

United Nations Group of Experts on Geographical Names
2025 Session

28 April – 2 May 2025

Item 5 (b) of the provisional agenda

Technical expertise: geographical names data management

Background document

Available in English only

World Geographical Names Database - Data Collection
Prepared by the UNGEGN Secretariat

World Geographical Names Database - Data Collection

1. Overview

Data submissions to the World Geographical Names Database (WGND) must comply with the following structure and content requirements. This ensures consistency, quality, and interoperability across contributions from Member States.

Submissions are requested in **GeoJSON** format, where each **Feature** represents a city or populated place, including multilingual toponymic (name) information. Countries are invited to discuss with the Secretariat (mark.iliffe@un.org) options for data submission if technical realities prevent submission of data within the GeoJSON format.

2. GeoJSON Structure Requirements

The overall structure must follow the GeoJSON standard:

```
json
{
  "type": "FeatureCollection",
  "features": [Feature, Feature, ...]
}
```

Each **Feature** must contain:

- **Geometry:** Point (longitude, latitude)
- **Properties:** Toponyms (names) related to the location

3. Detailed Field Requirements

Feature Object

Field	Type	Required	Notes
type	String	Yes	Must be "Feature"
geometry.type	String	Yes	Must be "Point"
geometry.coordinates	Array	Yes	[longitude, latitude] (in decimal degrees)
properties.toponyms	Array	Yes	Array of toponym objects (see below)

Toponym Object (inside `properties.toponyms`)

Each toponym must follow this structure:

Field	Type	Required	Description
nameShort	String	Yes	Short form of the geographical name
nameRomanised	String or Null	No	Romanized version of the name, if available
nameLong	String or Null	No	Full/official name, if different from <code>nameShort</code>
language	String	Yes	Language of the name (e.g., "French", "Arabic")
isCapital	Boolean	Yes	<code>true</code> if the location is a national capital
isRomanized	Boolean	Yes	<code>true</code> if the name is romanized
cityID	Integer	No	Provided by the Secretariat as part of the UUID initiative
toponymType	String	Yes	Type of name ("Endonym", "Exonym", "Variant name", etc.)

4. Examples

A simple valid submission would look like:

```
json
{
  "type": "FeatureCollection",
  "features": [
    {
      "type": "Feature",
      "geometry": {
        "type": "Point",
        "coordinates": [18.5646533294, 4.3661938576]
      },
      "properties": {
        "toponyms": [
          {
            "nameShort": "Bangui",
            "nameRomanised": null,
            "nameLong": null,
            "language": "French",
            "isCapital": true,
            "isRomanized": true,
            "cityID": 60,
            "toponymType": "Endonym"
          },
          {
            "nameShort": "بانغي",
            "nameRomanised": null,
            "nameLong": "",
            "language": "Arabic",
            "isCapital": true,
```

```

        "isRomanized": false,
        "cityID": 60,
        "toponymType": "Variant name"
    }
  ]
}

```

5. Data Quality Expectations

- **Coordinate Accuracy:** Ensure accurate decimal degree coordinates.
 - **Language Consistency:** Language names must be clearly specified (e.g., "English", "Spanish").
 - **Toponym Type:** Clearly distinguish between "Endonym", "Exonym", "Variant name", etc.
 - **Completeness:** Submissions should provide as many languages as possible (especially the six UN official languages, if available).
 - **Romanization:** Where non-Latin scripts are used, indicate whether a romanized version is available.
-

6. Submission Format

- Submit the final data in a **single .geojson file** or a **compressed .zip archive** if multiple files are necessary.
- File encoding must be **UTF-8** to support multilingual data.
- If submitting updates to existing cities, ensure consistency of the `cityID` as provided by the Secretariat.