

Table of Contents

1. Scope	4
2. Conformance	5
3. References	6
4. Terms and Definitions	7
4.1. coordinate reference system	7
4.2. coordinate system	7
4.3. dataset	7
4.4. feature	7
4.5. feature collection; collection	7
4.6. height	7
4.7. linestring	8
4.8. point	8
4.9. polygon	8
4.10. GeoPackage file	8
4.11. tile	8
4.12. Valid GeoPackage	8
4.13. vector and vector data	8
4.14. vector geometry	8
4.15. Version	9
5. Conventions	10
5.1. Identifiers	10
6. Results from Interoperability Experiments	11
7. Guidance on the use of GeoPackages	13
8. Flattening the Schema	14
9. Layer Names within a GeoPackage	17
Annex A: Conformance Class Abstract Test Suite (Normative)	19
A.1. Conformance Class A	19
A.1.1. Requirement 1	19
A.1.2. Requirement 2	19
Annex B: Title ({Normative/Informative})	20
Annex C: Revision History	21
Annex D: Bibliography	22

Open Geospatial Consortium

Submission Date: <yyyy-mm-dd>

Approval Date: <yyyy-mm-dd>

Publication Date: <yyyy-mm-dd>

External identifier of this OGC® document: <http://www.opengis.net/doc/{doc-type}/{standard}/{m.n}>

Internal reference number of this OGC® document: 19-066

Version: n.n

Category: OGC® Best Practice

Editor: Michala Hill

OGC Best Practices for Conversion of CDB Vector Data to Geopackage

Copyright notice

Copyright © <year> Open Geospatial Consortium

To obtain additional rights of use, visit <http://www.opengeospatial.org/legal/>

Warning

This document defines an OGC Best Practice on a particular technology or approach related to an OGC standard. This document is not an OGC Standard and may not be referred to as an OGC Standard. It is subject to change without notice. However, this document is an official position of the OGC membership on this particular technology topic.

Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Document type: OGC® Best Practice

Document subtype:

Document stage: Draft

Document language: English

License Agreement

Permission is hereby granted by the Open Geospatial Consortium, ("Licensor"), free of charge and subject to the terms set forth below, to any person obtaining a copy of this Intellectual Property and any associated documentation, to deal in the Intellectual Property without restriction (except as set forth below), including without limitation the rights to implement, use, copy, modify, merge, publish, distribute, and/or sublicense copies of the Intellectual Property, and to permit persons to whom the Intellectual Property is furnished to do so, provided that all copyright notices on the intellectual property are retained intact and that each person to whom the Intellectual Property is furnished agrees to the terms of this Agreement.

If you modify the Intellectual Property, all copies of the modified Intellectual Property must include, in addition to the above copyright notice, a notice that the Intellectual Property includes modifications that have not been approved or adopted by LICENSOR.

THIS LICENSE IS A COPYRIGHT LICENSE ONLY, AND DOES NOT CONVEY ANY RIGHTS UNDER ANY PATENTS THAT MAY BE IN FORCE ANYWHERE IN THE WORLD.

THE INTELLECTUAL PROPERTY IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NONINFRINGEMENT OF THIRD PARTY RIGHTS. THE COPYRIGHT HOLDER OR HOLDERS INCLUDED IN THIS NOTICE DO NOT WARRANT THAT THE FUNCTIONS CONTAINED IN THE INTELLECTUAL PROPERTY WILL MEET YOUR REQUIREMENTS OR THAT THE OPERATION OF THE INTELLECTUAL PROPERTY WILL BE UNINTERRUPTED OR ERROR FREE. ANY USE OF THE INTELLECTUAL PROPERTY SHALL BE MADE ENTIRELY AT THE USER'S OWN RISK. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR ANY CONTRIBUTOR OF INTELLECTUAL PROPERTY RIGHTS TO THE INTELLECTUAL PROPERTY BE LIABLE FOR ANY CLAIM, OR ANY DIRECT, SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES, OR ANY DAMAGES WHATSOEVER RESULTING FROM ANY ALLEGED INFRINGEMENT OR ANY LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR UNDER ANY OTHER LEGAL THEORY, ARISING OUT OF OR IN CONNECTION WITH THE IMPLEMENTATION, USE, COMMERCIALIZATION OR PERFORMANCE OF THIS INTELLECTUAL PROPERTY.

This license is effective until terminated. You may terminate it at any time by destroying the Intellectual Property together with all copies in any form. The license will also terminate if you fail to comply with any term or condition of this Agreement. Except as provided in the following sentence, no such termination of this license shall require the termination of any third party end-user sublicense to the Intellectual Property which is in force as of the date of notice of such termination. In addition, should the Intellectual Property, or the operation of the Intellectual Property, infringe, or in LICENSOR's sole opinion be likely to infringe, any patent, copyright, trademark or other right of a third party, you agree that LICENSOR, in its sole discretion, may terminate this license without any compensation or liability to you, your licensees or any other party. You agree upon termination of any kind to destroy or cause to be destroyed the Intellectual Property together with all copies in any form, whether held by you or by any third party.

Except as contained in this notice, the name of LICENSOR or of any other holder of a copyright in all or part of the Intellectual Property shall not be used in advertising or otherwise to promote the sale, use or other dealings in this Intellectual Property without prior written authorization of LICENSOR or such copyright holder. LICENSOR is and shall at all times be the sole entity that may authorize you or any third party to use certification marks, trademarks or other special designations to indicate compliance with any LICENSOR standards or specifications. This Agreement is governed by the laws of the Commonwealth of Massachusetts. The application to this Agreement of the United Nations Convention on Contracts for the International Sale of Goods is hereby expressly excluded. In the event any provision of this Agreement shall be deemed unenforceable, void or invalid, such provision shall be modified so as to make it valid and enforceable, and as so modified the entire Agreement shall remain in full force and effect. No decision, action or inaction by LICENSOR shall be construed to be a waiver of any rights or remedies available to it.

i. Abstract

<Insert Abstract Text here>

ii. Keywords

The following are keywords to be used by search engines and document catalogues.

ogcdoc, OGC document, <tags separated by commas>

iii. Preface

NOTE

Insert Preface Text here. Give OGC specific commentary: describe the technical content, reason for document, history of the document and precursors, and plans for future work. > Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

iv. Submitting organizations

The following organizations submitted this Document to the Open Geospatial Consortium (OGC):

Organization name(s)

v. Submitters

All questions regarding this submission should be directed to the editor or the submitters:

Name Affiliation

Chapter 1. Scope

NOTE

This Best Practice compliments the optional OGC CDB Extension that defines the behavior and requirements for encoding vector data in a GeoPackage container for use in a CDB data store. Both are grounded in the existing CDB Core requirements and the GeoPackage core requirements for vector data.

This document discusses the results and conclusions of the Interoperability Experiment conducted to determine the optimal use of GeoPackage in CDB, assuming relatively minor changes to the CDB current standard. It also offers general guidance on the use of GeoPackage in CDB, with details on the flattened attribute schema.

Before reading this standard, please remember that the idea is to restrict the encoding of a dataset to a single vector format per CDB Version. Since a “CDB” is made of one or more “Versions” (as specified by Configuration.xml), and that each CDB Version can have a different encoding for a given dataset, the result is that a “CDB” may pretty well exist with multiple encodings for the same dataset. This means that if you wish to use GeoPackage containers, you need to create a version that will just contain GeoPackages of the vector data and not include any Shapefiles.

Chapter 2. Conformance

This best practice defines [TBD] requirements / conformance classes.

The standardization targets of all conformance classes are "GeoPackages in CDB Data Store". The main requirements class is:

Core. [\[Add link to Annex A later\]](#)

The Core specifies requirements that all GeoPackages that are to be stored in a CDB store that shall be implemented.

Conformance with the standard and best practice shall be checked using all the relevant tests specified in Annex A (normative) of this document. The framework, concepts, and methodology for testing, and the criteria to be achieved to claim conformance are specified in the OGC Compliance Testing Policies and Procedures and the OGC Compliance Testing web site.

In order to conform to this OGC® interface best practice, a software implementation shall choose to implement the conformance levels specified in Annex A (normative)

All requirements-classes and conformance-classes described in this document are owned by the standard(s) identified.

Chapter 3. References

The following normative documents contain provisions that, through reference in this text, constitute provisions of this document. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the normative document referred to applies.

ISO / TC 211: ISO 19115-1:2014 Geographic information — Metadata — Part 1: Fundamentals (2014)

OGC: OGC 15-113r5, Volume 1: OGC CDB Core Standard: Model and Physical Data Store Structure (2017)

OGC: OGC 16-070r3, Volume 4: OGC CDB Best Practice use of Shapefiles for Vector Data Storage (2017)

OGC: OGC 12-128r15, OGC GeoPackage Encoding Standard - with Corrigendum (2018)

OGC: OGC ???, OGC Vector GeoPackage Extension (TBD)

Chapter 4. Terms and Definitions

This document uses the terms defined in Sub-clause 5.3 of [OGC 06-121r8], which is based on the ISO/IEC Directives, Part 2, Rules for the structure and drafting of International Standards. In particular, the word “shall” (not “must”) is the verb form used to indicate a requirement to be strictly followed to conform to this Best Practice.

For the purposes of this document, the following additional terms and definitions apply.

4.1. coordinate reference system

coordinate system that is related to the real world by a datum [ISO 19111]

4.2. coordinate system

set of mathematical rules for specifying how coordinates are to be assigned to points [ISO 19111]

4.3. dataset

collection of data, published or curated by a single agent, and available for access or download in one or more formats [DCAT]

Note: The use of 'collection' in the definition from [DCAT] is broader than the use of the term collection in this standard. See the definition of feature collection.

4.4. feature

abstraction of real world phenomena [ISO 19101-1:2014]

Note: If you are unfamiliar with the term 'feature', the explanations on Spatial Things, Features and Geometry in the W3C/OGC Spatial Data on the Web Best Practice document provide more detail.

4.5. feature collection; collection

a set of features from a dataset

Note: In this standard, 'collection' is used as a synonym for 'feature collection'. This is done to make this document easier to understand for those that are not geo-experts.

4.6. height

Distance of a point from a chosen reference surface measured upward along a line perpendicular to that surface. [ISO 19111]

Note 1 to entry: A height below the reference surface will have a negative value, which would embrace both gravity-related heights and ellipsoidal heights.

4.7. linestring

curve composed of straight-line segments (aka line) [ISO 19136:2007]

4.8. point

0-dimensional geometric primitive, representing a position [ISO 19136:2007]

Note to entry: The boundary of a point is the empty set.

4.9. polygon

planar surface defined by 1 exterior boundary and 0 or more interior boundaries [ISO 19136:2007]

4.10. GeoPackage file

a platform-independent SQLite database file that contains GeoPackage data and metadata tables with specified definitions, integrity assertions, format limitations and content constraints.

4.11. tile

a geometric shape with known properties that is the result of the tiling (tessellation) of a plane. A tile consists of a single connected "piece" without "holes" or "lines" (topological disc).

4.12. Valid GeoPackage

A GeoPackage that contains features per clause Features and/or tiles per clause Tiles and row(s) in the gpkg_contents table with data_type column values of "features" and/or "tiles" describing the user data tables.

4.13. vector and vector data

quantity having direction as well as magnitude

Note to entry: A directed line segment represents a vector if the length and direction of the line segment are equal to the magnitude and direction of the vector. The term vector data refers to data that represents the spatial configuration of features as a set of directed line segments. [ISO 19123:2005]

4.14. vector geometry

representation of geometry through the use of constructive geometric primitives [ISO 19107:2003]

4.15. Version

A collection of pure CDB Datasets and/or user-defined datasets. Please see section 3.2 of Volume 1: OGC CDB Core Standard: Model and Physical Data Store Structure for details on the CDB versioning strategy and structure.

Chapter 5. Conventions

This section provides details and examples for any conventions used in the document. Examples of conventions are symbols, abbreviations, use of XML schema, or special notes regarding how to read the document.

5.1. Identifiers

The normative provisions in this standard are denoted by the URI <http://www.opengis.net/spec/cdb-vector-geopackage/1.2>

All requirements and conformance tests that appear in this document are denoted by partial URIs which are relative to this base.

Chapter 6. Results from Interoperability Experiments

OGC CDB and GeoPackage are both standards increasingly used in M&S and GEOINT, but they both contain weaknesses and strengths when it comes to the combined needs of both industries. The combined technology of GeoPackage and OGC CDB is hoped to be a deterministic repository of easily read geospatial datasets suitable for storage, runtime access, and dissemination for live, virtual, constructive, gaming, and Mission Command (MC) systems. The use of CDB raised concerns in both industries because it utilizes Esri Shapefiles and Presagis OpenFlight, both of which are now open standards but were once proprietary. The primary issue is that Shapefiles leverage the dBase IV format to store feature attribution. To address the limitations of ESRI Shapefiles within CDB, potential designs were explored to integrate OGC GeoPackage with OGC CDB. GeoPackage technology was designed to ensure Mobile / Handheld devices could interoperate with systems within the larger Command Post and U.S. Army Enterprise. Current Geopackage adaptations can natively store 3D information, routing information, various types of data ‘tiles’, and other geographically related datasets. An implementation of OGC GeoPackage within CDB aligns CDB vector data storage with the U.S. Army architecture.

Four approaches were testing in the Interoperability Experiment:

- Experiment 1: Conversion of Shapefiles in one or more CDB data stores into GeoPackages as required for Experiments 2, 3, and/or 4.
- Experiment 2: One-to-one conversion of a Shapefile into a GeoPackage.
- Experiment 3: Store each CDB Level of Detail (LOD) as a layer in GeoPackage.
- Experiment 4: Store each Geocell of Vector Data as a layer in GeoPackage.

Experiment 2 was further broken down into Options 1a – 1d. Each experiment offered its own set of potential advantages, disadvantages, and lessons learned.

Utilizing Option 1c:

- Fully flattened instance, class, and extended level attribute tables
- There is a 4:1 to 7:1 reduction in the number of files.
- Some duplication of data, resulting in larger files.
- There is one layer per GeoPackage.
- The Feature Class and Extended Attributes are populated for each feature.
- Full “Off the Shelf” GeoPackage software compatibility.
- Best Experiment 2 option in terms of file size and performance time.

Utilizing Option 1d:

- Flatten CDB standard instance and class attribute – maximum GIS tools compatibility
- “Off the Shelf” GeoPackage software compatibility for CDB standard attributes.
- Table for extended attributes using Related Tables GeoPackage extension

- This approach utilizes a standard normalized relational database design, utilizing foreign keys.
- Some duplication of data, resulting in larger files.
- There is one layer per GeoPackage.
- The Feature Class and Extended Attributes are populated for each feature.

The final method selected for conversion of a Shapefile CDB to a GeoPackage CDB is a hybrid of Options 1c and 1d. Because “Off the Shelf” GeoPackage software will not be aware of the class and extended level attributes, we needed to “flatten” the class level attributes with the instance level attributes. It was not possible to effectively flatten the extended level attributes at this time, and those will remain in a separate table with regular table relationships. The decision was made not to use the Related Tables GeoPackage extension.

Several observations were made during the experiments. The output created by the recommended method of conversion allows backward compatibility. However, the GeoPackage CDB results in larger size on disk, which could be problematic. The GeoPackage format (especially as outputted by GDAL) is very space inefficient when it comes to encoding tiles with small quantities of data. But it does reduce the total number of files on disk significantly and eliminates the occurrence of dBase files containing 0 KB.

For more details on the experiments and information above, please see the OGC CDB Vector Data in GeoPackage Interoperability Experiment Engineering Report.

Chapter 7. Guidance on the use of GeoPackages

An GeoPackage is an open, standards-based, platform-independent, portable, self-describing, compact format for transferring geospatial information. It is a platform-independent SQLite database file. The GeoPackage specification was developed by OGC. This document describes the encoding of Vector Data in GeoPackage only.

Per Requirement 4 of the CDB Vector GeoPackage Extension, the position of all points is expressed using WGS-84 geographic coordinates (latitude, longitude, altitude), as explained in Volume 8: CDB Spatial Reference Systems Guidance.

As detailed in the following table of GeoPackage Geometry Codes (Table 30 of Appendix G of the OGC GeoPackage Specification), the following geometry types are supported in GeoPackage:

Code	Name
0	GEOMETRY
1	POINT
2	LINESTRING
3	POLYGON
4	MULTIPOIPT
5	MULTILINESTRING
6	MULTIPOLYGON
7	GEOMETRYCOLLECTION

Per Requirement 2 of the CDB Vector GeoPackage Extension, while the GeoPackage model supports encoding of 8 different types that can be stored in the same GeoPackage, the CDB standard requires a maximum of 3 feature geometry types per tile.

Chapter 8. Flattening the Schema

The CDB Shapefile standard provides three attribution schemas to represent attribution data. Those three have been flattened into two for use with GeoPackage.

- Instance-level attribution schema
- Extended-level attribution schema

As CNAM in the CDB Shapefile schema was only used as a relationship field, it has been removed from the flattened schema. Figure 1 shows the flattened instance and class-level attributes, and how the extended attributes link to a feature. I_CEI, I_GEAI, and I_VEI are all instance-level attributes and are linked to a specific feature. C_CEI, C_GEAI, and C_VEI are all class-level attributes and are linked to all features in a layer.

Example: Record ID 4 in the extended-level attributes table is related to record ID 1 in the linear features table and is an instance-level attribute. Record IDs 5 and 8 are also linked to record ID 4, and all apply to the same linear feature. Note that a 0 in the LNK column denotes the end of the applicable record list. Record 9 in the extended-level attributes table is a class-level attribute related to all features that are present in the highway layer.

All of the information that is needed to instance features is organized in accordance to the CDB tile structure. All the tiled GeoPackage dataset files are located in the same directory. The dataset's second component selector (CS2) is used to differentiate between files with the same extension or with the same Vector features. Table 1 presents the list of codes that are allocated. Note that Vector datasets do not necessarily use all of the Dataset Component Selector 2 reserved codes. Users of the CDB standard should refer to the appropriate section for an enumeration of the supported File Component Selector 2 codes as well as the ones specific to the Dataset.

The Vector dataset concept and the feature code concepts overlap somewhat; some of the Vector datasets are generalizations or specializations of feature codes. Note that the same feature should not have two representations.

Table 1: Component Selector 2 for Vector Dataset File Names

CS2	File Extension	Dataset Component Name	Supported Shape Type
001	*.gpkg	Point features	Point, PointZ, PointM, MultiPoint, MultiPointZ, MultiPointM
002	*.gpkg	Lineal features	PolyLine, PolyLineZ, PolyLineM
003	*.gpkg	Polygon features	Polygon, PolygonZ, PolygonM, Multipatch

004	*.gpkg	Lineal figure point features	Point, PointZ, PointM, MultiPoint, MultiPointZ, MultiPointM
005	*.gpkg	Polygon figure point features	Point, PointZ, PointM, MultiPoint, MultiPointZ, MultiPointM
006	*.gpkg	2D relationship tile connections	N/A
007	*.gpkg	2D relationship dataset connections	N/A

Table 2: Component Selector 2 for Vector Dataset Table Names

CS2	Dataset Component Name	Supported Shape Type
001	Point features	Point, PointZ, PointM, MultiPoint, MultiPointZ, MultiPointM
002	Point feature class-level attributes	N/A
003	Lineal features	PolyLine, PolyLineZ, PolyLineM
004	Lineal feature class-level attributes	N/A
005	Polygon features	Polygon, PolygonZ, PolygonM, Multipatch
006	Polygon feature class-level attributes	N/A
007	Lineal figure point features	Point, PointZ, PointM, MultiPoint, MultiPointZ, MultiPointM
008	Lineal figure point feature class-level attributes	N/A
009	Polygon figure point features	Point, PointZ, PointM, MultiPoint, MultiPointZ, MultiPointM
010	Polygon figure point feature class-level attributes	N/A
011	2D relationship tile connections	N/A
012	Deprecated	N/A
013	Deprecated	N/A

CS2	Dataset Component Name	Supported Shape Type
014	Deprecated	N/A
015	2D relationship dataset connections	N/A
016	Point feature extended-level attributes	N/A
017	Lineal feature extended-level attributes	N/A
018	Polygon feature extended-level attributes	N/A
019	Lineal Figure Point extended-level attributes	N/A
020	Polygon Figure Point extended-level attributes	N/A

Chapter 9. Layer Names within a GeoPackage

Each GeoPackage file shall contain one and only one feature table. The feature table shall include all attributes (fields) required by the CDB standard (Volume 1, 5.7.1.2. CDB Attribution). The single feature table shall be named identically to the GeoPackage filename (not including the ".gpkg" filename extension). A GeoPackage file may include an optional extended level attribute table. The extended level attribute table shall not contain geometry. The name of the extended level attribute table shall use the CS2 selector shown in Table 2. Table 3 shows the file naming convention.

Table 3: Tiled Dataset File Naming Convention 1

Field	Description
Lat	Geocell Latitude – Identical to the name of the directory defined in section 3.6.2.1, Directory Level 1 (Latitude Directory).
Lon	Geocell Longitude – Identical to the name of the directory defined in section 3.6.2.2, Directory Level 2 (Longitude Directory).
Dnnn	Character D followed by the 3-digit code assigned to the dataset.
Snnn	Character S followed by the 3-digit value of Component Selector 1.
Tnnn	Character T followed by the 3-digit value of Component Selector 2.
LOD	Level of Detail – As defined in section 3.3.8.5, Level of Detail.
Un	UREF – Identical to the name of the directory as defined in section 3.6.2.5, Directory Level 5 (UREF Directory).
Rn	RREF – A reference to the Right Index of a tile. Character R (Right direction) followed by the column number as described in this section.
xxx	File extension as per file type.

The example below and Table 4 show the filename structure for a road feature. GeoPackage Filename: N13E045_D201_S002_T003_L00_U0_R0.gpkg Feature Table Name: N13E045_D201_S002_T003_L00_U0_R0 Extended Attribute Table Name: N13E045_D201_S002_T017_L00_U0_R0

Table 4: Basic Filename Structure

Latitude	Longitude	Dataset Code	Selector (CS1)	Selector 2 (CS2)	LOD	UREF	RREF
N13	E045	_D	201	_S	002	_T	003

Figure 2 shows the flattened schema with filenames as they might appear for a road dataset. The feature and extended attribute tables are shown contained within the GeoPackage file.

Annex A: Conformance Class Abstract Test Suite (Normative)

NOTE

Ensure that there is a conformance class for each requirements class and a test for each requirement (identified by requirement name and number)

A.1. Conformance Class A

A.1.1. Requirement 1

Test id:	/conf/conf-class-a/req-name-1
Requirement:	/req/req-class-a/req-name-1
Test purpose:	Verify that...
Test method:	Inspect...

A.1.2. Requirement 2

Annex B: Title ({Normative/Informative})

NOTE

Place other Annex material in sequential annexes beginning with "B" and leave final two annexes for the Revision History and Bibliography

Annex C: Revision History

Date	Release	Editor	Primary clauses modified	Description
2016-04-28	0.1	G. Editor	all	initial version

Annex D: Bibliography

Example Bibliography (Delete this note).

The TC has approved Springer LNCS as the official document citation type.

Springer LNCS is widely used in technical and computer science journals and other publications

NOTE

- For citations in the text please use square brackets and consecutive numbers:
[1], [2], [3]

– Actual References:

[n] Journal: Author Surname, A.: Title. Publication Title. Volume number, Issue number, Pages Used (Year Published)

[n] Web: Author Surname, A.: Title, <http://Website-Url>

[1] OGC: OGC Testbed 12 Annex B: Architecture. (2015).