the HTML5 spec; this is the most precise one.)
>
> ---
>
> Nancy Harrison 3 February 2023
>
> Also, are we saying that the shape element is required, or not? The above says 'if it's left empty", but that's different from omitting the shape element completely, as is apparently the case in HTML5. If we're following their paradigm, the element should be optional.

Kris Eberlein, 07 February 2023

We discussed this at the TC call today. Per Nancy's comment, the DITA grammar files require the `<shape>` element.

Per Dawn's comment, I added an additional example to the `<imagemap>` topic that covers the use case of specifying the shape is "default">

**Disposition: Completed**

**Comment by Nancy Harrison on 03 February 2023**

Can we put in somewhere that this follows what's in HTML5? Otherwise it's looks like we're the ones being so messy and inconsistent. We could put it either here or in the area topic...

Kris Eberlein, 04 February

I've changed the paragraph that precedes the definition list to read as follows:

"The values for use in the element are defined by the HTML5 specification. The following values are supported as the content of the element:"

I then removed the duplicate sentence about HTML5 that was at end of the "Usage information" section.

**Disposition: Completed**

The values for use in the `<shape>` element are defined by the HTML5 specification.

## Specialization hierarchy

The `<shape>` element is specialized from `<keyword>`. It is defined in the utilities domain module.

## Attributes

The following attributes are available on this element: universal attributes and `@keyref`.

For this element, the `@translate` attribute has a default value of "no".

## Example

See `<imagemap>` (10).

**Comment by dstevens73**
I have no objection to this, but why didn't we do the same reference in area and coords?

Kris Eberlein, 31 January 2023

Both the examples in the `<area>` and `<coords>` topics could stand alone and only required a few lines of code. When we can avoid cross referencing to another topic – and emphasize the relevant element with bold highlighting, I think that's the best thing to do.

But here, for `<shape>`, I think the reader really needs the context of the whole `<imagemap>`code sample.

**Disposition: Closed**

## 2.5 &lt;sort-as&gt;

For elements that are sorted, the `<sort-as>` element provides text that is combined with the base sort phrase to construct the effective sort phrase.

---

**Comment by dstevens73**

This description is difficult to understand without reading the entire topic; in fact, I would say I had absolutely no idea what this meant until I read the whole topic. However, I don't think I still understand it enough to suggest what to do about my comment. :(

This sentence, from the Sorting topic, might help some. Perhaps something like this:

"The `<sort-as>` element can be used to specify an effective sort phrase when the base sort phrase is not appropriate for sorting; for example, Japanese characters or letters with diacritics. This element prepends text with the base content to construct an effective sort phrase."

Also it helped me to read other sites about language sorting to have these basics: "Each language has a set or sets of rules for how language strings should be sorted. The <sort-as> element transforms those rules into a standard Latin alphabetical order (actually I'm not entirely sure about this, but the example is giving a Latin word for the Chinese character, so I think that's what it's doing?)

I'm babbling, but I guess it just validates that I, for one, didn't get it.

---

Kris Eberlein, 31 January 2023

The `<sort-as>` element is a difficult, complex element. Sort is inherently complicated, especially since details of how it is handled are left up to the processor! And it is very difficult to write a solid short description for very complex elements or concepts ...

Part of the difficulty, I think, is that the short description includes phases ("base sort phrase" and "effective sort phrase" that are defined elsewhere.

---

Nancy Harrison 3 February 2023

I find the term 'combined with' to be a bit confusing; my understanding of sorting, which I once needed to know a fair amount about and which seems to be confirmed by later content, is that the combination is done by pre-pending the sort-as value to the base sort content, so it would make more sense to me to replace 'combined with' above by 'prepended-to'. My sense is that the only folks who will read this are those who need to understand sorting, and they will need to learn the basics somewhere else.

---

Kris Eberlein, 04 February 2023

Good point about replacing "combined with the base sort phrase" to "preprended to the base sort phrase"

---

Kris Eberlein, 14 February 2023

Recast the short description as "The <sort-as> element can be used to specify an effective sort phrase when the base sort phrase is not appropriate for sorting, for example, Japanese characters. This element prepends text to the base sort phrase in order to construct an effective sort phrase."

**Disposition: Accepted**

---

**Comment by Zoë Lawson on 5 Feb 2023**
Double-checking that the title of the topic was mangled by accident and isn't in the actual repo.

## Usage information

Sort text can be specified in the content of the `<sort-as>` element or in the `@value` attribute on the `<sort-as>` element. The `<sort-as>` elementis also useful for elements where the base sort phrase is inadequate or non-existent, such as an index entry for a Japanese Kanji phrase.

If a `<keyword>` element is used within `<sort-as>`, the `@keyref` attribute can be used to set the sort phrase. If a `<keyword>` uses `@keyref` and would otherwise also act as a navigation link, the link aspect of the `@keyref` attribute is ignored.

Some elements in the base DITA vocabulary are natural candidates for sorting, including topics, definition list entries, index entries, and rows in tables and simple tables. Authors are likely to include `<sort-as>` elements in the following locations:

- For topics, the `<sort-as>` element can be included directly in `<title>` or `<titlealt>` when the different forms of title need different effective sort phrases. If the effective sort phrase is common to all the titles for a topic, the `<sort-as>` element can be included as a direct child of the `<prolog>` element in the topic.
- For glossary entry topics, the `<sort-as>` element can be included directly in `<glossterm>` or as a direct child of the `<prolog>` element.
- For topic references, the `<sort-as>` element can be included directly in the `<titlealt>` element with a `@title-role` of `navigation`, such as `<navtitle>`, within `<topicmeta>` or as a child of `<topicmeta>`.

- For definition list items, the `<sort-as>` element can be included in the `<dt>` element.
- For index entries, the `<sort-as>` can be included as a child of `<indexterm>`. In a multilevel `<indexterm>` element, the `<sort-as>` element only affects the level in which it occurs.

## Processing expectations

If the `@value` attribute is not specified and the `<sort-as>` element does not contain content, then the `<sort-as>` element has no effect.

As a specialization of `<data>`, the `<sort-as>` element is allowed in any context where `<data>` is allowed. However, the presence of `<sort-as>` within an element does not necessarily indicate that the containing element should be sorted. Processors can choose to sort any DITA elements for any reason. Likewise, processors are not required to sort any elements.

| 003 (19) | Processors **SHOULD** expect to encounter `<sort-as>` elements in the above locations. Processors that sort **SHOULD** use the following precedence rules: |
|---|---|
| | • A `<sort-as>` element that is specified in a title takes precedence over a `<sort-as>` element that is specified as a child of the topic prolog. |
| | • Except for instances in the topic prolog, processors only apply `<sort-as>` elements that are either a direct child of the element to be sorted or a direct child of the title- or label-defining element of the element to be sorted. |
| | • When an element contains multiple, direct-child, `<sort-as>` elements, the first direct-child `<sort-as>` element in document order takes precedence. |
| | • It is an error if there is more than one `<sort-as>` child for a given `<indexterm>` element. |
| | • Sort phrases are determined after filtering and content reference resolution occur. |
| 004 (19) | When a `<sort-as>` element is specified, processors that sort the containing element **MUST** construct the effective sort phrase by prepending the content of the `<sort-as>` element to the base sort phrase. This ensures that two items with the same `<sort-as>` element but different base sort phrases will sort in the appropriate order. |
| | For example, if a processor uses the content of the `<title>` element as the base sort phrase, and the title of a topic is "24 Hour Support Hotline" and the value of the `<sort-as>` element is "twenty-four hour", then the effective sort phrase would be "twenty-four hour24 Hour Support Hotline". |

## Specialization hierarchy

The `<sort-as>` element is specialized from `<data>`. It is defined in the utilities-domain module.

## Attributes

The following attributes are available on this element: universal attributes, `@name`, and `@value`.

For this element,

- The `@name` attribute has a default value of "sort-as".
- The `@value` attribute specifies text to combine with the base sort phrase to create an effective sort phrase. When the `<sort-as>` element has content and the `@value` attribute is specified, the `@value` attribute takes precedence.

## Example

**Comment by nancylph**

The following examples illustrate how a glossary entry for the Chinese ideographic character for "big" might specify an effective sort phrase of "dada" (the Pin-Yin transliteration for Mandarin):

### Figure 1: <sort-as> within <glossterm>

In the following code sample, the `<sort-as>` element is located within the `<glossterm>` element:

```
<glossentry id="gloss-dada">
  <glossterm>
    <sort-as value="dada"/>
    &#x5927;&#x5927;
  </glossterm>
  <glossdef>Literally "big big".</glossdef>
</glossentry>
```

### Figure 2: <sort-as> within <prolog>

In the following code sample, the `<sort-as>` element is located within the `<prolog>` element of the glossary entry topic:

```
<glossentry id="gloss-dada">
  <glossterm>&#x5927;&#x5927;</glossterm>
  <glossdef>Literally "big big".</glossdef>
  <prolog>
    <sort-as>dada</sort-as>
  </prolog>
</glossentry>
```

**Related concepts**

Sorting (3)

Processors can be configured to sort elements. Typical processing includes sorting glossary entries, index entries, list items, and tables based on the contents of cells in specific columns or rows.

# A Aggregated RFC-2119 statements

This appendix contains all the normative statements from the DITA 2.0 specification. They are aggregated here for convenience in this non-normative appendix.

| Item | Conformance statement |
|---|---|
| 001 (3) | Processors that perform sorting **SHOULD** explicitly document how the base sort phrase is determined for a given element. |
| 002 (4) | When a `<sort-as>` element is specified, processors that sort the containing element **MUST** construct the effective sort phrase by prepending the content of the `<sort-as>` element to the base sort phrase. This ensures that two items with the same `<sort-as>` element but different base sort phrases will sort in the appropriate order.<br><br>For example, if a processor uses the content of the `<title>` element as the base sort phrase, and the title of a topic is "24 Hour Support Hotline" and the value of the `<sort-as>` element is "twenty-four hour", then the effective sort phrase would be "twenty-four hour24 Hour Support Hotline". |
| 003 (17) | Processors **SHOULD** expect to encounter `<sort-as>` elements in the above locations. Processors that sort **SHOULD** use the following precedence rules:<br><br>• A `<sort-as>` element that is specified in a title takes precedence over a `<sort-as>` element that is specified as a child of the topic prolog.<br>• Except for instances in the topic prolog, processors only apply `<sort-as>` elements that are either a direct child of the element to be sorted or a direct child of the title- or label-defining element of the element to be sorted.<br>• When an element contains multiple, direct-child, `<sort-as>` elements, the first direct-child `<sort-as>` element in document order takes precedence.<br>• It is an error if there is more than one `<sort-as>` child for a given `<indexterm>` element.<br>• Sort phrases are determined after filtering and content reference resolution occur. |
| 004 (17) | When a `<sort-as>` element is specified, processors that sort the containing element **MUST** construct the effective sort phrase by prepending the content of the `<sort-as>` element to the base sort phrase. This ensures that two items with the same `<sort-as>` element but different base sort phrases will sort in the appropriate order.<br><br>For example, if a processor uses the content of the `<title>` element as the base sort phrase, and the title of a topic is "24 Hour Support Hotline" and the value of the `<sort-as>` element is "twenty-four hour", then the effective sort phrase would be "twenty-four hour24 Hour Support Hotline". |

# Index

**B**

base sort phrase 3

**C**

collation 3

**D**

definitions
    base sort phrase 3
domains
    utilities 6

**E**

effective sort phrase 3
examples
    effective sort phrase 3

**G**

grouping 3

**P**

processing
    sorting 3
processing expectations
    base sort phrase, documentation of 3

**S**

sorting 3

**U**

utilities domain
    `<area>` 6
    `<coords>` 8
    `<imagemap>` 10
    `<shape>` 12
    `<sort-as>` 15