

## Metadata

Name	Content	Scheme
Coverage	UK	
Creator	esd-toolkit – administration@esd-toolkit.org	
Date.Created	2005-06-30	
Date.Issued	2005-10-09	
Date.Modified	2006-11-15	
Description	Technical guidance notes on the structure of the WebPortal and LocalDirectgov schemas for describing in an XML format an agency's services and how they are delivered.	
Language	Eng	ISO639-2
Publisher	Porism Limited, SW9 8BJ info@porism.com	
Rights.Copyright	<a href="http://www.hmso.gov.uk/copyright.htm">http://www.hmso.gov.uk/copyright.htm</a> Crown copyright	
Status	Version 1.02	
Subject	Information management	IPSV
Title	WebPortal and LocalDirectgov XML Schema – Technical Guidance Notes	
Source	<a href="http://www.esd.org.uk/standards/LocalDirectgov/Documents/LocalDirectgovSchemaTechnicalGuidance.pdf">http://www.esd.org.uk/standards/LocalDirectgov/Documents/LocalDirectgovSchemaTechnicalGuidance.pdf</a>	

## Contents

Metadata.....	1
1 Purpose.....	2
2 Explanation of the Schemas.....	2
2.1 Distinction Between the WebPortal and LocalDirectgov Schemas.....	2
2.2 The Files.....	2
2.3 Schema Design Criteria .....	3
2.4 Root Element .....	4
2.5 Metadata Element .....	5
2.6 Agency Element.....	7
2.7 Service and Interaction Elements.....	8
2.7.1 Service Elements.....	11
2.7.2 Interaction Elements .....	12
2.8 FAQ Elements.....	16
3 Special Data Types Explained .....	16
3.1 Metadata Elements Populated From Controlled Lists .....	16
3.2 Blocks Of Text That May Be Structured .....	17
4 Validations Outside the Schema .....	19
5 Sample Data Conversions .....	19
Appendix A – Full Schema Diagram.....	20

Ref 05044

## 1 Purpose

This document describes the schemas issued by esd-toolkit in conjunction with the Local Directgov Project to define the structure of XML data describing the services delivered by a local authority or other agency.

This document should be read in conjunction with the Business Guidance Notes at:

<http://www.esd.org.uk/standards/LocalDirectgov/Documents/LocalDirectgovSchemaBusinessGuidance.pdf>

## 2 Explanation of the Schemas

### 2.1 Distinction Between the WebPortal and LocalDirectgov Schemas

The WebPortal schema defines the structure of XML data defining the delivery of services by any agency and suitable for inclusion in a portal covering a wide variety of public service agencies.

The LocalDirectgov portal adds further restrictions to the structure of the XML data so that it is suitable for the Local Directgov project, feeding data to the Directgov portal. The LocalDirectgov schema requires use of local government controlled lists to populate certain values which are not subject to this constraint in the WebPortal schema.

XML that complies with the LocalDirectgov schema will always comply with the WebPortal schema. The converse is not necessarily true.

### 2.2 The Files

Version 1.3 of the LocalDirectgov schema can be found at the address:

<http://www.esd.org.uk/standards/XmlSchemas/LocalDirectgov-v1.3.xsd>.

Version 1.2 of the WebPortal schema can be found at the address:

<http://www.esd.org.uk/standards/XmlSchemas/WebPortal-v1.2.xsd>.

The schemas define the structure of an XML ‘file’ (used in this document to mean any electronic source of XML data) containing information for a single public sector agency.

The XML file defines the services offered by the agency and associated guidance in the form of Frequently Asked Questions (FAQs).

Services are defined in a hierarchical structure going from broad service areas at the top of the hierarchy to more narrow services and, at the bottom (also known as ‘leaf’) layer, interactions supporting service delivery.

The number of levels in the hierarchy is not prescribed so that the service data of most councils can readily be transformed into the LocalDirectgov format whether that information is held in a deep hierarchical structure or just as a flat list of services and associated interactions (or just interactions). Practically a hierarchy (typically of about three levels) means that searches may result in broad service headings being found from where one may drill down to a specific service interaction. Such searches lead to only a manageable number of

matching results and expose users (citizens, businesses, other agencies, etc) to information on related services about which they might not have known.

FAQ information is defined separately from services and cross-referenced with services.

Both services (including related interactions) and FAQs contain metadata that cross-references with national standard lists at [www.esd.org.uk/standards](http://www.esd.org.uk/standards). Some tagging against these lists is required and more tagging is supported.

Additional tagging against local vocabularies is supported. By extending the complex type ItemReferenceStructure, an agency can declare its own scheme names for metadata elements.

Section 3.1 below describes the ItemReferenceStructure, how it is used for particular values and how it is extended for values from specific controlled lists.

### 2.3 Schema Design Criteria

The schema is designed according to GovTalk schema guidelines (see [http://www.govtalk.gov.uk/schemasstandards/developerguide\\_document.asp?docnum=946](http://www.govtalk.gov.uk/schemasstandards/developerguide_document.asp?docnum=946)), which encourage use of existing external schemas where appropriate.

Hence the schema uses Dublin Core (see <http://www.dublincore.org/>) standards and conventions where possible. Metadata elements are prefixed with ‘dc:’ or ‘dcterms:’ where they exist within Dublin Core or Qualified Dublin Core. Metadata elements that only exist in eGMS are prefixed with ‘egms:’

The same XML structure is used for metadata as Dublin Core, but the Dublin Core schemas (see <http://dublincore.org/schemas/xmls/>) are not used because some complex types are expressed as restrictions of simple types. Tools such as Altova’s XML Spy do not support this practice. Metadata encoding guidelines at <http://dublincore.org/documents/2003/04/02/dc-xml-guidelines/> are followed.

Where XML element types are not catered for by Dublin Core, the LocalDirectgov schema includes any suitable schemas from the Government Data Standards Catalogue (GDSC, see <http://www.govtalk.gov.uk/gdsc/>).

New XML element types are introduced where there is nothing suitable in GDSC. These will be submitted for inclusion in GDSC.

Schemas are consistent with those published at the esd-standards site (<http://www.esd.org.uk/standards>) for controlled lists, mappings between lists and structured content use by the Local Government LAWs and Knowledge Management projects. These share applicable common types.

The schema file has metadata in the standard structure given by GovTalk, as illustrated below.

```
<xs:annotation>  
  <xs:appinfo xmlns:gms="http://www.govtalk.gov.uk/CM/gms">
```

```

<Metadata xmlns="http://www.govtalk.gov.uk/CM/gms-xs">
  <Audience>e-service developers</Audience>
  <Creator>esd-toolkit http://www.esd-toolkit.org</Creator>
  <Date>
    <Created>2005-05-13</Created>
  </Date>
  <Description>Root element of a Local Direct Goc XML</Description>
  <Format>
    <MediaType>text/xml</MediaType>
    <Syntax>http://www.w3.org/2001/XMLSchema</Syntax>
    <Description>XML schema, W3C Recommendation 2001</Description>
  </Format>
  <Identifier>LocalDirectgov -v1.0</Identifier>
  <Language>eng</Language>
  <Publisher>Porism Limited http://www.porism.com</Publisher>
  <Rights>
    <Copyright>Crown Copyright 2005</Copyright>
  </Rights>
  <Title>Root element of a Local Direct Goc XML</Title>
</Metadata>
</xs:appinfo>
</xs:annotation>

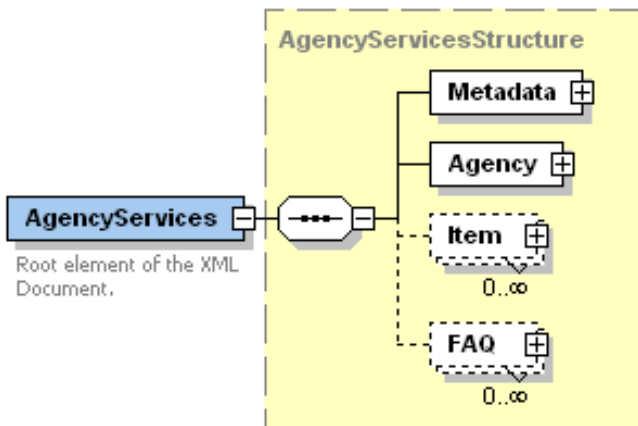
```

The schema as a whole is validated using GovTalk’s Schema QA Stylesheet at [http://www.govtalk.gov.uk/schemasstandards/developerguide\\_document.asp?docnum=944](http://www.govtalk.gov.uk/schemasstandards/developerguide_document.asp?docnum=944).

## 2.4 Root Element

The root element of the XML file is called “AgencyServices” and is of type “AgencyServicesStructure”. This has four elements, each with its own complex type:

- Metadata – describing the file’s creation, publishing, status and scope
- Agency – describing the agency whose services the file describes
- Item – defining services and interactions in a hierarchical structure
- FAQ – defining FAQs issued by the agency



In addition to metadata applicable to the file as a whole, metadata is defined for the Item and FAQ elements. All metadata is drawn from a common set of metadata types.

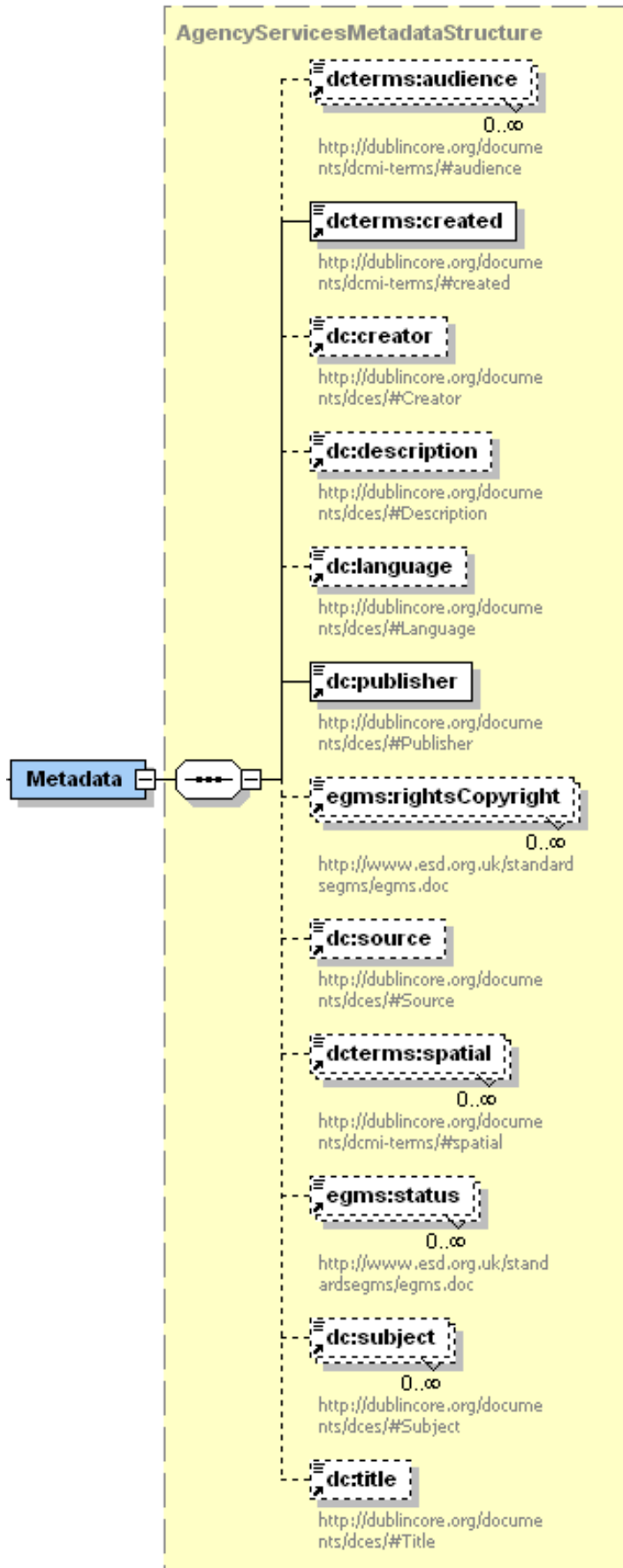
Each of the four main complex types making up AgencyServices is defined in the following subsections of this document.

## 2.5 Metadata Element

The complex type `AgencyServicesMetadataStructure` describes metadata applicable to the services XML file as a whole.

Elements are:

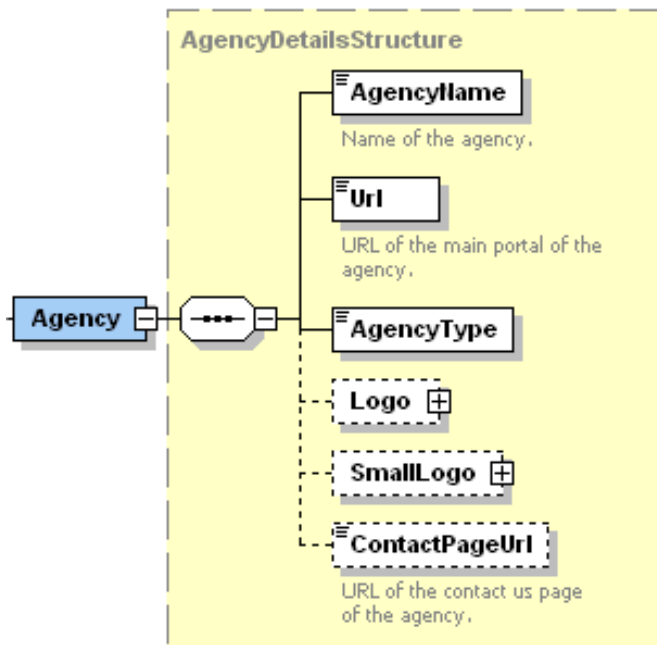
- `dc:terms:audience` – the audience for the services. This is of type `SimpleLiteralStructure`, normally extended to type `LGAL`. Audience should not be specified for agencies, such as Local Authorities, whose audience is not limited to a particular range of service users
- `dc:terms:created` – the date the file was created
- `dc:creator` – the organisation that created the data. This is of type `SimpleLiteralStructure`, normally extended to type `ONS-SNAC` to specify a particular administrative organisation.
- `dc:description` – A free text description of the content of the data such as “Definition of the services delivered by *AgencyName*, in format defined by `LocalDirectgov`”
- `dc:language` – the language of the data. The `Xsi:type` can be made “ISO639-2” to identify that the value comes from this specific scheme. There may be only one language element. If an agency produces data for more than one language, separate files should be used – one for each language.
- `dc:publisher` – the organisation that published the data. This is of type `SimpleLiteralStructure`, normally extended to type `ONS-SNAC` to specify a particular administrative organisation.
- `egms:rightsCopyright` – copyright information
- `dc:source` – a reference to the resource from which the data is derived (eg a specific CMS)
- `dc:terms:spatial` – spatial (ie geographic) coverage of the data. The element contains the name of an administrative body (eg a council covering a specific area). The `Xsi:type` is from an encoded list with just “ONS SNAC” at present. The `Id` is the ONS SNAC code for the administrative unit.
- `egms:status` – status of the file. This is of type `SimpleLiteralStructure`, normally extended to type `IeeeStatus` to specify status from a controlled list. The status of ‘Final’ should be used for a complete XML file ready for issue. The type may be extended to “Version” to provide a numeric version number.
- `dc:subject` – Most appropriate term(s) (the element is repeatable) describing the subject matter of the data. The `Xsi:type` is from an encoded list including “IPSV”. The IPSV value of “Information management” should normally be used. At least one `dc:subject` element with an IPSV term should be used. `dc:subject` elements referencing other controlled lists are accepted.
- `dc:title` - Short free text description of the content of the data such as “Services provided by *AgencyName*”



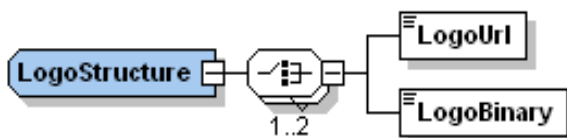
## 2.6 Agency Element

The AgencyDetailsStructure defines the agency whose data is described.

- Name – the agency name of type AgencyReferenceStructure. Normally assigned type ONS:SNAC
- URL – the URL of the agency Web site of type xs:anyURI
- AgencyType – the type of the agency from LGAtL.
- Logo – the agency logo that should have a maximum height and width of 200 pixels. This logo is intended for use on web pages and other resources that credit the agency.
- SmallLogo – a small version of the agency logo. , which should have a maximum height of 80 pixels and width of 100 pixels. This logo is intended for use on individual rows of a Web page or other resource that credits the agency.
- ContactPageUrl – the URL of the ‘Contact Us’ Web page of the agency of type xs:anyURI. Note that this element is *optional* in the WebPortal schema and *required* in the LocalDirectgov schema

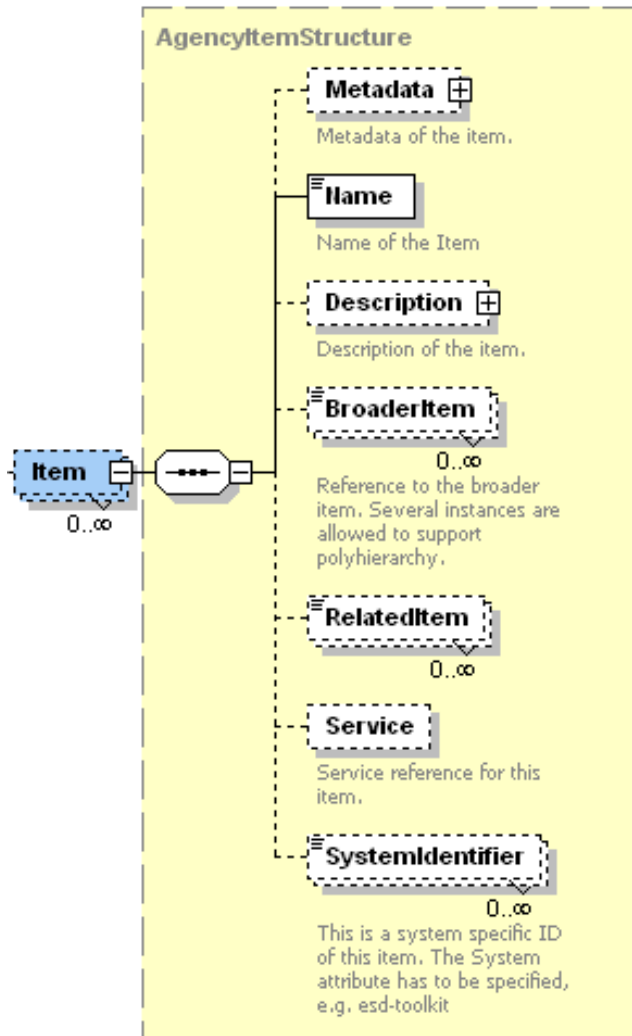


Both logo elements are of type LogoStructure which supports either a URL to the logo or a binary copy of it. If a URL is used, it must represent an absolute (not relative) address.



## 2.7 Service and Interaction Elements

Services and interactions are defined by Item elements of the AgencyItemStructure complex type.



The AgencyItemStructure is extended to either an AgencyServiceStructure (see section 2.7.1 below) or services or an AgencyInteractionStructure for interactions (see section 2.7.2 below).

Items of type “Interaction” should all have at least one Service and one interaction type in their subject metadata. For the LocalDirectgov schema, service should be populated from LGSL and interaction type from LGIL.

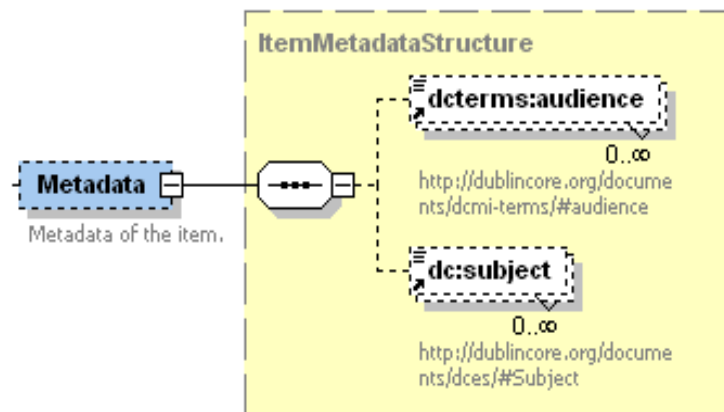
Items with no BroaderItem element are considered at the top of the hierarchical structure and should all have at least one IPSV term in their subject metadata.

Items have an ‘Id’ attribute, which takes an integer used for internal cross referencing between items. Item ids may change between instances of the data for an agency and so should only be used for cross-referencing within one XML file.

The AgencyItemStructure complex type has the following elements:

- Metadata accepting two elements of metadata:
  - Audience – identifying audience group(s) to whom the service is particularly applicable. Services applicable to most or all audience groups should not be tagged. The Xsi:type is from an encoded list with just “LGAL” initially. This tagging will be interpreted as overriding the audience assignments made centrally between LGSL and LGAL.
  - Subject – identifying the subject matter of the service. Attributes are:
    - Xsi:type which can be:
      - IPSV for subject terms
      - LGSL for standard services
      - LGIL for standard interaction types
      - A local scheme for a local vocabulary  
A local authority can decide to use its own scheme name. To do so, it is necessary to declare a complex type with the scheme name which extends the complex type “ItemReferenceStructure” (see section 3.1 below).
      - Omitted for words from no controlled vocabulary. Hence the attribute is optional
    - Id – giving the id for LGSL, LGIL and other schemes for which use of the Id is recommended

Subject terms from no controlled vocabulary (and hence with no Xsi:type or Id) are used for local proper names applied to services (eg “Cleanaway”, representing the name of a waste management service provider).



Note that item metadata is optional because the schema does allow for services *beneath* the top level of the hierarchical structure but **above** specific interactions having no metadata. However, in most circumstances metadata should be assigned to each item. In particular one or more IPSV terms *must* be assigned to top level items and an LGIL interaction type must be assigned to each interaction.

- Name – free text short name for the service/interaction
- Description – Descriptive text of complex type DescriptionStructure, described in section 3.2 below.

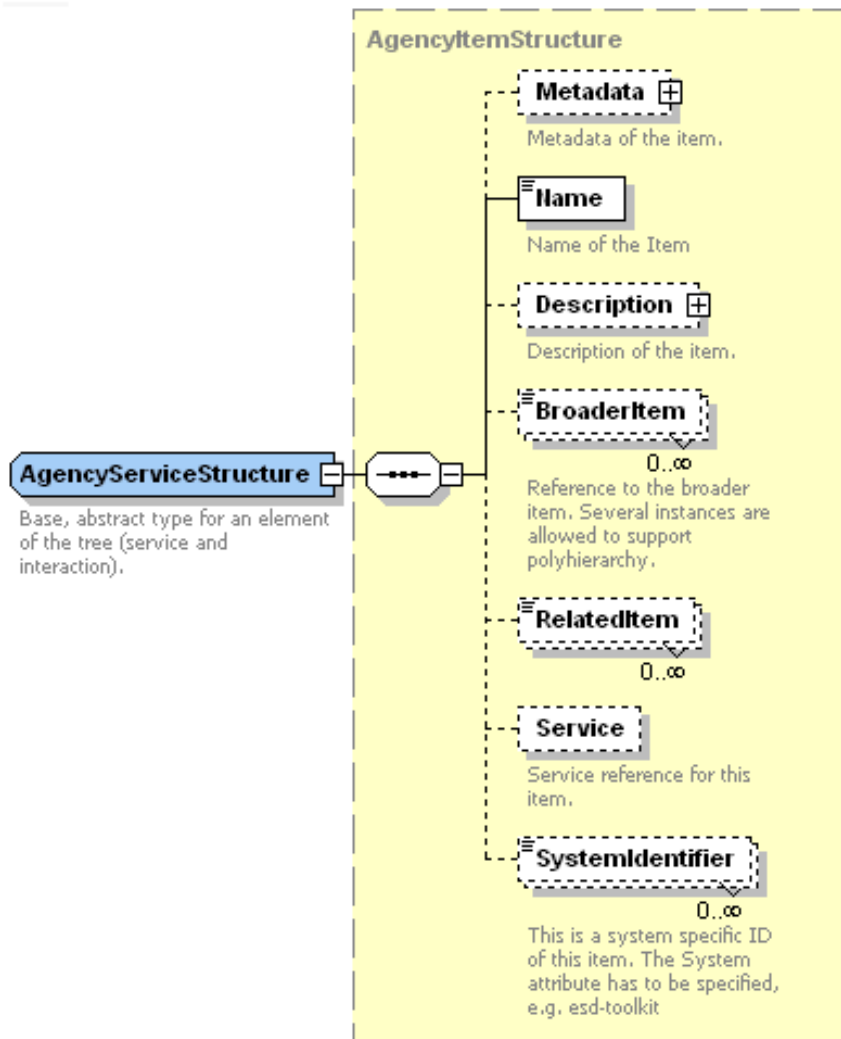
- BroaderItem – gives the service above the item in the hierarchical structure. The tag takes the name of the broader item. There are two attributes:
  - Id – the numeric id reference of the broader item
  - Default – a Boolean value used (as in hierarchical controlled lists) to allow rendering of the hierarchy of services by software that only supports mono-hierarchical lists. Hence, multiple broader items are permitted and one, but only one, must have a ‘Default’ value of true for each item

An item can have more than one BroaderItem if it exists twice within the hierarchical structure of services. In this way, a poly-hierarchical structure of services is supported. For example, a service to take out refuse for people with mobility problems might have broader items covering *both* waste management *and* services for people with disabilities.

- RelatedItem – gives a reference to another service item in the hierarchical structure. The tag takes the name of the related item and should have the same Id attribute as that item. Related items may be seen as representing ‘See also...’ references within a citizen interface.
- Service – gives an optional service reference. For the LocalDirectgov schema, service must come from LGSL.
- SystemIdentifier – any number of identification references to uniquely identify the item to a computer system, such as esd-toolkit. The System attribute, naming the computer system, is required. The SystemIdentifier allows data to be uploaded to the system identifying which existing data records require updating.

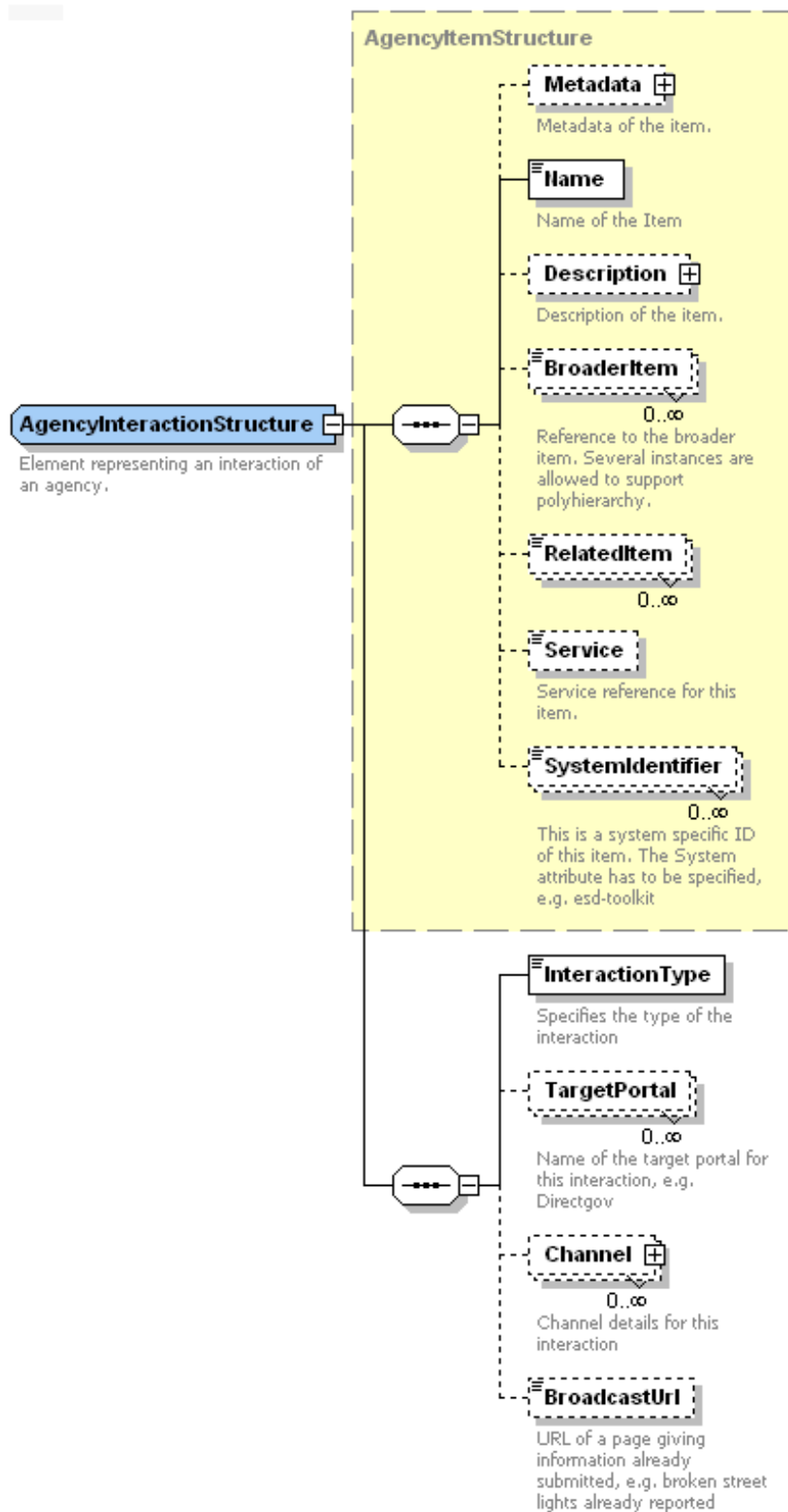
### 2.7.1 Service Elements

For services, the AgencyServiceStructure extends the AgencyItemStructure. The AgencyServiceStructure add no further elements in version 1.0 of the schema.



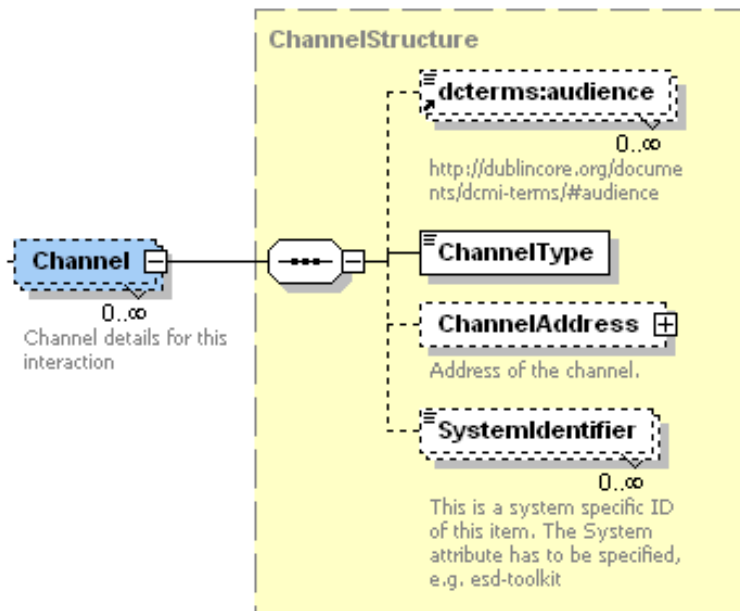
## 2.7.2 Interaction Elements

For interactions, the AgencyInteractionStructure restricts the AgencyItemStructure, making the Service element mandatory.



AgencyInteractionStructure adds these elements:

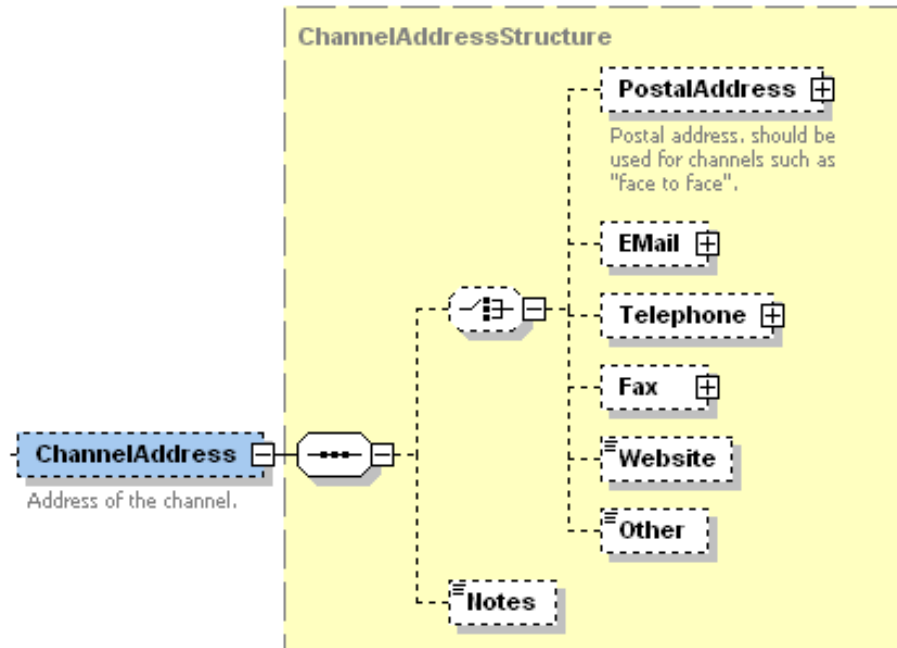
- Interaction type – gives the type of interaction. For the LocalDirectgov schema this must come from LGIL
- TargetPortal – gives the name of the portal an interaction is to be used for. For Local Directgov where more than one interaction of the same LGSL service and LGIL interaction type exists, the value of “Directgov” should be used against the one interaction to be used. Where there is no duplication of LGSL/LGIL combinations interactions will be used by Directgov irrespective of the TargetPortal value.
- Channel – gives the type and address of any channels through which the service or interaction is made available. An interaction may have no channels if the information within the XML data is sufficient for the user (citizen, business, etc) not to have to go further to complete the interaction (eg an interaction publishing information on domestic refuse collection days might provide that information within the XML data). However, the XML file is not intended to fully replace a content management system.



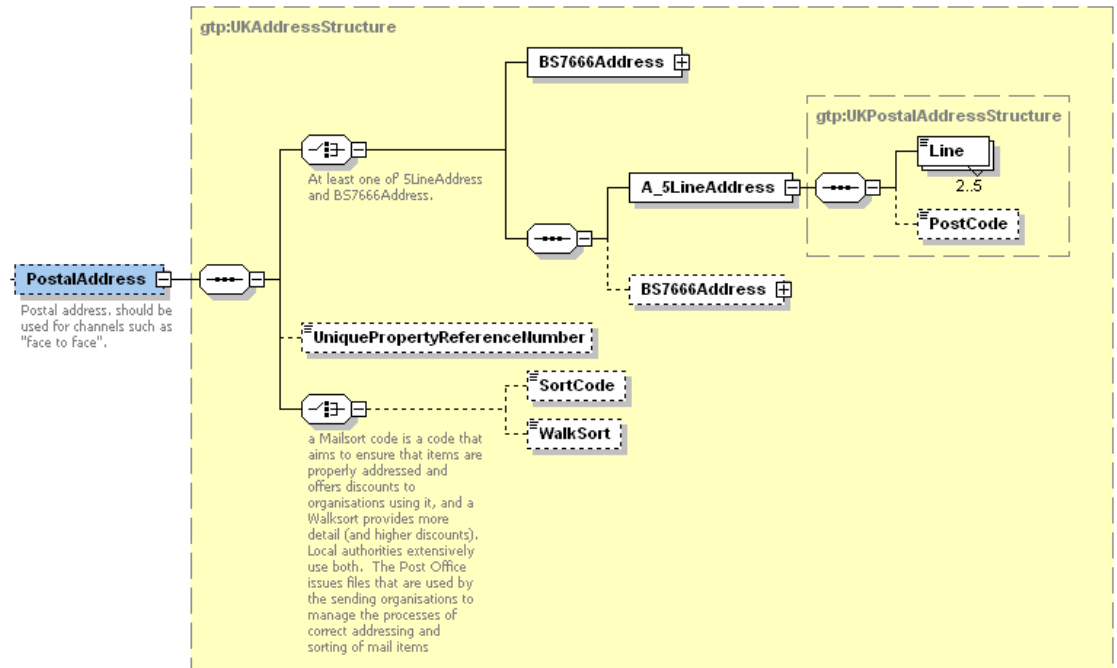
Channel is of type ChannelStructure and has these elements:

- dcterms:audience – optional audience for the channel
- ChannelType – This is of type ItemReferenceStructure, normally extended to type LGChL for a channel from the Local Government Channel List. For the LocalDirectgov schema, ChannelType must be populated from LGChL.

- ChannelAddress – with one (or none) of six elements and optional notes. A channel address may have notes but no address, for example if there is an explanation that a channel is not supported.

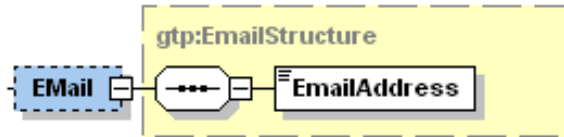


- PostalAddress of the UKAddressStructure (from the BS7666 address standard) complex type

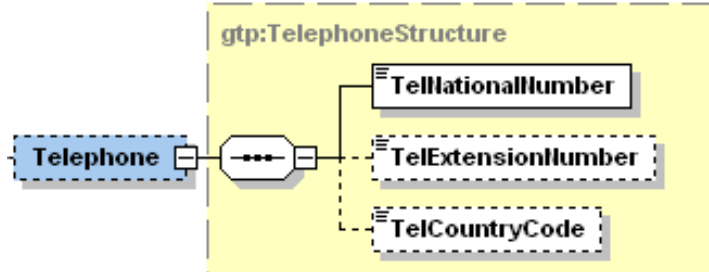


This supports either the full BS7666Address format or a simple address of two to five lines with an optional post code. It also supports the Unique Property Reference Number (UPRN) which should be used wherever available.

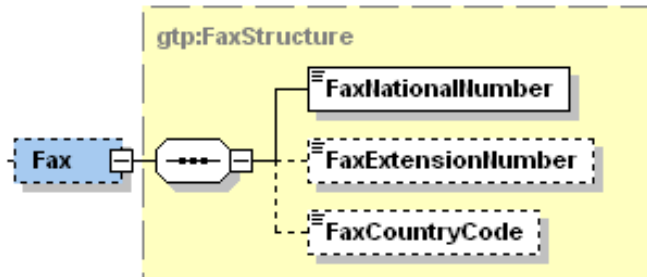
- Email of the govtalkpeople:EmailStructure type



- Telephone of the govtalkpeople:TelephoneStructure type



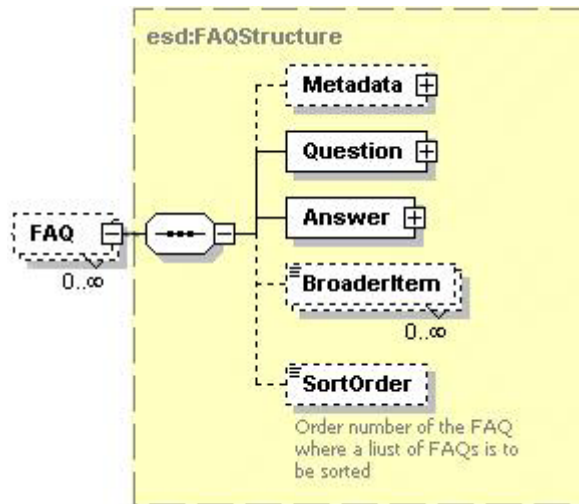
- Fax of the govtalkpeople:Fax type



- Website of type `xs:anyURI`
- Other – a simple string
- `BroadcastUrl` – The URL of a page giving information already submitted (eg broken street lights already reported) which the user may wish to view before continuing with an interaction.
- `SystemIdentifier` – any number of identification references to uniquely identify the item to a computer system, such as esd-toolkit. The `System` attribute, naming the computer system, is required. The `SystemIdentifier` allows data to be uploaded to the system identifying which existing data records require updating.

## 2.8 FAQ Elements

The XML will hold any number of FAQ items of complex type FAQStructure. The ‘LocalOnly’ Boolean attribute defines if the FAQ is only applicable to the agency’s delivery of services or is more broadly applicable.



These have the following elements:

- Metadata – the subject and audience metadata of the FAQ, supporting metadata elements of the same structure as the service/interaction metadata. This information allows FAQs to be grouped by content and associated with particular services. For example, the Department of Work and Pensions or the Local Authorities Coordinators of Regulatory Services (LACORS) may define FAQs that are applicable to LGSL services not delivered directly themselves, but referenced by LGSL subject metadata.
- Question – question text of the complex type DescriptionStructure, described in section 3.2 below
- Answer – answer text of the complex type DescriptionStructure, described in section 3.2 below
- BroaderItem – the Item reference(s) of the service(s) or interaction(s) to which the FAQ relates

## 3 Special Data Types Explained

### 3.1 Metadata Elements Populated From Controlled Lists

All metadata elements that *could* be populated from a controlled list are of type SimpleLiteralStructure, which is extended via the ItemReferenceStructure to a type specific to each controlled list.

The complex type “SimpleLiteralStructure” is an extension of “xs:string”. The complex type “ItemReferenceStructure” is an extension of “SimpleLiteralStructure” and includes an attribute “Id”.

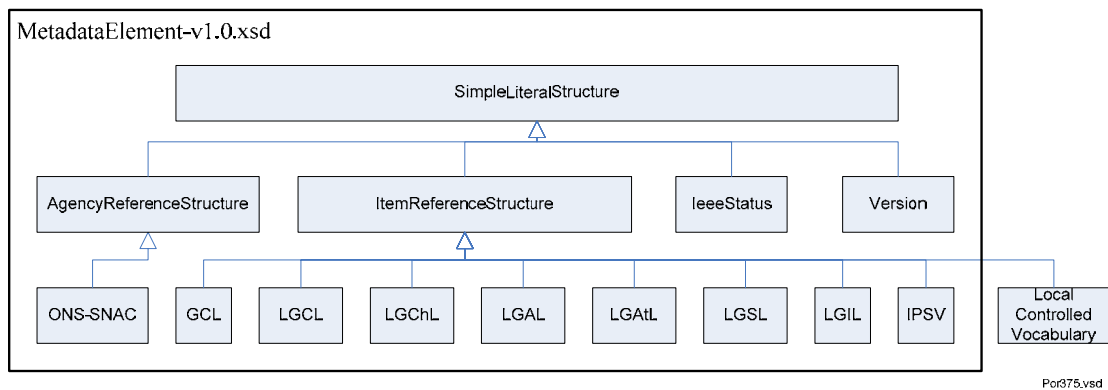
For example:

`<dc:subject>rubbish</dc:subject>` is an instance of the “SimpleLiteralStructure”.

`<dc:subject xsi:type="IPSV" Id="309">Waste disposal</dc:subject>` is an instance of the “ItemReferenceStructure”, by virtue of the `xsi:type` of “IPSV” extending the “SimpleLiteralStructure”.

When a scheme name has to be used, a complex type of the same name is declared extending the “SimpleLiteralStructure” (which may be extended for controlled list which do not have a numerical identifier) or “ItemReferenceStructure”.

The file “MetadataElement-v1.0.xsd” already contains the declaration of the main schemes used in the public sector, including IPSV, LGAL, LGAtL, LGChL and LGSL.



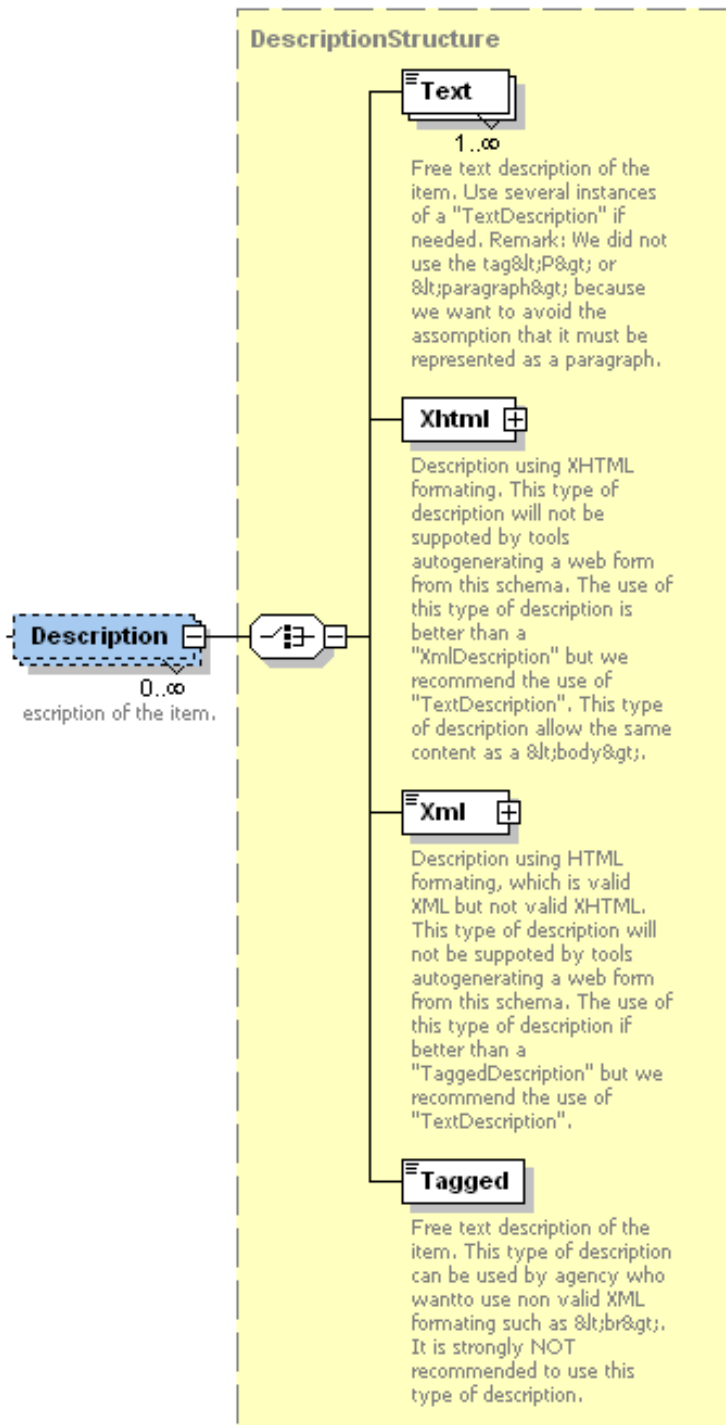
Note that the complex types for specific controlled lists (eg IPSV) do not have names ending in ‘Structure’ because such names might confuse people reading the file when used for `xsi:type` values in XML.

When an agency needs to use its own controlled list, it should create a schema containing the new Scheme declaration. See the example below.

```
<xs:include schemaLocation="LocalDirectgov -v1.3.xsd"/>
<xs:complexType name="CustomScheme">
  <xs:annotation>
    <xs:documentation>Scheme name...</xs:documentation>
  </xs:annotation>
  <xs:simpleContent>
    <xs:extension base="ItemReferenceStructure"/>
  </xs:simpleContent>
</xs:complexType>
```

### 3.2 Blocks Of Text That May Be Structured

Blocks of text are described by the complex type DescriptionStructure. This has four alternative (but exclusive) elements for defining text, given in order of desirability (most desirable first). The structure allows the XML file to use text structured by well or poorly defined rules if plain text is not available.



The four alternative elements are:

- Text - Pure text without structuring. Multiple 'Text' elements may be used for multiple blocks of text. Use of 'Text' for descriptions is recommended to support multiple output channels and leave control over structuring to display software, where accessibility issues may be managed
- Xhtml – Text formatted according to the W3C XHTML schema defining the structure of the Body element of an XHTML document. This type is provided where text is only available structured or loses meaning without structuring, but that structuring is XHTML compliant

- Xml – Text structured with well formed XML but not complying with XHTML. This type is provided where text is only available structured or loses meaning without structuring, but that structuring is XML (but not XHTML) compliant
- Tagged – Text including mark-up tagging which is not XML compliant. This type is provided for use where text is only available in a non-XML tagged format (eg simple HTML). HTML tags are represented as text with tags opened by “&lt” and closed by “&gt”.

#### 4 Validations Outside the Schema

The following validations cannot be performed by the schema:

- Check values for particular schemes (eg language).
- Top level items (ie ones with no BroaderItem) should have at least one IPSV term. \*
- Items of the AgencyInteractionStructure should all have at least one LGIL interaction type in their subject metadata. \*
- The metadata for the file must have at least one subject term from IPSV. \*
- Where items have one or more BroaderItems, one (and only one) should be set as the default. \*
- Appropriate complex types for controlled lists are used for specific metadata fields (eg LGAL is not used for Subject)
- Image size checks.

\* denotes validations that could be performed by a Schematron schema, which is currently not supplied by DirectLocalGov.

#### 5 Sample Data Conversions

Data conversions via XSLT from sample XML outputs from three Somerset local authorities are given at these addresses:

Source XML data	XSL Transformation	XML data compliant with LocalDirectgov .xsd
<a href="http://www.esd.org.uk/standards/LocalDirectgov/Examples/scc.xml">www.esd.org.uk/standards/LocalDirectgov/Examples/scc.xml</a>	<a href="http://www.esd.org.uk/standards/LocalDirectgov/Examples/SdcAsk_To_LocalDirectgov.xslt">http://www.esd.org.uk/standards/LocalDirectgov/Examples/SdcAsk_To_LocalDirectgov.xslt</a>	<a href="http://www.esd.org.uk/standards/LocalDirectgov/Examples/ScclocalDirectgov.xml">www.esd.org.uk/standards/LocalDirectgov/Examples/ScclocalDirectgov.xml</a>
<a href="http://www.esd.org.uk/standards/LocalDirectgov/Examples/SedgemoorDistrictCouncil.xml">www.esd.org.uk/standards/LocalDirectgov/Examples/SedgemoorDistrictCouncil.xml</a>		<a href="http://www.esd.org.uk/standards/LocalDirectgov/Examples/SedgemoorLocalDirectgov.xml">www.esd.org.uk/standards/LocalDirectgov/Examples/SedgemoorLocalDirectgov.xml</a>
<a href="http://www.esd.org.uk/standards/LocalDirectgov/Examples/TauntonDeane.xml">www.esd.org.uk/standards/LocalDirectgov/Examples/TauntonDeane.xml</a>		<a href="http://www.esd.org.uk/standards/LocalDirectgov/Examples/TauntonDeaneLocalDirectgov.xml">www.esd.org.uk/standards/LocalDirectgov/Examples/TauntonDeaneLocalDirectgov.xml</a>

The transformations use standard controlled list XML files available from [www.esd.org.uk/standards](http://www.esd.org.uk/standards) to derive Ids for subject metadata elements populated by controlled lists. They also use the LGCL to IPSV mapping available from that site to create IPSV subject metadata from source LGCL data.

