Engineer

GDAL

地理空間データを読み、変換するライブラリー

どのようなものか

GDALは、Geospatial Data Abstraction Libraryの略で、GDAL/OGRともよく表記されるオープンソースのソフトウェアです。一つの形式の地図データを開いて別の形式へ書き出し、要求された場合は、その過程で再投影も行います。プロジェクト自身の文書では、変換ライブラリーと説明しています。

空間データには、二つの形があります。ラスターは、衛星画像、航空写真、標高モデルなどの、セルの格子です。ベクターは、住所、道路、区画の境界など、属性付きの点、線、多角形です。GDALは、呼び出すプログラムへ、形ごとに一つの抽象モデルを提供し、その背後に、形式ごとのドライバーを置きます。ラスターにはGeoTIFF、JPEG 2000、NetCDF、ベクターにはShapefile、GeoJSON、GeoPackage、KML、PostgreSQLのテーブルなどがあり、どちらにも、ほかに多くあります。GDAL向けに書いたプログラムは、同じ呼び出しで、すべてを読めます。

ベクター側は、長くOGRと呼ばれ、GDAL 2.0で統合するまで、ドライバーの構造が別でした。そのため、ベクターのプログラムは、今も古い接頭辞を使います。`ogrinfo`がベクターデータセットを説明し、`ogr2ogr`が変換します。Frank Warmerdamが、1998年の終わりごろに開発を始めました。MIT系のライセンスで配布する、Open Source Geospatial Foundation、OSGeoのプロジェクトです。

使い方は三つです。最初は、ラスターの`gdalinfo`、`gdal_translate`、`gdalwarp`、ベクターの`ogrinfo`、`ogr2ogr`などのコマンドラインです。バージョン3.11からは、`gdal raster`や`gdal vector`などのサブコマンドを持つ、一つの`gdal`コマンドもあり、まだ暫定としています。二つ目は、Pythonなどのバインディングから、コードで使う方法です。三つ目は、ほかのソフトウェアに組み込まれた形で使う方法です。例えばQGISは、ドライバーでベクター形式を読み、

Processingツールボックスで、GDALのプログラムを提供します。典型的な変換は一行です。次は、ShapefileをGeoPackageへ変え、デンマーク全体で使うUTMゾーンへ、再投影します。

```
ogr2ogr -f GPKG -t_srs EPSG:25832 parcels.gpkg parcels.shp
```

執筆時の現行版は3.13.3です。この層でGDALと競合するツールは少数です。ArcGISは自前の変換ツールを持ち、個々の形式にも、小さい専用の読み取りツールがあります。しかし、顧客の送るものを、何でも受け取るチームは、GDALに依存する別のライブラリーを通す場合も含め、結局、構成へ入れることが多くなります。

概念を判断に生かす

「GDALとは？」を実際の判断に生かしましょう。その概念を、開発または購入するシステムに結び付けて考えるお手伝いをします。 船舶、車両、区画、センサーなど、場所に関わる情報を、検索・計測・共有できる地図にまとめます。

サービスを見る お問い合わせ →

用語の背後にある前提を検証する

解説とリンク先の実績を参考に、プロジェクトについての質問を整理してください。ツールを選ぶ前に、前提を一緒に検討できます。 地理情報パイプラインを50倍に高速化 — 50x 地理データパイプラインの高速化

地理情報パイプラインを50倍に高速化 お問い合わせ →

なぜ重要なのか

どの地理空間システムにも、外のデータが入る境界があります。自治体のShapefile、測量者のGeoTIFF、協力先のGeoJSON、誰かがブラウザーで描いたKMLなどです。GDALに関わる判断は、誰が、その境界に立つかです。デスクトップツールをクリックする人か、毎回同じことをするスクリプトかです。GDALは、スクリプトを書く負担を小さくします。

`ogr2ogr` や `gdalwarp` のコマンドにした変換は、レビューで読め、バージョン管理でき、次の納品でも繰り返せます。手動の変換は、手順を覚えた人が、一度だけ行うものです。毎月、バッチで届くデータでは、その違いが、品質を保てるかを決めます。

ライセンスも重要です。利用者数に応じた料金はかからず、販売元の承認も必要ありません。同じツールを、ノートPC、ビルドサーバー、コンテナー内で使えます。

三つ目の理由は、ストレージ上のデータへ直接アクセスできることです。仮想ファイルシステムにより、ドライバーはデータをその保存先から読み取れます。`/vsicurl/` なら、HTTPで必要なバイト範囲を求め、大半のドライバーは、リモートのデータセット全体をダウンロードせず、部分ごとに読み取ります。`/vsis3/` は、Amazon S3へ同じことをします。タイル化し、縮小した概観画像を持ち、要求回数を抑えるよう配置したCloud Optimized GeoTIFFと組み合わせれば、何ギガバイトものラスターをオブジェクトストレージに置いたまま、範囲を一つずつ読めます。

負担もあります。GDALは、任意で追加する依存先が多い、大きなネイティブライブラリーです。入るドライバーは、コンパイルの方法によるため、同じく「GDALがある」二台でも、読める形式が違う場合があります。変換は忠実で、出力に選んだ形式の限界まで、忠実に引き継ぎます。

実際の活用例

このサイトの二つのプロジェクトに、異なる役割で登場します。

2015年から2022年のGISソフトウェア開発のチームリーダーの仕事では、三十を超えるGISプロジェクトを納品し、それぞれを本番でも支えました。空間データは、PostGIS付きのPostgreSQLに置き、GDAL/OGRで、Shapefile、GeoJSON、GeoTIFF、KML、ベクタータイルの間を移しました。QGISとJOSMがデータそのものの作業を担い、そのデータを使う地図はMapbox GLとLeafletで作りました。GeoTIFFがラスター、ほかはベクターなので、ライブラリーのラスターとベクターの両方の機能を使いました。仕事は、2D、3Dのデジタル地図とLiDARの処理から、幾何補正、ベクター化、屋内地図、ナビゲーションまでに及びました。

2022年から2024年の、大規模な地理空間分析プロジェクトのデータエンジニアの役割では、GDALは誤りを見つけて修正する道具の一つでした。GDAL、QGIS、ArcGISで、空間データの誤りを診断し、修正しました。同じデータクレンジングの処理が繰り返し必要になるため、Bashスクリプトで自動化しました。PostGISが空間の検索と索引を、PL/pgSQLとTransact-SQLのプロシージャーが、PostgreSQLとSQL Server内の変換を担当し、クレンジングした結果を、Mapboxで作った対話的な可視化へ渡しました。GDALを含む道具で、1,000を超える問題を直し、空間クエリーの実行時間が35%減った理由の一つにもなりました。

問題が起きるところ

GDALは、指示どおりに動きます。多くの問題は、意図と少し違う指示から始まります。

点が、別の大陸へ行きます。 GDAL 3.0からは、座標系を定義する機関が定めた軸の順序に従います。緯度と経度でおなじみのEPSG:4326は、緯度が先です。GDAL 2向けのコードとGeoJSONは、経度が先です。混ぜると座標の組がすべて逆順になり、コペンハーゲンの点がアラビア海へ現れます。プログラムの境界で一つに決め、明示します。経度を先にするコードでは、空間参照の `SetAxisMappingStrategy(OAMS_TRADITIONAL_GIS_ORDER)` が、文書の示す切り替えです。

列名が切れ、日付から時刻が消えます。 Shapefileの属性名は10文字までです。長ければ、書くときに縮め、衝突すれば、8文字へ切り、連番を付けます。`abcdefghijkl` は `abcdefgh_1` になります。この形式が保存するのは日付だけで、時刻は含まれません。他のソフトウェアとの相互運用のため、GDALは `.shp` と `.dbf` を2 GB以下に保つよう勧めます。コードページを示す `.cpg` がなく、文字コードを誤認すると、デンマーク語やキリル文字が文字化けすることがあります。受け取る側が認めるならGeoPackageを書き、Shapefile必須なら、変換前に、列名を意図的に変えます。

リモートのラスターを、丸ごと取得します。 接頭辞なしの `https://` では、HTTP疑似ドライバーが、他のドライバーが読む前に、全体をメモリーへ入れます。小さいものには合っても、全国のオルソ画像には不向きです。 `/vsicurl/` を付け、必要な範囲だけを取り、大きいラスターをCloud Optimized GeoTIFFで保存して、範囲の数も抑えます。その後は、変えません。COGファイルを直接編集すると、その内部配置が壊れます。GDAL 3.8からは、配置が壊れることを明示して許可しない限り、更新用に開くことを拒みます。

CSVの数値が、文字列になります。 指示がなければ、全フィールドを文字列で読みます。人口の列が文字順に並び、後段で合計を求めると、型エラーになります。開く際に `AUTODETECT_TYPE=YES` を指定するか、各列の型を示す `.csvt` を、CSVの隣へ置きます。

更新で、Pythonのインポートが壊れます。 PyPIの `gdal` は、マシン上のGDALを使ってビルドするため、版が合う必要があります。プロジェクトのインストール手順では、インストール済みのGDALに合わせてPythonパッケージのバージョンを固定します。

```
pip install gdal[numpy]=="$(gdal-config --version).*"
```

参考資料

- GDALの文書 — プロジェクト自身による、変換ライブラリー、ドライバー、ライセンスの説明です。
- プログラム — `gdalinfo` や `ogr2ogr` から、新しい `gdal` までの、全コマンドラインツールです。
- GDALの仮想ファイルシステム — `/vsicurl/` 、 `/vsis3/` など、データのある場所で読む接頭辞です。
- OGRの座標参照系と座標変換のチュートリアル — GDAL 3.0で変わった、軸順についての節を含みます。
- COG：Cloud Optimized GeoTIFF生成ツール — ファイルの内部配置と縮小画像、ファイルを直接編集するとそれらが崩れる理由を説明します。

用語集

実際の制約に合わせて選ぶ

業務の流れ、既存のインターフェース、変更を必要とする利用者をお知らせください。検証可能な納品内容と、お客様のチームが保守を続けられる引き継ぎ内容を定められます。

料金を見る お問い合わせ →

用語を実務に結び付ける

「GDALとは？」がお客様の組織にどう当てはまるかをご相談ください。現在のシステムの具体例が一つあれば、話を始められます。

サービスを見る

welcome@engineer.company +45 50 10 61 84 デンマーク、コペンハーゲン

このページを持ち歩く

連絡先カード

心を込めて丁寧に制作： D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

著作権 © 2025 – 現在 · Engineer ApS · すべての権利を留保します。

印刷元： https://engineer.company/ja/glossary/gdal/