

DISCOVERING METADATA WITH ASTUN TECHNOLOGY

Presented by Jo Cook with Eoghan Hennessy

We are **Astun Technology**, and we create world-class
cloud-hosted geospatial solutions on open source
foundations

Today we'll be talking about **data discovery and
metadata**

Hi, I'm Jo Cook



I'm the Data Discovery Lead at Astun Technology

Eoghan Hennessy is co-hosting



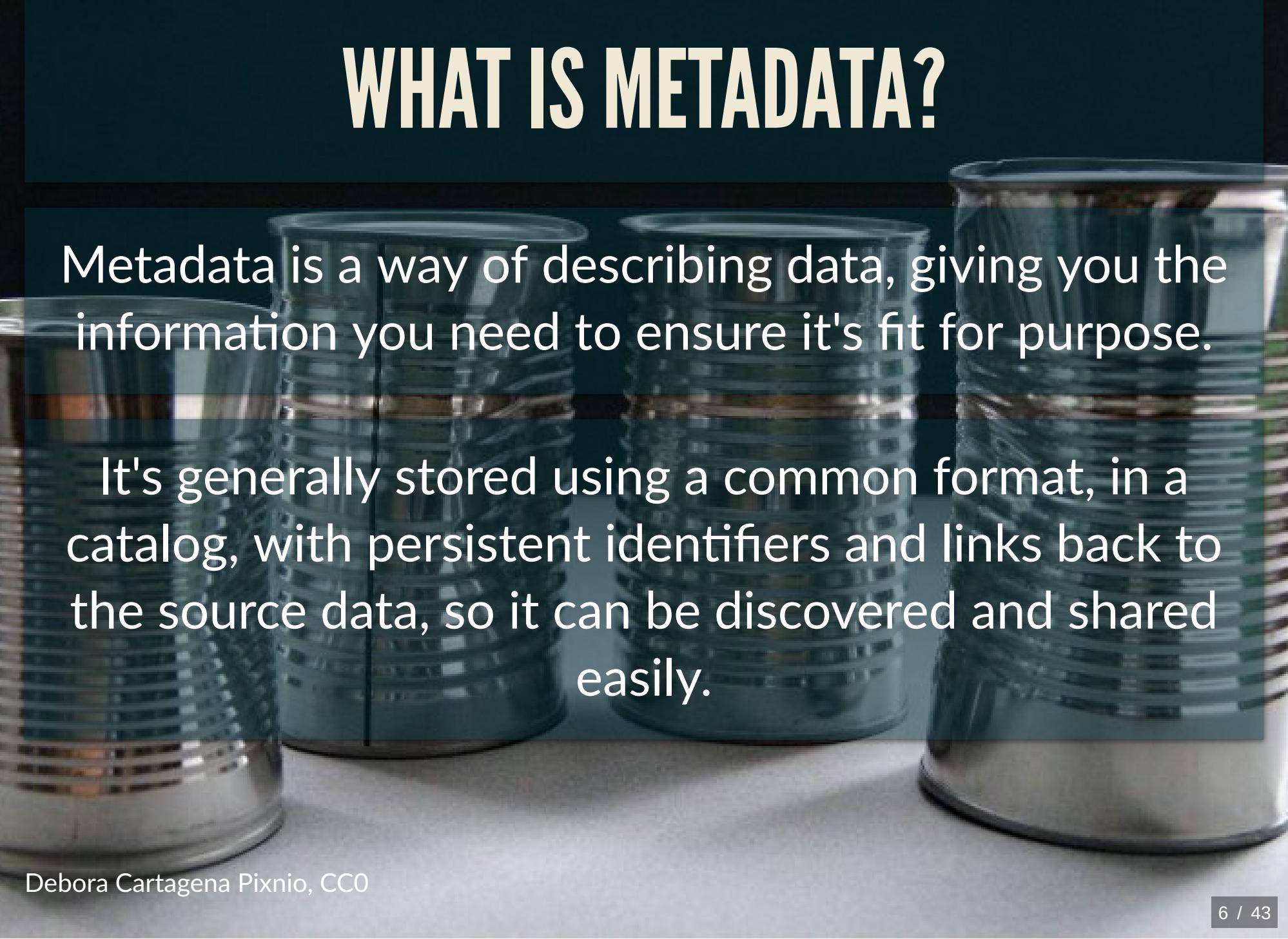
He's a Project Manager at Astun with responsibility for iShare in the Cloud and Metadata Enterprise Solutions

SOME BACKGROUND

Astun Technology manage and support metadata catalogs for Scottish Government, the Environment Agency, DEFRA, and many regional authorities

Our work includes developing the plugin to allow GeoNetwork to support the Gemini 2.3 metadata standard, and we are Project Steering Group members and committers to both code and documentation.

WHAT IS METADATA?

A photograph showing a stack of several large, cylindrical metal barrels or drums. They are stacked in a staggered pattern, creating a sense of depth. The barrels have a metallic texture and some visible rust or wear. The background is dark, suggesting an indoor warehouse or storage facility.

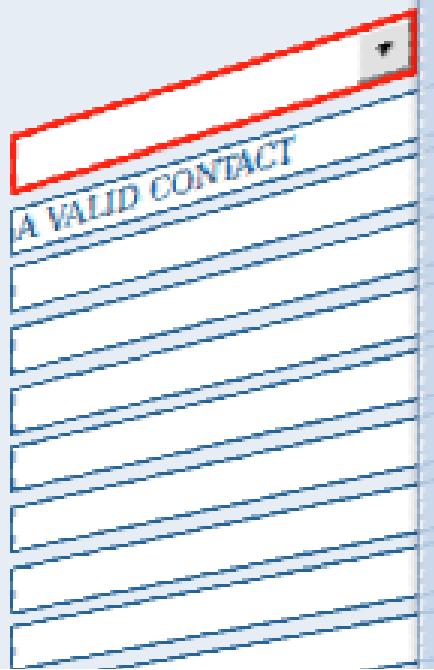
Metadata is a way of describing data, giving you the information you need to ensure it's fit for purpose.

It's generally stored using a common format, in a catalog, with persistent identifiers and links back to the source data, so it can be discovered and shared easily.

BUT...

Metadata is often viewed as a legislative box-ticking exercise that is hard, costs time and money, and seems to have little direct benefit.

cough, INSPIRE, cough



Compliance to metadata standard (XML Schema validation report)

- cvc-datatype-valid.1.2.1: " is not a valid value for 'gmd:dimensions'.
- cvc-type.3.1.3: The value " of element 'gco:Integer' is not valid. (Element: gmd:numberOfDimensions)
- cvc-type.3.1.3: The value " of element 'gco:Integer' is not valid. (Element: gmd:enumeration-valid)
- cvc-enumeration-valid: Value " is not facet-valid with respect to enumeration [center, lowerLeft, lowerRight, upperRight, upperLeft]. It must be a value from the enumeration. (Element: gmd:MD_PixelOrientationCode with parent element: gmd:pointInPixel)
- cvc-type.3.1.3: The value " of element 'gmd:MD_PixelOrientationCode' is not valid. (Element: gmd:MD_PixelOrientationCode with parent element: gmd:pointInPixel)
- cvc-datatype-valid.1.2.1: " is not a valid value for 'double'. (Element: gco:Real with parent element: gmd:minValue)
- cvc-type.3.1.3: The value " of element 'gco:Real' is not valid. (Element: gco:Real with parent element: gmd:maxValue)
- cvc-datatype-valid.1.2.1: " is not a valid value for 'double'. (Element: gco:Real with parent element: gmd:density)

#5:ERROR-XPath: /gmd:dimension/gmd:MD_

Metadata recommendations (Schematron)

TIMES HAVE CHANGED!



We've gone beyond creating data and metadata for internal use. Now, data is shared on corporate intranets, on public-facing portals, and in national data catalogs, and efficient use of data is crucial in solving many of today's crises...

Covid: how Excel may have caused loss of 16,000 test results in England

Public Health England data error blamed on limitations of Microsoft spreadsheet

Astun have enhanced and expanded our metadata offering to meet these new challenges, and to help you meet emerging best practice guidelines for data management recently issued by the Geospatial Commission

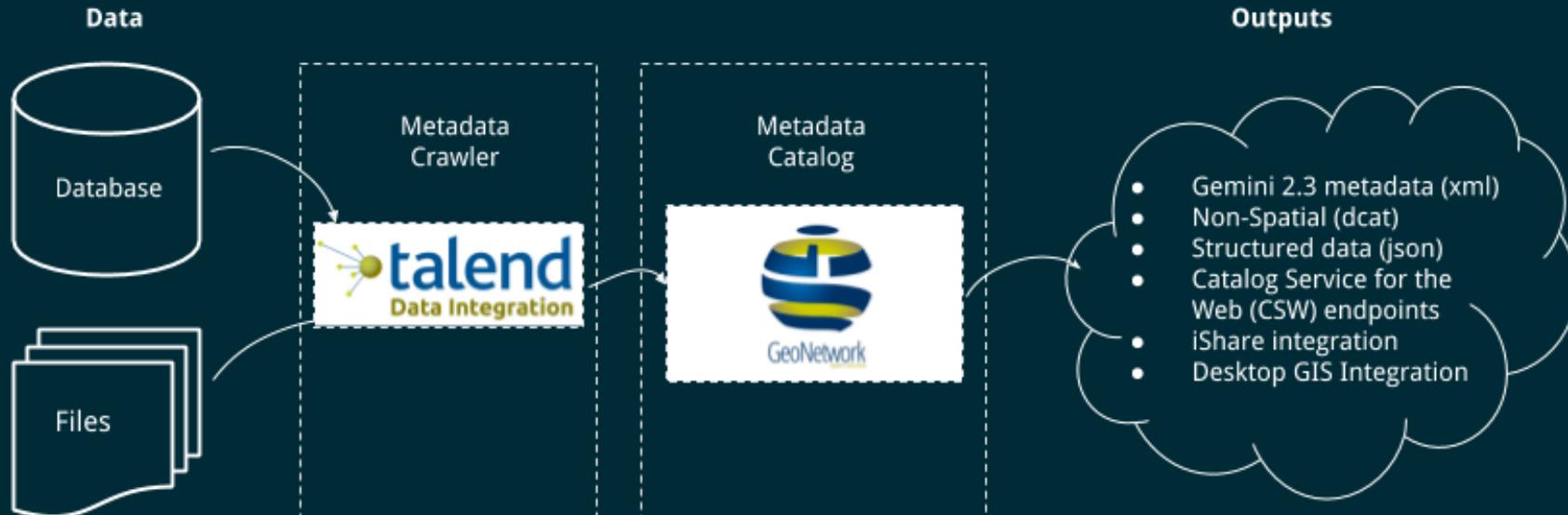
Search engine optimisation
(SEO) for data publishers:
Best practice guide

EXPANDED PRODUCT

We use GeoNetwork, the leading open source geospatial metadata catalog, as the core of our product. We've now added Metadata Crawler, also open source, based on Talend ETL, for bulk metadata creation.

We've also expanded the functionality of GeoNetwork, which we'll talk more about shortly

OVERVIEW



BEST PRACTICE

Geospatial Commission Best Practice Guides state that data should be Findable, Accessible, Interoperable, and Reusable (FAIR)



FINDABLE

Searches start in search engines, not in portals, so we need to make sure our datasets are search-engine-optimised for both good rankings and well-displayed results

We've added structured data into Gemini 2.3,
Created reports showing records with data quality
issues,
Improved GeoNetwork SEO

[Back to search](#)< PreviousNext >[Download](#) ▾[Display mode](#) ▾

Dedicated Woodlands

Description:

The Dedication Scheme (Basis I & II) was introduced in 1947 in order to encourage landowners to retain their land in forestry and to introduce good forestry practice. Basis III was introduced in 1974, providing grants for new planting and additional supplements for broadleaves.

The Dedication Scheme was closed to new applications in 1981. Land still under Dedication could continue to be within the scheme but Dedication would terminate on a change of ownership.

Dedication schemes without a Plan of Operations and therefore receiving no grant, are deemed to be under a Negative Covenant.

Dataset Attributes:

Descriptor Dataset name
Case_Name Dedication Scheme name
Case_No Dedication Scheme reference number
Basis Basis number of scheme (I, II or III)
Date_Appr Date woodland became Dedicated
PlanOfOps Period of current (or last) Plan of Operations
Covenant Scheme in 'Positive' or 'Negative' Covenant
Grid_Ref Grid Reference of property
Total_Area Total area of scheme (* OL not always included)
Man_Area Total area of scheme under management

Download and links



Scottish Forestry Open Data

Data download site <https://open-data.gov.scot/datasets/dedicated-woodlands>[Open link](#)

Spatial extent



#1 RECORD IN GEONETWORK

Temporal extent

Creation date

2002-09-01

Revision date

2019-11-01

Data sets



Dedicated Woodlands

1 warning



Missing field 'creator' (optional)

type

Dataset

id

<https://spatialdata.gov.scot/geonetwork/srv/api/records/4268B281-F173-4B7E-B2AE-1583C659253C>

includedInDataCatalog

type

DataCatalog

url

<https://spatialdata.gov.scot/geonetwork/srv/search#>

#2 STRUCTURED DATA

inLanguage

eng

name

Dedicated Woodlands

alternateName

FCPRODUCT.S_DEDICATION

description

Description: The Dedication Scheme (Basis I & II) was introduced in 1947 in order to encourage landowners to retain their land in forestry and to introduce good forestry practice. Basis III was introduced in 1974, providing grants for new planting and additional supplements for broadleaves. The Dedication Scheme

Google

Dedicated Woodlands

Last updated Download format Usage rights Topic Free Saved data sets

21 data sets found

D **Dedicated Woodlands**
data.gov.uk
spatialdata.gov.scot
Updated Jul 23, 2020

D **Woodland Dedication**
data.gov.uk
www.europeandataportal.eu
Updated Aug 25, 2020

G **Open Government Licence for Public Sector Information**
lle.gov.wales
wggt-lle-landing.azurewebsites.net
zip
Updated Jun 26, 2011

Forestry Commission **Dedication Woodland Grant Scheme England**
data.gov.uk
www.europeandataportal.eu
+2more
Updated Aug 28, 2020

Dedicated Woodlands

Explore at data.gov.uk Explore at spatialdata.gov.scot

16 scholarly articles cite this data set (View in Google Scholar)

Data set updated Jul 23, 2020

Data set provided by Scottish Government SpatialData.gov.scot

Licence Open Government Licence 3.0
Licence information was derived automatically

Description

Description:
The Dedication Scheme (Basis I & II) was introduced in 1947 in order to encourage landowners to retain their land in forestry and to introduce good forestry practice. Basis III was introduced in 1974, providing grants for new planting and additional supplements for broadleaves.
The Dedication Scheme was closed to new applications in 1981. Land still under Dedication could continue to be within the scheme but Dedication would terminate on a change of ownership.
Dedication schemes without a Plan of Operations and therefore receiving no grant, are deemed to be under Negative Covenant.

Dataset Attributes: Descriptor Dataset name Case_Name Dedication Scheme name Case_No Dedication Scheme reference number Basis Basis number of scheme (I, II or III) Date_Appr Date woodland became Dedicated PlanOfOps Period of current (or last) Plan of Operations Covenant Scheme in 'Positive' or 'Negative' Covenant Grid_Ref Grid Reference of property Total_Area Total area of scheme (* OL not always included) Man_Area Total area of scheme under management

#3 GOOGLE SEARCH

18 / 43

#4 SOCIAL SHARING

A screenshot of a Twitter post. The profile picture is a woman with short brown hair. The tweet text is: "Look at this awesome woodlands dataset <https://spatialdata.gov.scot/geonetwork/srv/api/records/4268B281-F173-4B7E-B2AE-1583C659253C>". The tweet card shows a summary for "Dedicated Woodlands": "Description: The Dedication Scheme (Basis I & II) was introduced in 1947 in order to encourage landowner...". Below the card is the text "Everyone can reply". At the bottom are standard Twitter interaction icons: image, GIF, reply, smiley face, and retweet, followed by a blue "Tweet" button.

Look at this awesome woodlands dataset
<https://spatialdata.gov.scot/geonetwork/srv/api/records/4268B281-F173-4B7E-B2AE-1583C659253C>

Dedicated Woodlands

Description: The Dedication Scheme (Basis I & II) was introduced in 1947 in order to encourage landowner...
spatialdata.gov.scot

Everyone can reply

| Tweet

DATA QUALITY REPORTS

Show all records where the title is greater than 50 characters in length, or the abstract is less than 150

READY ▶ 🔍 📁 ⚙

grid chart pie line scatter settings ▾

uuid	title	titlenght	abstractlength	owner	
a4845b40-c937-4e98-8412-f7fbc023a84c	Aberdeen Primary School Catchments	34	78	aberdeen	Export all data as csv
59fb4b40-d308-46a4-bb22-c2125f90105d	Local Authority boundaries - Scotland in Gemini 2.3 format	58	222	address	Export visible data as csv
31de199f-6261-468c-bach-e2ad3018d851	Conservation Areas and Article 4 Directions - Scotland	54	1921	address	Export all data as excel
938cfaac-3940-4a12-92ad-fb2495ca22d3	All Sources (within agglomerations) Noise Contours Lden Round 2	63	488	aecom	Export visible data as excel
7283a81e-4722-4c67-bf8d-e5852c699081	Industry (within agglomerations) Noise Contours Lden Round 1	60	445	aecom	Columns:
346bec4d-a12a-4b35-8f55-91090cc68e99	All Sources (within agglomerations) Noise Contours Lnigh Round 2	65	485	aecom	✓ uid
24bda07b-1821-4ed7-90ae-a2bc9b71df24	Roads (within agglomerations) Noise Contours Lnigh Round 1	59	440	aecom	
d5402a9e-8de3-4547-86f3-20fdac4e1e06	Industry (within agglomerations) Noise Contours Lnigh Round 1	62	442	aecom	
dac98505-1a00-4336-ba8f-d4d209b9082c	Roads (within agglomerations) Noise Contours Lnigh Round 2	59	440	aecom	
d92e0d81-603d-4adc-96ae-b5e1321af9b3	All Sources (within agglomerations) Noise Contours Lnigh Round 1	65	485	aecom	
1164aef-e5a9-4fd0-a836-ad4b9cb2d003	Roads (within agglomerations) Noise Contours Lden Round 1	57	443	aecom	
2e67d1be-d32e-463e-b6e5-6c086293c96a	Railway (within agglomerations) Noise Contours Lden Round 2	59	446	aecom	
b9f58d3c-8ecc-4bbb-bb84-ec2b27685ea3	All Sources (within agglomerations) Noise Contours Lden Round 1	63	488	aecom	
37f84402-e9f5-45a9-8482-9bb4d5887b3a	Roads (within agglomerations) Noise Contours Lden Round 2	57	443	aecom	

For our iShare customers, we've made it easy to link layers in iShare Maps or GIS to metadata records so you can provide additional context for the data you display online

File Help

Map Sources

- MyMaps
 - Maps
 - Restricted
 - NLC Planning
 - NLC Building Control
 - NLC Environmental Health
 - NLC Environment Team
 - NLC Highways
 - NLC Economic Development
 - NLC Children
 - NLC Adult Services
 - NLC Data
 - NLC Land and Property
 - NLC Local Plan
 - NLC Single Business View
 - NLC Emergency Planning
 - NLC Shield
- Maps
- Restricted
- NEL Call Centre
- NEL Planning
- NEL Standard User
- NEL POC
- NEL Local Plan
- Metadata
- BaseMaps
- MyHouse
- Logger

Data Share Connections

Workflow

Languages

Roles Editor

LINKING IN ISHARE

+ Restricted [nelincs_mapsources/nelincs_restricted]

+ NEL Standard User [nelincs_mapsources/nelincs_standarduser]

- Maps [nlincs_mapsources/nlincs_AllMaps]

Automatically find metadata for all layers

Remove all automatic metadata from layers

Remove all metadata from layers

- Boundaries

Layer Name	Metadata Name	Link Method	Layer Metadata Actions	
Localities			Find	
Lsoa	1	Isoa	Manual	Replace Remove
Msoa			Find	
North Lincolnshire		northlincolnshire	Manual	Replace Remove
Oa			Find	
Parishes		parishes	Manual	Replace Remove
Postcodes		postcodes	Manual	Replace Remove
Wards			Find	

+ Council

+ Democracy

+ Education and Learning

+ Environment

+ Flooding and Drainage

+ Highways

ACCESSIBLE

Metadata solutions need to be easy to use, for creators and users, and it should be easy to access the required data

We use Metadata Crawler to do the hard work of creating initial metadata records for all of your datasets so you don't have to do it manually, contribute to accessibility improvements in GeoNetwork, and have made the editor in Gemini 2.3 much easier to use

「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」

→「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」→「アカウント登録用のURLを取得する」

チ・キヨリ4のシメナハセ

・レジストリのURLを取得する。レジストリのURLを取得する。レジストリのURLを取得する。

アカウント登録用のURLを取得する

・ソースコードを読み取る。ソースコードを読み取る。ソースコードを読み取る。ソースコードを読み取る。
ソースコードを読み取る。ソースコードを読み取る。ソースコードを読み取る。ソースコードを読み取る。
ソースコードを読み取る。ソースコードを読み取る。ソースコードを読み取る。ソースコードを読み取る。

・ロジックを理解する。ロジックを理解する。ロジックを理解する。ロジックを理解する。

メモリ4の支N〇ネ

アカウント登録用のURLを取得する

アカウント登録用のURLを取得する。アカウント登録用のURLを取得する。アカウント登録用のURLを取得する。

アカウント登録用のURLを取得する。アカウント登録用のURLを取得する。アカウント登録用のURLを取得する。

メモリ4、支N〇ハセ

(Not metadata crawler)

GEMINI 2.3 EDITOR IMPROVEMENTS

▼ Metadata

File identifier *

ae0e855d-f0a2-438e-855c-6ef5400f4ef3

To support the operation of UK Location and INSPIRE, discovery metadata records must include a File Identifier for the resource.

1. File Identifier must be populated with a unique identifier. This may be a UUID generated by a metadata creation tool.
2. Once created and published to UK Location, the File Identifier must not be changed.
3. A change to the File Identifier would represent the creation of a new resource.
4. For more guidelines see [Gemini Guidance](#)

Metadata Language *

English

The purpose of this is to identify the language used in a multi-lingual metadata service, for example in the INSPIRE geo-portal.

1. For more guidelines see [Gemini Guidance](#)

Parent identifier

 Add parent record

Supports parent-child relationships in metadata. Allows navigation from a dataset record to the series record of which it forms a part.

1. Only to be used if the dataset is part of a series
2. For more guidelines see [Gemini Guidance](#)

Resource type *

Dataset

Scope to which metadata applies. To distinguish between datasets, series, and services

1. Choose the appropriate value from the list.
2. For more guidelines see [Gemini Guidance](#)

Scope to which metadata applies. To distinguish between datasets, series, and services

1. Choose the appropriate value from the list.
2. For more guidelines see [Gemini Guidance](#)

Metadata Date *

2019-09-13T16:59:27

Date on which the metadata was last updated, or was confirmed as being up-to-date, or if not updated, then the date it was created.

1. This should be updated whenever the metadata is updated due to a change in the dataset, or when it is reviewed and confirmed as wholly correct.
2. For more guidelines see [Gemini Guidance](#)

Metadata standard name

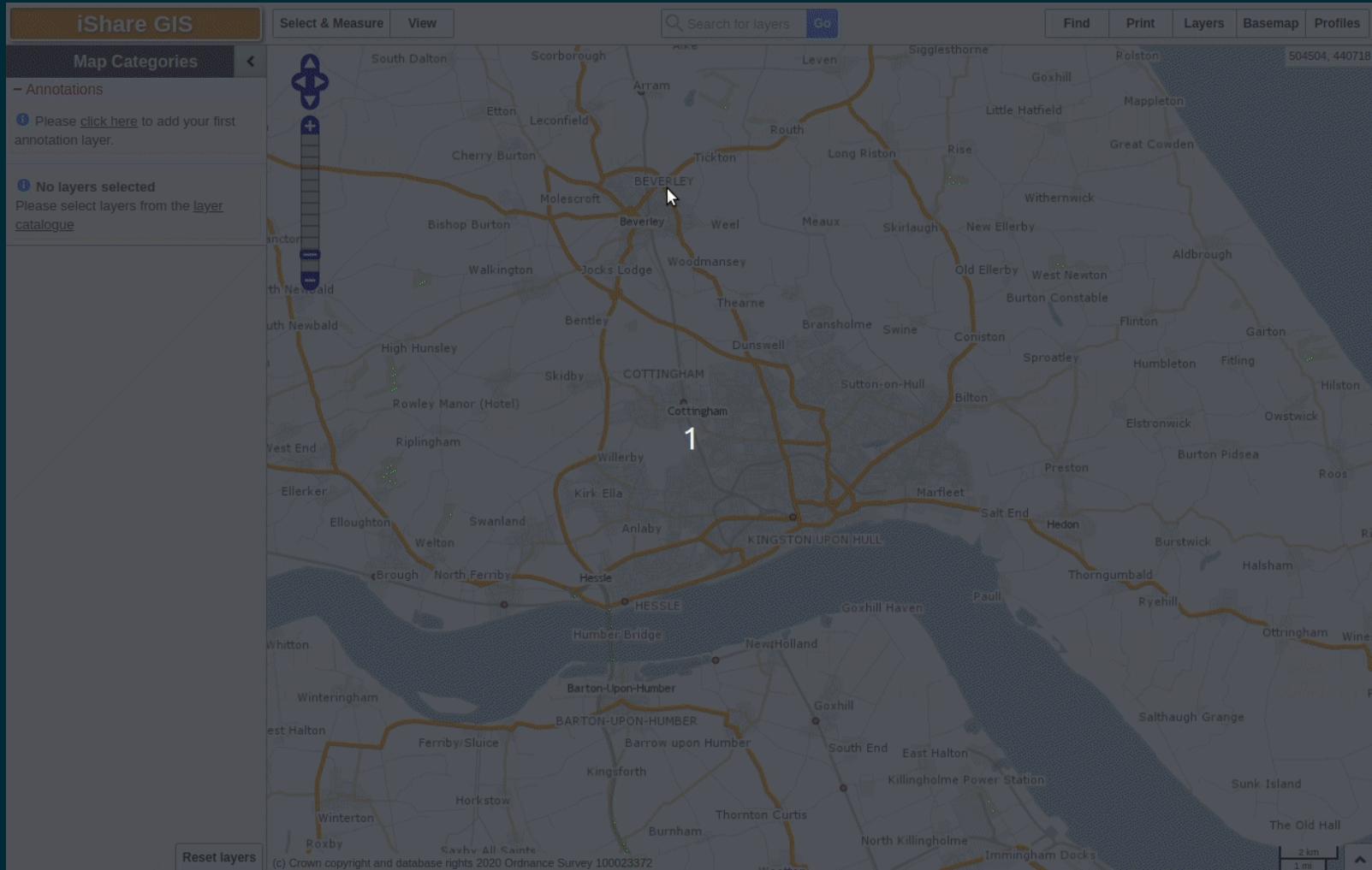
UK GEMINI

Metadata standard version

2.3

In iShare GIS, you can search for metadata directly in the application, helping you identify and display the data you need, quickly and easily

METADATA IN ISHARE GIS



INTEROPERABLE

Metadata and data need to be in a format that other people or machines can consume

We've implemented support for non-spatial metadata,
support for feature catalogs (ISO19110),
and have a tool for converting online vocabularies
into a format GeoNetwork can use for keywords

NON-SPATIAL METADATA

DEFRA non-spatial record TEMPLATE Joanna 14.04.202 ...

Categories



Cancel

Save & close

Save metadata



Identification



Validation



Title * DEFRA non-spatial record TEMPLATE Joanna 14.04.2020

The title of the resource (mandatory)

Alternate title

TEST record for the Defra Non-spatial template

An alternative title for the resource (optional)

Date *

Creation

14 / 04 / 2019

x

-- : --



Date *

Revision

14 / 04 / 2020

x

-- : --



The date the resource was created (mandatory), and the most recent revision date

Abstract *

 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

 Curabitur pretium tincidunt lacus. Nulla gravida orci a odio. Nullam varius, turpis et commodo pharetra, est eros bibendum elit, nec luctus magna felis sollicitudin mauris. Integer in mauris eu nibh euismod gravida. Duis ac tellus et risus vulputate vehicula. Donec lobortis risus a elit. Etiam tempor. Ut ullamcorper, ligula eu tempor congue, eros est euismod turpis, id tincidunt sapien risus a quam. Maecenas fermentum

A description of the resource (mandatory)

- Identification
- Point of contact
- Descriptive keywords
- Temporal Extent
- Distribution Information
- OnLine resource
- Lineage

FEATURE CATALOG



PlaceHolderURL

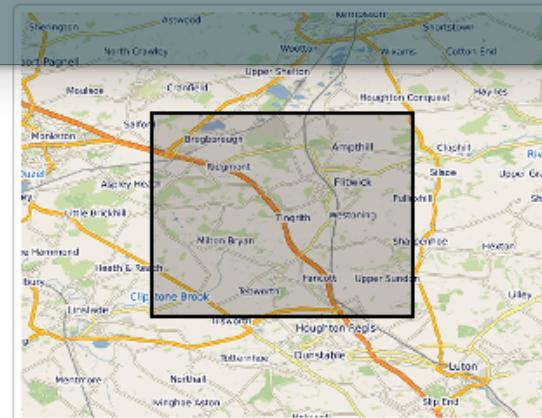
PlaceHolderDesc <https://www.centralbedfordshire.gov.uk/>

thirdparty.woodland_trust_wood

This dataset is published in the database PG:dbname=cbc-lob host=127.0.0.1 user=**** password=**** in table thirdparty.woodland_trust_wood.



Spatial extent



Feature catalog



Feature catalog cbc-lob.thirdparty.woodland_trust_wood

Technical information

Name	Definition
ceasedte (String)	
effectdte (String)	
gid (Integer64)	
namefull (String)	
notes (String)	
refno (String)	
type (String)	
typefull (String)	
featcode (Integer64)	
originator (String)	
plantechselected (Integer64)	



Temporal extent

Creation date

2000-01-01

Revision date

2020-07-17



Publishing Body



Metadata Updated:

3 months ago

Share on social sites



History

In iShare, records can have a direct reference back to the Spatial Data Warehouse, and the iShare View and Download Service URLs, linking everything together in a way that's easy to maintain

LINKING BACK TO ISHARE

[Back to search](#)< PreviousNext >EditDeleteManage recordDownloadDisplay mode

Point data representing location of cemeteries in North Lincolnshire.

On going

Download and links



This dataset is published in the database dbname=sdw
host=[REDACTED] port=5432 table=[REDACTED]



This dataset is published in the download service (WFS) available at http://mapsonline.northlincs.gov.uk/GetOWS.ashx?VERSION=1.1.0&FORMAT=image%2Fpng&MAPSOURCE=nlincs_mapsources%2Fnlincs_AllMaps&typename=Cemeteries&service=WFS&REQUEST=GetFeature.

Add to map

Zipped shapefile download

Download

Zipped shapefile download http://mapsonline.northlincs.gov.uk/GetOWS.ashx?MAPSOURCE=nlincs_mapsources%2Fnlincs_AllMaps&outputformat=SHAPEZIP&VERSION=1.1.0&SERVICE=WFS&REQUEST=GetFeature&TYPENAME=Cemeteries&SRNAME=EPSG:27700

No ratings

See all feedbackAdd your review

Spatial extent



Temporal extent

Publication date

1996-04-01

Revision date

2019-09-01

Publishing Body



REUSABLE

Data should be in formats that can be shared, with quality metadata to ensure it is signed off for publication, used correctly, and is fit for purpose

We can check that links to View and Download services are working correctly, can send automated reminders when datasets are due or overdue an update and are working on snippets to allow you to indicate data is fit for release

Associated resources



LINK CHECKER

+ Add ▾

Online resources

+ Add



OGC:WMS-1.1.1-http-get-map

<http://gis.astuntechnology.com/mapserver/ms64?map=E:/isharedata/live...>



Addresses_OGC

Addresses

Service

+ Add

[Astun Default Layers](#)



Suggestions

Suggestions

http://gis.astuntechnology.com/mapserver/ms64?map=E:/isharedata/live/_mapserverConfig/mapsources__default.map&language=English& cannot be reached. Run this task to remove it.

OVERDUE REMINDERS

Metadata records overdue an update - [REDACTED] Metadata Portal. Inbox x



geonetwork@astuntechnology.com via amazones.com

to me ▾

Dear [REDACTED]

The following metadata records are overdue for review, please login and check them, and mark as updated:

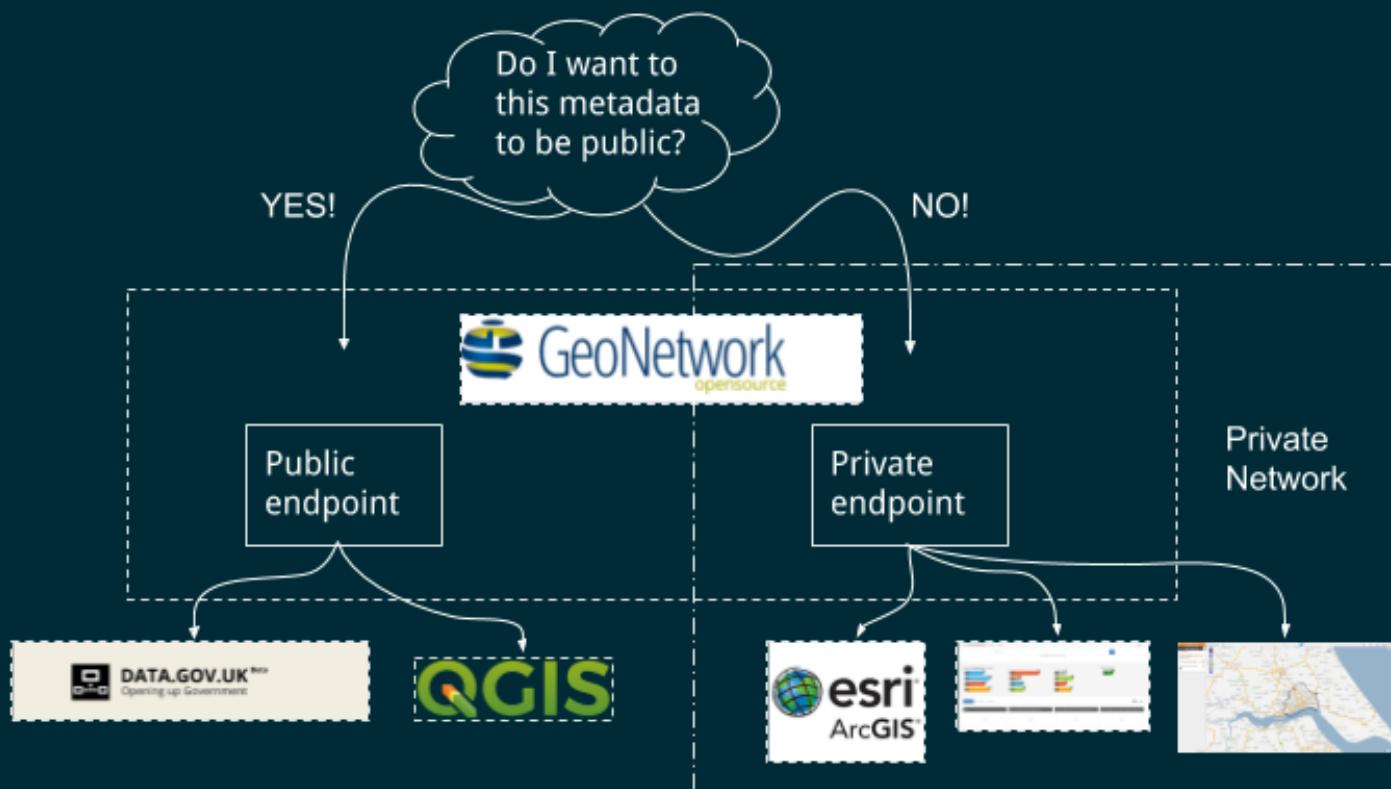
Test hdc_District Boundaries <https://metadata.astuntechnology.com/geonetwork/> [REDACTED] /eng/catalog_search#/metadata/e5b9cb40-bf94-488e-b33a-dee3f71c8534

Regards,

Metadata Team.

For our all users, including iShare, we can provide separate public and private catalog endpoints, giving a clear separation between metadata they wish to make publicly available, and that which should be accessible only within their organisation.

PUBLIC AND PRIVATE ENDPOINTS



All of this can be deployed quickly and easily on our resilient cloud infrastructure, whether you're an iShare customer or not, and whether you have existing metadata or not!

MAKING YOUR DATA...

FINDABLE with structured data and GeoNetwork improvements,

ACCESSIBLE with Metadata Crawler and Gemini 2.3 editor enhancements,

INTEROPERABLE with non-spatial and feature catalogs,

REUSABLE by ensuring links to data are correct and sending update reminders

THANK YOU!
ANY QUESTIONS?

Contact sales@astuntechnology.com for more details