

Guidelines for creating 'Gigateway Approved' metadata

For metadata records on www.gigateway.org.uk to be 'Gigateway Approved', the Information Services Team will be regularly checking all metadata entries to ensure that they continue to reach a high level of data quality. The four steps to creating and publishing the improved metadata are as follows: -

1. Follow the guidelines given in Checklist 1 to help you create good quality metadata that conforms to the NGDF standard.
2. The abstract should adequately and concisely describe the data with enough detail to ensure that potential users can make an initial decision regarding its suitability. See Checklist 2 for the 11 points to consider including in the abstract. Example answers for raster, vector and documented metadata information are given to the right of the table and the compiled abstracts are shown at the end of this document.
3. All the fields in MetaGenie must be completed with valid, correctly spelled and accurate entries. More than one instance of certain elements is allowed. If multiple instances are desired then each instance should have a separate entry.
4. All the metadata records which pass the data quality checks will be known as 'Gigateway Approved' and will be posted on the gigateway website in the form of a spreadsheet which summarises each approved metadata record. The gigateway team will regularly review these approved records to ensure that they continue to meet all the criteria.

Checklist 1 – Checklist for creating valid metadata entries

Key: M = Mandatory, O = Optional, C = Conditional, 1 = One entry only allowed, N = One entry or more allowed

Field Title	Validation checks	Other
DATA DESCRIPTION		
Title	<ul style="list-style-type: none"> Title should encapsulate the subject, temporal and spatial coverage of the dataset. Should be short (aim to keep it under 100 characters). Should not contain terms or jargon to make it incomprehensible. If data has multiple titles then the name by which the dataset is most commonly known should be used (other discipline specific titles or titles with acronyms in can be entered as an Alternative Title) If no intelligible title exists the data producer must create one by which the dataset will be known in the public domain. <p>EXAMPLE: Voter Participation in Liverpool Local Elections, 1994</p>	M-1
Alternative Title	<ul style="list-style-type: none"> Titles translated into others languages can be entered here. Should be short (aim to keep it under 100 characters). Use this field for titles which have an acronym in common use, or where a discipline specific name exists. Enter 'not applicable' if no alternative title exists. 	O-N
Data originator	<ul style="list-style-type: none"> Should contain information about the person or body with prime responsibility for the intellectual content of the data. 'Data originator' is not necessarily the same person or organisation named as the contact for the data, or indeed, the person or organisation responsible for publishing the data. The originator(s) are not necessarily those who own the Intellectual Property Rights in the dataset. The general rule is that if the work is produced as a result of the policy of a corporate body and at the expense of that corporate body, the corporate body has prime responsibility, not the individual who actually produced it. Insert an abbreviation for the organisation in brackets if appropriate. <p>EXAMPLES: The Royal Society for the Protection of Birds (RSPB) <i>or</i> Environment Agency (EA)</p>	O-N
Abstract	<ul style="list-style-type: none"> Must concisely describe the content and purpose of dataset. Aim to write under 100 words ideally, but certainly no more than 300 words as a guide. Make sure you have followed all the points mentioned in Checklist 2. <p>EXAMPLES: See abstract examples for raster, vector and documented information at the end of these guidelines.</p>	M-1
Gigateway keywords	<ul style="list-style-type: none"> Ensure that at least one relevant gigateway keyword has been selected. <p>VALID ENTRIES: Agriculture/Farming, Aquaculture/Fishery, Base Maps, Biota, Cadastral & legal, Climatology & Atmosphere, Communications, Earth Cover/Land Use, Economy, Elevation Derived Products, Environment, Geoscientific Information, Health, Imagery, Infrastructure, Inland waters, Military Infrastructure, Oceans, Planning, Political Boundaries, Society, Utilities</p>	M-N
DATA CAPTURE PERIOD		
Status of Start	<ul style="list-style-type: none"> This date relates to when the information was collected, NOT when the 	M-1

Date	<p>data was digitised or converted to a particular format.</p> <ul style="list-style-type: none"> • Only select 'not applicable' if the data has been continuously updated for a number of years and does not contain an historic record. • If the start date is known then enter the word 'known'. If 'known' is selected the 'Start Date of Data Capture' must also be completed. • If 'not known' is entered the 'Start date of data capture' should be left blank. <p>VALID ENTRIES: known, not known, not applicable</p>	
Start date of data capture	<ul style="list-style-type: none"> • This is the date that the dataset was first collected and is conditional on the entry to the 'Status of start date' being 'known'. • Format of this element is CCYY-MM-DD, i.e. 2004-09-29 • Where any part of the date is not known '01' should be used instead. For example, if only the year and month were provided then the entry should be 2004-09-01 for example. • The 'Start date of data capture' should come before the 'End date of data capture' • 'Start date of data capture' should not be in the future. <p>EXAMPLE: 2004-09-29 (to represent the 29th September 2004) or 2004-09-01 to represent September 2004 (Notice that '01' is used here to show that the date is unknown)</p>	C-1
Status of End Date	<ul style="list-style-type: none"> • This date relates to when the data collection finished, NOT when the data was finally digitised or converted to a particular format. • Only select 'not applicable' if the data has been continuously updated for a number of years and does not contain an historic record. • If the start date is known then enter the word 'known'. If 'known' is selected the 'End Date of Data Capture' must also be completed. <p>VALID ENTRIES: known, not known, not applicable, ongoing</p>	M-1
End Date of Data Capture	<ul style="list-style-type: none"> • This is the date on which data collection ended and is conditional on the entry to the 'Status of End Date' being 'known'. • Format of this element is CCYY-MM-DD, i.e. 2004-09-29 • Where any part of the date is not known '01' should be used instead. For example, if only the year and month were provided then the entry should be 2004-09-01 for example. • 'End date of data capture' should not be in the future. <p>EXAMPLE: 2004-09-29 (to represent the 29th September 2004) or 2004-09-01 to represent September 2004 (Notice that '01' is used here to show that the date is unknown)</p>	C-1
Frequency of Data Update	<ul style="list-style-type: none"> • If data capture is continuous then the period of update should be given. • Enter 'continuous' if the dataset is updated continuously and the user would obtain the data in whatever state it happened to be on the date of purchase • Enter 'irregular' if the data continues to be updated but there is no regular period of update. <p>VALID ENTRIES: hourly, daily, weekly, fortnightly, monthly, biannually, annually, biennially, triennially, quinquennially, decennially, continuous, irregular, never, other</p>	M-1
Date of update of metadata	<ul style="list-style-type: none"> • This mandatory element can be used to state when the discovery metadata itself was last updated. • Format of this element is CCYY-MM-DD, i.e. 2004-09-29 • Where any part of the date is not known '01' should be used instead. For example, if only the year and month were provided then the entry 	M-1

	<p>should be 2004-09-01 for example.</p> <ul style="list-style-type: none"> • 'Date of update of metadata' should not be in the future. • To remain as a 'Gigateway Approved' record the metadata will need to be updated at least once a year. This also applies to static datasets which will need to be checked to see that all the fields are still current. <p>EXAMPLE: 2004-09-29 (to represent the 29th September 2004) or 2004-09-01 to represent September 2004 (Notice that '01' is used here to show that the date is unknown)</p>	
DATA GEOGRAPHIC EXTENT		
National Extent	<ul style="list-style-type: none"> • This element is conditional on geographic identifiers rather than geographic coordinates being used to define the spatial extent of the data. • Where the data extent covers more than one nation then multiple entries are allowed. <p>VALID EXAMPLES: England, Northern Ireland, Scotland, Wales, Isle of Man, Channel Islands, Great Britain, Offshore, United Kingdom</p>	M-N
Administration Area Extent	<ul style="list-style-type: none"> • The use of this element is conditional on geographic identifiers rather than geographic coordinates being used to define the spatial extent of the data. • Where the data extent covers more than one administrative area then multiple entries are allowed. <p>ADMINISTRATIVE AREA EXAMPLES: Islington London Borough Council, Cornwall County Council,</p>	O-N
Post Code Extent	<ul style="list-style-type: none"> • The use of this element is conditional on geographic identifiers rather than geographic coordinates being used to define the spatial extent of the data. • Valid entries are postcode areas (e.g. CB), districts (e.g. CB1), sectors (e.g. CB1 2) and also postcode units (e.g. CB1 2AD) • Where the data extent covers more than one postcode then multiple entries are allowed. • Remember to use a space in the postcode if appropriate. <p>POSTCODE EXAMPLES: CB, CB1, CB1 2 or CB1 2AD</p>	O-N
Spatial Reference System (to define extent of dataset)	<ul style="list-style-type: none"> • To indicate the geographic extent of the data then the name of the system of spatial referencing should be selected from the list below. • More than one instance is allowed for datasets with more than one contiguous footprint. • The same extents can be repeated using different spatial reference systems. <p>EXAMPLES: National Grid of Great Britain, Irish Grid or Latitude and Longitude</p>	M-N
North bounding coordinate	<ul style="list-style-type: none"> • Enter the northern most coordinate of the limit of the dataset extent using coordinates from the National Grid of Great Britain, Irish Grid or Latitude and Longitude. • Use ShowBox (http://www.gigateway.co.uk/mapserver/mapui_gi/mg_showbox.pl) to find out the Latitude/Longitude or grid references for the National Grid of Great Britain. <p>Latitude and Longitude: the valid ranges are +180 degrees to -180 degrees (for longitude) and +90 degrees to -90 degrees (for latitude) National Grid of Great Britain: Grid references as used by Ordnance</p>	M-N

	<p>Survey GB Irish Grid: Grid references as used by Ordnance Survey Ireland</p>	
West bounding coordinate	<ul style="list-style-type: none"> Enter the western most coordinate of the limit of the dataset extent using coordinates from the National Grid of Great Britain, Irish Grid or Latitude and Longitude. <p>Latitude and Longitude: the valid ranges are +180 degrees to -180 degrees (for longitude) and +90 degrees to -90 degrees (for latitude) National Grid of Great Britain: Grid references as used by Ordnance Survey GB Irish Grid: Grid references as used by Ordnance Survey Ireland</p>	M-N
South bounding coordinate	<ul style="list-style-type: none"> Enter the southern most coordinate of the limit of the dataset extent using coordinates from the National Grid of Great Britain, Irish Grid or Latitude and Longitude. <p>Latitude and Longitude: the valid ranges are +180 degrees to -180 degrees (for longitude) and +90 degrees to -90 degrees (for latitude) National Grid of Great Britain: Grid references as used by Ordnance Survey GB Irish Grid: Grid references as used by Ordnance Survey Ireland</p>	M-N
East bounding coordinate	<ul style="list-style-type: none"> Enter the eastern most coordinate of the limit of the dataset extent using coordinates from the National Grid of Great Britain, Irish Grid or Latitude and Longitude. <p>Latitude and Longitude: the valid ranges are +180 degrees to -180 degrees (for longitude) and +90 degrees to -90 degrees (for latitude) National Grid of Great Britain: Grid references as used by Ordnance Survey GB Irish Grid: Grid references as used by Ordnance Survey Ireland</p>	M-N
DATA ACCESS		
Access constraints	<ul style="list-style-type: none"> This defines any restrictions placed on access to the data. A financial constraint would imply that the data had to be purchased. A legal constraint implies that some security or contractual constraint applies to the dataset. You may like to specify here who are/are not allowed to access the data. <p>VALID EXAMPLES: financial, legal, other, not known, none</p>	M-N
Use constraints	<ul style="list-style-type: none"> When copyright or other legal restrictions protect the data this element is used to define the terms of the constraint including re-use fees, reproduction restrictions and acknowledgements. Do not provide the actual fees payable here, just the fact that a fee would be payable is enough. If the use constraints are unknown then enter 'not known'. 	O-1
Level of Spatial Detail	<ul style="list-style-type: none"> This element should give some indication of the resolution, the degree of disaggregation or generalisation of the data. The nature of the information provided will depend on the type of data. Where the level of Spatial Resolution is unknown enter 'not known'. <p>For statistical data the smallest unit to which data is disaggregated could be used - EXAMPLES: enumeration district, postcode, county</p> <p>For imagery the pixel size expressed in the equivalent ground measurement could be used – EXAMPLE: 10 metre square pixels</p>	O-1

	<p>For data based upon topographical mapping nominal or capture scale could be used – EXAMPLES: 1:25000, 1:50,000</p> <p>For data captured using a GPS receiver you need to state the accuracy after post-processing – EXAMPLE: 2-5 metres, <1 metre</p>	
Supply media	<ul style="list-style-type: none"> This element is used to define the format of the media in which the data can be provided <p>EXAMPLES: Paper (data can be supplied in printed form), Magnetic (floppy disc, DAT tape), Online (data can be downloaded electronically), Optical (CD-ROM), Other</p>	M-N
Data format	<ul style="list-style-type: none"> This can help to identify which application to use the data in. Where relevant, additional information concerning the format of the data may be provided, such as the version or level of NTF. <p>EXAMPLES: NTF, DXF, Microsoft Word, Microsoft Excel, Other, not known, not applicable (i.e. only if supplied on paper)</p>	O-N
Presentation type	<ul style="list-style-type: none"> This is the mode in which the data is represented. <p>VALID ENTRIES: <i>Image</i> – Photographic or satellite image <i>Graphic</i> – Drawing <i>Map</i> – Data presented in map form <i>Numeric</i> – Set of numbers or statistical data <i>Text</i> – Free text</p>	M-N
Spatial Reference System (within dataset)	<ul style="list-style-type: none"> This element is used to define the spatial referencing system used within the dataset. It should not be confused with the spatial referencing system used to define the spatial extent of the dataset. The spatial referencing system may be based on coordinates e.g. National Grid of Great Britain, Irish Grid, latitude and longitude or geographic identifiers e.g. postcodes, postal addresses, administrative units or countries. If the Spatial Reference System is unknown then enter 'unknown'. <p>VALID ENTRIES: National Grid of Great Britain, Irish Grid, Latitude and Longitude, postcode, parish, ward, electoral constituency, census area, local authority (county/unitary/district/borough), health authority area, travel-to-work area, other area type.</p>	M-N
MISCELLANEOUS		
Additional Source Information	<ul style="list-style-type: none"> Use this to provide an indication of where additional information about the dataset can be found. You should not use this element as a place to add additional information about the dataset. Where there are no additional sources of data then enter 'None'. <p>EXAMPLES: contact address, the dataset handbook, a research paper on the dataset, a URL, etc.</p>	O-N
Dataset Associations	<ul style="list-style-type: none"> Where similar or related datasets are available then they can be referenced in this element. A related dataset may be one where data for the same geographic extent exists but for a different time frame. Alternatively, an identical dataset may exist for an adjacent geographic area. The information is meant as a pointer to related datasets and not as a description of such datasets and therefore should be kept as short as possible. 	O-N

	<ul style="list-style-type: none"> Where there are no associated datasets enter 'None'. 	
Samples	<ul style="list-style-type: none"> This element can be used to provide a sample of the data to the enquirer. Where the dataset is graphical, i.e. a map, image or graphic then a sample of the dataset can be provided here in the form of an image file. If the dataset were statistical then a tabular sample of the data might be provided. 	O-N
DATA SUPPLIER DETAILS		
Contact Name or Title	<ul style="list-style-type: none"> In this element the name or preferably the job title of the person who answers to data enquiries and requests should be inserted here. By inserting the job title of the person it avoids having to update the metadata record so often if people change roles in an organisation. Ensure that the named person agrees to their name being included. <p>EXAMPLES: GIS Officer, Information Manager, Mr Alan Brindley, etc.</p>	M-1
Full Postal Address	<ul style="list-style-type: none"> The postal address for the contact organisation to be inserted. The address should contain: Organisation name, Building name, number, street name, Postal town or city, Country. If a postal address is not provided then enter 'not applicable' 	M-1
Post code	<ul style="list-style-type: none"> Enter the postcode (with a space) or zip code in this field. If a postal address is not provided then enter 'not applicable'. <p>EXAMPLE: EC2A 2BB</p>	M-1
Telephone Number	<ul style="list-style-type: none"> Insert a telephone number for the contact person. The number given should be the full international dialling number. If a telephone number is not provided then enter 'not applicable'. 	O-1
Facsimile Number	<ul style="list-style-type: none"> The facsimile number for the contact. The number given should be the full international dialling number. If a facsimile number is not provided then enter 'not applicable'. 	O-1
Email	<ul style="list-style-type: none"> The email address for the contact. If an email address is not provided then enter 'not applicable'. 	O-1
Website Address	<ul style="list-style-type: none"> Where the contact organisation has a World Wide Web site which can be used to obtain further information about the dataset or from which the data can be purchased then the URL should be provided here. If a web address is not provided then enter 'not applicable'. 	O-1
USER DEFINED FIELDS		
User Defined Value: Name	<ul style="list-style-type: none"> Use this element, in combination with the next element, to add your own extra metadata details. This element is the name part of the user defined name value pair. <p>EXAMPLE: Location of file on server</p>	O-N
User Defined Value: Value	<ul style="list-style-type: none"> This element is the value part of the user defined name value pair. If the previous 'Name' element is completed then an entry for this element is mandatory, otherwise it can be left blank. <p>EXAMPLE: \\ebcgis\esrigis\working_data\Administrative_boundary\Elmbridge_wards.bds</p>	O-N

Checklist 2 – Checklist, with examples, for what to include in the abstract

	Abstract guideline	Example answers (R = Raster, V = Vector, D = Document)		Check
1	What does the data essentially show?	R	<i>Landsat-7 Enhanced Thematic Mapper Plus Imagery</i>	Yes/No
		V	<i>Potentially Contaminated Land</i>	
		D	<i>UK Newspaper Reports of Earthquakes</i>	
2	What area does the data cover? <i>e.g. district council, county, country, town, etc.</i>	R	<i>Worldwide</i>	Yes/No
		V	<i>Billington Borough Council</i>	
		D	<i>Mainly earthquakes in Great Britain</i>	
3	Period of time the data covers and frequency of update? <i>with month and year for period of time</i>	R	<i>August 1998 to July 2001, every 16 days</i>	Yes/No
		V	<i>January 1983 to January 2005, annually</i>	
		D	<i>January 1703 to August 1997, irregular</i>	
4	What scale was the data captured at or what is the resolution of the data?	R	<i>Bands 1 to 5 and 7 at 30m resolution, Band 6 at 60m resolution, Panchromatic band at 15m resolution</i>	Yes/No
		V	<i>1:2500</i>	
		D	<i>N/A</i>	
5	Data capture method? <i>e.g. GPS survey, aerial photography, satellite imagery, digitised from other mapping (please state the source.....), other method (please state.....)</i>	R	<i>Satellite imagery - Landsat-7 Enhanced Thematic Mapper Plus (Landsat-7 ETM+)</i>	Yes/No
		V	<i>Digitised from OS MasterMap</i>	
		D	<i>N/A</i>	
6	Suggested uses for data? <i>e.g. trend analysis, planning, routing and logistics information, environmental analysis, land use mapping, other (please state.....)</i>	R	<i>Global change research (desertification, agricultural development, water resource monitoring, urbanisation, deforestation, pollution, cataclysmic volcanic activity, etc.)</i>	Yes/No
		V	<i>Land use mapping and environmental analysis</i>	
		D	<i>Report writing</i>	
7	Tick all the features that can be identified with the data? <i>e.g. boundaries, roads, railways, land features, water features, historical, other (please state.....)</i>	R	<i>Landuse (e.g. water, vegetation, urban areas, deserts, etc.)</i>	Yes/No
		V	<i>garages, spoil heaps, train stations and factories</i>	
		D	<i>N/A</i>	
8	Are there any limitations in the data? <i>e.g. some data missing (please state...), some data out of date (please state.....),</i>	R	-	Yes/No
		V	<i>Data missing for Jan 2001, 2002 data incomplete.</i>	

	<i>some attribute data missing (please state.....), data in paper form, other (please state.....)</i>	D	<i>Reports are mainly for earthquakes over 5 on the Richter scale.</i>	
9	Does the data have any key attribute data that could be used for linking in other datasets? <i>e.g. TOID, UPRN, USRN, other (please state.....)</i>	R	-	Yes/No
		V	<i>TOIDs</i>	
		D	<i>TOIDs and UPRNs could quite legitimately be used in the text (but in this case they weren't)</i>	
10	If you are not the originator of the data, please state who is and whether you made any changes to the data before you supplied it to us?	R	<i>Data originally from U.S. Geological Survey (USGS) Earth Resources, Observation & Science (EROS) Data Center and purchased for research by the Natural Environment Research Council (NERC) Earth Observation Data Centre (NEODC).</i>	Yes/No
		V	<i>Data originally supplied by Billington County Council in 1983 and Billington Borough Council has been updating the overlay since then.</i>	
		D	<i>The Independent, The Times, etc.</i>	
11	Is the data in vector, raster or another format? <i>e.g. vector, raster, tabular, document, other (please state.....).</i>	R	<i>Raster</i>	Yes/No
		V	<i>Vector</i>	
		D	<i>Document</i>	

Example extracts for vector, raster and documented information (using Checklist 2 as a guide)

VECTOR DATA

Potentially Contaminated Land data for Billington Borough Council from January 1983 to January 2005, captured annually on 1:2500 scale and digitised using OS MasterMap as the basemap. Suggested uses for this polygon data are land use mapping and environmental analysis. These are 'potentially' contaminated sites as they highlight all the garages, spoil heaps, train stations and factories which are possible sources of pollution and therefore may not have contaminated the environment at all. The limitations are that data for 2001 is missing and 2002 is incomplete. There are legal constraints too. Data was originally supplied to Billington Borough Council from Billington County Council in 1983 and Billington Borough Council have been updating the data annually since then. TOIDs are incorporated with the attribute information for all the polygons relating to buildings. Desktop GIS use is recommended for the vector data which can be supplied in .mid/.mif format.

RASTER DATA

Imagery scenes showing worldwide coverage of landuse captured with Landsat-7 Enhanced Thematic Mapper Plus (Landsat-7 ETM+) and containing spectral bands 1 through 8. Raster data has been updated every 16 days since August 1998 to July 2001. Bands 1 to 5 and 7 at 30m resolution, band 6 at 60m resolution and panchromatic band at 15m resolution. The consistency of the data in terms of data acquisition, geometry, spatial resolution, calibration, coverage and spectral characteristics makes it ideal for global change research concerning desertification, agricultural development, water resource monitoring, urbanisation, deforestation, pollution, cataclysmic volcanic activity and other natural and anthropogenic events. No charge for use in non-commercial academic research, but a charge is made for commercial use. Data originally from U.S. Geological Survey (USGS) Earth Resources, Observation & Science (EROS) Data Center and purchased for research by the Natural Environment Research Council (NERC) Earth Observation Data Centre (NEODC). Desktop GIS use recommended for the raster data which can be supplied in GeoTiff or FastL7A format.

DOCUMENTED INFORMATION

'UK Newspaper Reports of Earthquakes' is a collection of extracts from British newspapers describing earthquakes (mostly British). Collection began in Jan 1703 and update is ongoing but irregular. Original clippings, photocopies, and in a few cases, manual transcriptions are available.