

24 April 2023

English

---

## United Nations Group of Experts on Geographical Names

2023 session

New York, 1 – 5 May 2023

Item 3 of the provisional agenda \*

Reports of the Chair and the Secretariat

### World Geographic Names Database

Submitted by UNGEGN Secretariat\*\*

#### Summary

The UNGEGN World Geographical Names Database is a repository of the short and full names of countries (193 UN member states and two non-Member States observers), their capitals, and major cities (with population over 100,000) in a multilingual, multi-scriptural and geo-referenced format. Authoritative city endonyms are provided mainly by national name authorities and sound files are being added to assist users with pronunciation. To date, the Database contains toponyms, endonyms, and exonyms of features detailing over 2,700 country names, some 6,100 names for 3,362 cities, with more than 970 audio files.

The development of the Database is guided by UNGEGN's mandated decisions. The first iteration of the Database was built after 2007. Due to challenges around the technology used to manage the data provided to the Group of Experts, this iteration of the Database was closed in 2019. In response, the secretariat mobilised surplus resources from other information technology projects within the 2020 budgetary year to revive the Database utilising modern web technologies and adhering to the prevailing technological guidelines of the United Nations.

At its 2021 session, the Group of Experts adopted decision 2/2021/2 (E/2021/69), which “appreciated the work done by the secretariat in maintaining the website of the Group of Experts and reviving the online World Geographical Names database, noted that the completion of those modifications and access to the database as soon as possible would be of great benefit”. Moreover, the development of the Database has been conducted to ensure alignment with the principles of the Secretary-General's Data Strategy and compliance with the prevailing environment of publishing maps and web applications within the United Nations (ST/AI/189).

The full report details the activities of the secretariat in reviving the Database, including: reviewing the functionality of the Database; discussing the technological approach taken; examining the developmental journey of the Database to ensure its compliance with prevailing

---

\* GEGN.2/2023/1

\*\* The full report was prepared by the secretariat

guidelines, norms, and geospatial standards of both the Group of Experts and the United Nations; providing a set of recommendations for improving geographic coverage of the Database, including through engaging Member States to provide toponymic data and through the integration of other forms of authoritative data, such as from the United Nations Demographic Yearbook. The report also features a discussion that considers the future expansion of the Database, to include features such as a mechanism that enables the Group of Experts to work on the nexus of linked data and geographic names, and the toponyms of indigenous and other geographies.

## I. Introduction

1. This report summarises the main activities undertaken by the Secretariat of the United Nations Group of Experts on Geographical Names (UNGEGN) to redevelop the World Geographic Names (WGN) Database. The WGN Database is a repository of the short and full names of countries (193 UN member states and two non-Member States observers), their capitals, and major cities (with populations over 100,000) in a multilingual, multi-scriptural and geo-referenced format. Authoritative city endonyms within the Database have been mainly provided by national name authorities, and some sound files are available to assist users with pronunciation. To date, the Database contains toponyms, endonyms, and exonyms of features detailing over 2,700 country names, some 6,100 names for 3,362 cities, with more than 970 audio files.

2. The need for an authoritative depository of geographic names under the purview of the Group of Experts is a topic with a long and storied history within UNGEGN. For example, at its 22nd session in July 2004, the Group of Experts recommended the development of an authoritative database on country and major city names<sup>1</sup>. Moreover, at the ninth Conference on the Standardization of Geographical Names in August 2007, the Conference passed resolution IX/62, which considered “the ongoing efforts carried out by the secretariat of the Group of Experts, within the United Nations Statistics Division (UNSD), towards the building of a database that includes country names and major city names of the world in a multilingual, multi-scriptural and geo-referenced format”. Moreover, as the Group of Experts implements its strategic plan and programme of work for the period 2021-2029, the WGN Database offers the technological foundation with which to manage an authoritative depository of geographic names, not just for cities and countries, but also to help the Group to demonstrate how geographical names connect people to geographic location and why this matters as part of the broader recognition of culture, heritage and language.

3. At its 2021 session, the Group of Experts adopted decision 2/2021/23, which “appreciated the work done by the secretariat in maintaining the website of the Group of Experts and reviving the online WGN Database, noted that the completion of those modifications and access to the database as soon as possible would be of great benefit”. The WGN Database needed to be revived due to the progressive obsolescence of its technology platform. Crucially, the existing data which underpinned the previous iteration of the WGN Database was retrievable, leading to no loss of data provided by Member States.

4. Thus, this present report details the activities of the secretariat in reviving the WGN Database, including: reviewing the functionality of the Database; discussing the technological approach taken; examining the developmental journey of the Database to ensure its compliance with prevailing guidelines, norms, and geospatial standards of both the Group of Experts and the United Nations; providing a set of recommendations for improving geographic coverage of the Database, including through engaging Member States to provide toponymic data and through the integration of other forms of authoritative data, such as from the United Nations Demographic Yearbook, the United Nations Terminology Database “UNTERM”, or the gazetteer maintained by the United Nations. The report also features a discussion that considers the future expansion of the Database, including features such as a mechanism that enables the Group of Experts to work on the nexus of linked data and geographic names, and the toponyms of indigenous and other geographies.

---

<sup>1</sup> see [GEGN/22](#) Report of the United Nations Group of Experts on Geographical Names on the work of its twenty-second session

<sup>2</sup> see [E/CONF.98/136](#) Ninth United Nations Conference on the Standardization of Geographical Names

<sup>3</sup> see [E/2021/69](#) Report of the United Nations Group of Experts on Geographical Names on its 2021 session

## II. The World Geographic Names Database

5. Geographic names provide a basis from which to integrate data at a specific geographic location and at a specific geographic scale (i.e., a region, country, city, or other types of geography). In this regard, the previous iteration of the WGN Database was initially developed and launched by the secretariat in 2009 and was operational until 2019.

6. The management of the WGN Database became increasingly difficult over the years before it formally ceased operation. There was no formal administration interface, and all updates had to manually inputted using complex database queries. The user interface of the Database was outdated and perhaps unwieldy, in comparison to more recent interfaces, which potentially further discouraged its use. Despite these challenges, the secretariat persevered in maintaining the Database until late 2019. However, several underlying technologies used in the Database, such as Adobe Flash 8 and ASP.net 2.0, became deprecated by their developers, posing potentially severe security and privacy concerns. As a result, the secretariat was forced to place the first version of the Database in an “end-of-life read-only” state to avoid any future issues that could arise from the continued use of outdated technology and ensure the integrity of the data within the Database.

7. Yet, by identifying and repurposing surplus resources previously dedicated to technological development in other areas of UNSD in the intersessional period, the secretariat was able to redeploy these resources to redevelop the WGN platform, aligned to prevailing technological practices and common geospatial standards. In this respect, the secretariat was guided by the Group of Expert’s decisions<sup>4</sup> and the prevailing technology environment of the UN. Moreover, the secretariat liaised closely with the UNGEGN Expanded Bureau to identify user requirements and ensure that the mandates and requirements of UNGEGN were technically translated into the new WGN Database.

### The various dimensions of data

8. The data model of the WGN Database captures all data in a multilingual, multi-scriptural and geo-referenced format. The secretariat welcomes the report and discussion paper presented by the Working Group on Geographical Names Data Management “On geographical names data modelling and transfer related standards, manuals, or guidelines” at this present session. The discussion paper discusses data models for geographical names in detail<sup>5</sup>. Despite the parallel development of the Database and this discussion paper, the secretariat acknowledges and welcomes the significant alignment between the concepts detailed within the discussion paper and the technical reality of the Database.

9. In this regard, the modelling of a geographic name in the WGN Database context is primarily linguistic. For example, Annex 1 contextualises the relationships of “Antananarivo” in Madagascar and its various linguistic dimensions associated with that geography. The Database has been designed to collect the linguistic dimensions of city and country and includes a third, generic, ‘geography’ datatype. This allows for the Database to be easily extended, perhaps to consider the toponyms of indigenous geographies or maritime areas.

10. Complementing the toponymic element is time. The WGN Database now includes functionality that can manage the temporal dimension of toponyms, enabling the cataloguing of toponymic changes to geographies over time. This extends the previous iteration of the Database, which only allowed for

---

<sup>4</sup> see UNGEGN Decisions 2/2021/2, 1/2019/2 and UNCSGN resolution resolution IX/6

<sup>5</sup> see [GEGN.2/2023/99/CRP.99](#) On geographical names data modelling and transfer related standards, manuals, or guidelines

the dissemination of current toponyms. The discussion of these concepts is continued in subsequent sections of this report and highlights opportunities with other concepts under the purview of the Group of Experts, such as Linked Data.

### **The current state of the WGN Database**

11. The current version of the WGN Database discussed at this present session of the Group of Experts is an advanced work in progress. There are pieces of work remaining that, when completed, will significantly contribute to the functionality and running of the Database. Primarily, these are grouped into 1. Improvements to the underlying data; and, 2. Improvements to the system itself. Regardless, the redevelopment of the Database now provides Member States and the general public with an opportunity to disseminate and use, in both human- and machine-readable formats, the ability to harmonise, integrate and interrogate authoritative geographic names. The new Database is anchored by three separate, but interconnected, applications:

- (a) **An administration module**
- (b) **A public web API**
- (c) **A public dashboard**

12. These three applications serve two main purposes. First, they allow the secretariat to maintain and manage data in the WGN Database effectively. Second, they provide public users, including the UNGEGN community, the UN system, and other geographic names stakeholders, with a visually appealing way to visualise and query the data in the Database. This facilitates the dissemination and use of toponymic data in both human- and machine-readable formats. These applications also offer the ability to harmonise, integrate, and interrogate authoritative geographic names, enabling better coordination and analysis of geographic name data.

13. Technically, the WGN Database was redeveloped using MS SQL and C#, following the Model-View-Controller paradigm. It will be launched on the UN Azure Cloud and the codebase is managed on a GIT-enabled internal UN source code repository. These design decisions were taken to ensure the sustainability of the source code after the initial development of the application.

### **The Administration Module**

14. The WGN Administration Model enables the secretariat to manage the WGN Database in a user-friendly and efficient manner. Functionality highlights include:

- (a) **Data Entry and Editing:** The administration module provides a user-friendly interface for entering and editing geographic name data in the Database. This includes features that ensure data accuracy and consistency by validating data formatting and other quality measures related to data input.
- (b) **A Temporal Module:** that brings the temporal dimension to toponyms, capturing the toponymic changes to geographies over time. For example, the Kingdom of Eswatini changed its name from “Kingdom of Swaziland” to its present form on 11 May 2018.
- (c) **Version Control:** The administration module enables the secretariat to track changes made to the Database and revert to previous versions if needed. This helps maintain data integrity and provides an audit trail, if needed, to identify the provenance of updates made to the Database.

(d) **Search and Retrieval:** The administration module includes advanced search and retrieval functionalities to enable the secretariat to retrieve and view data from the Database efficiently. This includes various search filters, sorting options, and export capabilities to facilitate data management.

(e) **User Access Management:** The administration module allows the secretariat to manage user accounts and permissions, including creating new accounts, modifying access levels, and revoking access as needed. This would help ensure that only authorised members of the secretariat can make updates to the Database.

(f) **System Configuration and Maintenance:** The administration module provides options to configure and maintain the technical aspects of the Database, such as managing server settings, backups, and system updates.

(g) **Security and Authorisation:** The administration module includes robust security measures, such as user authentication, role-based access control, and encryption of sensitive data, to protect the integrity and confidentiality of the Database.

15. Importantly, the Administration Module can only be accessed by Secretariat staff while on UN premises. This is to ensure the integrity of the data within the WGN Database and prevent the possibility of non-Secretariat staff from gaining access to the Database.

### **The WGN Application Programming Interface (API)**

16. The WGN API is an open and publicly available, read-only, web-based API that enables all to access the data maintained within the WGN Database. The API has been developed following the good practices extolled by the UN Committee of Experts on Global Geospatial Information Management (UN-GGIM) Guide to the Role of Standards in Geospatial Information Management<sup>6</sup>; namely, to ensure that geospatial information is FAIR (Findable, Accessible, Interoperable, and Reusable). Moreover, the API grounds itself in the ambition of the Secretary General’s Data Strategy<sup>7</sup> in that the data is openly available for Member States, the UN system, and the public to use and consume the data within the WGN Database.

17. The API shares data in a format called “GeoJSON” through different endpoints, as listed in Annex 2. These endpoints provide standardised geospatial information that can be used to share different dimensions and features of the data. The API provides read-only access to the WGN Database by design.

18. Once data is input or updated within the Administration Module, it is automatically available for consumption in the API. The WGN Dashboard utilises the API as a means of displaying information. In effect, this has also helped ‘road test’ the data flows and overall WGN Database architecture.

### **WGN Dashboard**

19. The WGN Dashboard is a web-based interface that offers public users the ability to query toponyms in all languages deposited within the WGN Database, with the user interface available in

---

<sup>6</sup> Revised Guide to the Role of Standards in Geospatial Information Management <https://ggim.un.org/Revised-Standards-Guide/>

<sup>7</sup> “We also need to become better in governing greater exchange of data, not just publishing more Open Data for open access, but also developing licenses and data sharing agreements that enable partners to integrate more deeply with us, in responsible ways” Secretary General’s Data Strategy - [https://www.un.org/en/content/datastrategy/images/pdf/UN\\_SG\\_Data-Strategy.pdf](https://www.un.org/en/content/datastrategy/images/pdf/UN_SG_Data-Strategy.pdf)

the six official languages of the United Nations — Arabic, Chinese, English, French, Russian and Spanish. Key components of a WGN Dashboard include:

- (a) **Data Visualisation:** The Dashboard follows a ‘map first’ interface, allowing users to zoom and pan to explore the geographical distribution of toponyms. The Dashboard includes tools that enable a ‘drilling down’; into the data, zooming in or out on maps, or clicking on data points to retrieve more detailed information about specific geographic names.
- (b) **Interactive Features for Data Exploration:** The Dashboard allows users to interact with the data and visualisations, providing options to filter and customise the displayed data based on their linguistic preferences or requirements – such as viewing all toponymic labels as endonyms or in a specific language.
- (c) **Data Updates:** The Dashboard reflects in real-time the data of the WGN Database, ensuring that users have access to the most up-to-date information about geographic names.

20. Cartographically, to contextualise and visualise data provided by Member States in an accessible manner and to meet the instructions provided by UNGEGN’s decisions, the basemap of the WGN Database is provided by the UN Clear Map service of the Office of Information Communication Technology (OICT), UN Secretariat. This is a web-mapping service prepared in accordance with OICT to meet the guidelines set out by ST/AI/1898. Specifically, the basemap is an adapted version of Clear Map, where all map labels have been removed and replaced by data from the Database. This approach was undertaken in consultation with the Office of Information Communication and Technology’s Operations Support Division - Geospatial Information Section (née Cartographic Section – OSD-GIS). This continues the relationship between UNSD and OSD-GIS, as initially highlighted in resolution IX/6, which recommended: “that UNSD, with the support of the United Nations Cartographic Section for base maps and the Second Administrative Level Boundaries project for administrative divisions data sets, work in collaboration with the United Nations Group of Experts of Geographical Names and Member States to further develop, populate and maintain the geographical names database of the United Nations Group of Experts on Geographical Names, initially containing names of countries, capitals and major cities”.

### III. The Process of Redevelopment

21. In February 2021, surplus resources were identified by the UNGEGN Secretariat and were requisitioned for use. These resources were not monetary; they were made available due to surplus developer time in an UN-approved contractor (Trigyn Technologies Limited) becoming available. Over the period of February to October 2021, these resources were channelled into the development of a WGN Database prototype, developing the base of the three applications mentioned above. In October 2021, in part due to a surplus of monetary resources arising from underspend on mission travel related to the COVID-19 pandemic, the secretariat secured \$36,750 to continue the work of the contractor to finalise the development of the Administration Module and API while significantly progressing the Dashboard.

22. As the main application that users will interact with, the Dashboard proved itself as a particularly challenging component of the WGN Database architecture. Although there are many web technologies that enable engaging and user-friendly mapping experiences, these technologies do not allow for dynamic interaction with map labels or enable the user to select the language in which they would want the labels to be displayed. Through discussions with technologists in the geospatial

---

<sup>8</sup> see [ST/AI/189/Add.25/Rev.1](#) Regulations for the Control and Limitation of Documentation – Guidelines for the Publication of Maps

community, the Secretariat and Mapbox (a technology company that creates mapping APIs and other map-based visualisation libraries for developers), in an informal collaboration, developed the Dashboard using their Mapbox Studio and Mapbox GL libraries. Currently, anticipating the use of the WGN Database, the ‘free’ tier of service from Mapbox is being used. If this situation changes, the secretariat will investigate how to resource the continued running of the application.

23. In April 2022, the engagement with the contractor was finalised, and the secretariat assumed control of the codebase and worked to implement the WGN Database within the UN’s ICT infrastructure. As of the submission of this report, the UNGEGN Secretariat is awaiting the finalisation of the UN’s IT application onboarding process. This process is mandated by OICT and is a necessary step that must be taken before the official launch of the application. This IT onboarding process consists of several processes and tests to ensure compliance with the UN’s prevailing guidelines on cyber security, the testing of its architecture and components, agreeing on backup procedures and other IT considerations. Pragmatically, the official launch of the application is a formality, and the link to the Dashboard will be made available on the website of the Group of Experts when OICT deems the onboarding process to be completed.

24. Regrettably, the UNGEGN Secretariat does not have dedicated continuing resources for significant future developments to the WGN Database. Instead, their existing resources are primarily focused on the operation of the Database, with limited capacity to address bugs and resolve issues that may arise. This may impact the speed at which the secretariat can respond to and resolve such bugs. This is an item that the Group of Experts may wish to consider further, particularly in light of the limited resources available to the secretariat.

#### **IV. Discussing the Future of the World Geographic Names Database**

25. Throughout the development of the WGN Database, the secretariat collaborated closely with leaders in the Group of Experts to ensure that their requirements were effectively translated into a suite of applications that form the Database. These and other discussions raised several opportunities for the potential future of the Database.

#### **Enabling progress towards the 2030 Agenda for Sustainable Development**

26. The Secretary-General’s [Sustainable Development Report 2022](#), considers millions of data points provided by over 200 countries and areas, and in this respect paints a sobering picture. Using the latest available data on the SDG indicator framework, the report reveals that the 2030 Agenda for Sustainable Development is in grave jeopardy due to multiple, cascading, and intersecting crises. In addition to this, many countries are still unable to provide the data needed for informed policy analysis and decision-making. In this regard, the integrative power of geospatial information is a vital tool to harmonise all forms of data and statistics.

27. Through linking geographic names with SDG indicators, it is possible to analyse progress towards the SDGs at the regional or national level. This can help countries identify areas that require further attention and prioritise interventions accordingly. In this regard, while the SDGs are global in nature, their achievement largely depends on local action. Local governments, communities, and individuals play a critical role in implementing and localising the SDGs by tailoring them to the specific context, needs, and priorities of their local areas. Three SDG indicators directly necessitate



reporting of SDG indicators at the city level: SDG 11.2.19, SDG 11.6.110, and SDG 11.7.111; yet almost all indicators are relevant at the local city level, underscoring the importance of the Group of Experts work in this area.

### **Working towards a ‘complete’ dataset – Updating the current toponyms, receiving new toponymic data and integrating other authoritative datasets**

28. The development of the WGN Database of applications provides the United Nations, Member States and the general public with an opportunity to, in both human- and machine-readable formats, the ability to analyse, integrate, disseminate and use authoritative geographic names. But the data within the Database is not geographically complete and has other areas of potential improvement. While all UN Member States are included at the country level, the geographic coverage of cities is of variable quality. This is further complicated by incomplete toponymic data for cities that are included.

29. By comparing with authoritative datasets, it is possible to estimate the potential gap that exists within the dataset. Currently, there are 3,362 cities within the WGN Database. Comparing this with the United Nations Demographic Yearbook<sup>12</sup>, there are an additional 2,735 cities that could be incorporated, offering toponyms in English and French. In most instances, while these names are not provided by National Names Authorities, the data is provided by the National Statistical Office and thus confers a level of national authority. The toponyms could then be augmented in the six official languages of the UN by integrating data from the UN Terminology Database (UNTERM) as it is a multilingual database of official terminology relevant to the work of the UN. This is one mechanism that could be used to reach geographic completeness and capture all cities globally with a population of over 100,000 people and will be actioned by the secretariat in due course.

30. Moreover, in the intersessional period, the secretariat has liaised with the Secretariat for the UN Code for Trade and Transport Locations<sup>13</sup> (UN/LOCODE) to ascertain whether data from both datasets could be exchanged in a meaningful manner. UN/LOCODE is a five-character code system that provides a coded representation for the names of ports, airports, inland clearance depots, inland freight terminals and other transport-related locations which are used for the movement of goods for trade. While there exists an opportunity to establish a liaison group between UN/LOCODE and UNGEGN at the strategic level to explore interoperability between the two datasets, there are perhaps more technical opportunities that are more aligned with the work programme of the Group of Experts that can be explored in the short term, particularly given that UN/LOCODE is focused on hyper-local geographies.

31. In the operation and testing of the WGN Database, the secretariat has observed underlying issues with the data. These include spelling errors, positional accuracy issues and other geospatial quality and consistency issues. In the coming intersessional period, the secretariat aims to work with OSD/GIS to resolve these issues, with the oversight of relevant stakeholders from the Group of Experts. Collaboration with OSD/GIS will also be undertaken as the secretariat works to integrate data aligned with its prevailing mandates.

---

<sup>9</sup> Specifically, the data series “SP\_TRN\_PUBL” - Proportion of population that has convenient access to public transport (%)

<sup>10</sup> Specifically, data series “EN\_REF\_WASCOL” - Municipal Solid Waste collection coverage, by cities (%)

<sup>11</sup> Specifically, the data series “EN\_URB\_OPENSP” - Average share of the built-up area of cities that is open space for public use for all (%)

<sup>12</sup> The Demographic Yearbook, a publication compiled by UNSD since 1948 which collates authoritative data from over 230 National Statistics Offices on population size and composition, births, deaths, marriage and divorce, as well as respective rates, on an annual basis.

<sup>13</sup> UN/LOCODE : <https://unece.org/trade/uncefact/unlocode-ag>

## A Unique Identifier for the City Level

32. At the database level, all data within the WGN Database has a unique identifier. For countries, the M49 classification is used. The M49 classification is an UN-developed standard that classifies countries and areas for statistical and analytical purposes. Each M49 code consists of a three-digit number which serves to identify each country or territory uniquely. These numbers were initially obtained by arranging the countries in alphabetical order, using their names in the English language, and assigning a number following this order<sup>14</sup>. Specifically, the M49 classifier consists of three levels of classification:

- (a) **Regions:** The first level divides the world into regions such as Africa, the Americas, Asia, Europe, and Oceania.
- (b) **Subregions:** The second level further divides regions into subregions, such as Eastern Africa, Southern Asia, Northern Europe, etc.
- (c) **Countries or areas:** The third level classifies individual countries or areas within subregions using a three-digit numerical code. For example, the code “004” represents Afghanistan, “156” represents China, and “840” represents the United States of America.

33. Notably, cities are absent from this classification; there is an opportunity to promote the unique identifier assigned to cities within the WGN Database as a standardised, unique identifier for cities, in a similar manner to M49.

## The opportunity of Linked Data in the World Geographic Names Database

34. The development of the WGN suite of applications provides Member States, the UN system and the general public with an opportunity to disseminate and use, in both human- and machine-readable formats, the ability to harmonise, integrate and interrogate authoritative geographic names. For several years, the Group of Experts has been striving to standardise geographical names to make them easily accessible and findable at national and international levels. The ways in which geographical names can be made accessible have evolved over time, with the recognition of the potential of utilising the internet for this purpose dating back to resolution VII/9 of the Seventh Conference in 1998<sup>15</sup>. This resolution recommended providing data on toponymic websites free of charge to promote international standardisation. Recently, the opportunity of the Semantic Web and Linked Data has been considered by the Group of Experts, at its previous second session in 2021<sup>16</sup> and at this present session<sup>17</sup>.

35. The secretariat has been experimenting with integrating Unique Resource Identifiers (URI) into the data model of the WGN Database. This allows geographic names and related toponymic data to be queryable as Linked Data, which promotes interoperability. The secretariat has also investigated how the Linked Data can help organise and structure concepts related to geographic names into a hierarchical taxonomy. This approach helps define concepts and their relationships, such as mapping the abbreviated geographic name “USA” to the toponym “United States of America”. This enables a more precise and standardised representation of geographic names, facilitating data integration and

---

<sup>14</sup> see [ST/ESA/STAT/SER.M/49/Rev.4](#) Standard Country or Area Codes for statistical Use

<sup>15</sup> see [E/2007/89](#) Report of the United Nations Group of Experts on Geographical Names on the work of its twenty-fourth session

<sup>16</sup> see [GEGN.2/2021/6/CRP.6](#) Linked Data considerations for geographical names standardisation

<sup>17</sup> see [GEGN.2/2023/75/CRP.75](#) Linked open data developments within the United Nations Group of Experts on Geographical Names

exchange. The secretariat looks forward to engaging with the relevant functional groups of the Group of Experts to explore how the WGN Database can be further enabled in this area.

### **Ongoing Governance for the WGN Database**

36. As the secretariat works to navigate the final technical procedures leading to the formal launch of the WGN Database, it is now an appropriate time for the Group of Experts to consider an appropriate governance mechanism for it. The secretariat takes note of the report<sup>18</sup> of the Working Group on Geographical Names Data Management<sup>19</sup>, which provides a series of recommendations relevant to the Database. With respect to these potential future opportunities, other concerns for the consideration of the Group of Experts include how the data should be licensed, how toponymic pronunciations can be captured as audio files, how National Name Authorities should be engaged to capture endonyms and exonyms, and other how other geographies, such as those relating to indigenous, maritime, and are not Countries or Cities, can be managed within the Database. Moreover, how can the WGN Database be suitably promoted and communicated to the broader global audience? The secretariat welcomes further discussions and direction from the Group of Experts in this regard.

## **V. Summary and Call for Action**

37. With the revised WGN Database, the Group of Experts has a tool with which to collate and manage toponymic data at a global scale. Now the global geographic names community can work to populate the Database with authoritative, multilingual, multi-scriptural and geo-referenced toponymic data across many geographical dimensions. In the coming months after this present session, the Secretariat will integrate authoritative data from across the United Nations system with the view to capture all major cities in the six official UN languages, completing existing data on 193 countries. Beyond this immediate future, there is a significant and realisable opportunity to use the WGN Database as a means of respecting and promoting the value of geographical names as significant elements of cultural heritage, language and identity, including those of indigenous peoples and regional and minority groups. This is an exciting development for the Group of Experts and calls for enhanced engagement within the Geographical Names community and beyond.

## **VI. Points for Discussion**

38. **The Group of Experts is invited to:**

- (a) Take note of the present report and express its views on the redevelopment and future direction of the World Geographic Names Database;**
- (b) Welcome and invite Member States and their National Names Authorities to submit and update toponymic data to the secretariat;**
- (c) Consider and provide further guidance regarding:**
  - i. The development of a unique identifier for cities;**

---

<sup>18</sup> see [GEGN.2/2023/73/CRP.73](#) Report of the Working Group on Geographical Names Data Management for 2021 and 2022

<sup>19</sup> Working Group on Geographical Names Data Management of UNGEEN <https://wiki.gdi-de.org/display/wgtdfg/Working+Group+on+Geographical+Names+Data+Management+of+UNGEEN>

**ii. An appropriate governance mechanism for the World Geographic Names Database, inclusive of guidance on the Database's promotion and communication;**

**(d) Express its views on how Linked Data and the semantic web can be integrated into the World Geographic Names Database**

## Annex 1 The example of Antananarivo, Madagascar

	Data Header	Arabic	Chinese	English	French	Russian	Spanish	Endonyms		
<b>Region</b>	<ul style="list-style-type: none"> <li>M49: 22</li> <li>ISO3</li> <li>Geography (Bounding Box / Coordinates)</li> </ul>	شرق أفريقيا	东非	East Africa	Afrique de l'Est	Восточная Африка	África Oriental			
<b>Country</b>	<ul style="list-style-type: none"> <li>M49: 450</li> <li>SKOS</li> <li>ISO3: MDG</li> <li>Geography (Bounding Box)</li> </ul>	مدغشقر	马达加斯加	Madagascar	République de Madagascar	Мадагаскар	Madagascar	Malagasy	Madagasikara	Repoblikan'i Madagasikara
<b>City</b>	<ul style="list-style-type: none"> <li>No current identifier</li> <li>Geography (x- and y-coordinates)</li> </ul>	أنتاناناريفو	塔那纳利佛	Antananarivo	Antananarivo	Антананариву	Antananarivo	Antananarivo	Antananarivo	

## Annex 2

### WGN API Endpoints

#### Cities

- GET /API/GetAllCities
- GET /API/GetCity/{ID}
- GET /API/GetAllCities/{toponymTypeID}
- GET /API/GetCity/{ID}/{toponymTypeID}
- GET/API/GetAllCities/GetAllCapitals
- GET /API/GetCitiesWithSearch/{pageSize}/{pageNumber}/{searchBy}/{searchName}
- GET /API/GetCitiesDetails/{id}
- GET /API/GetCitiesNameDetails/{id}

#### Countries

- GET /API/GetAllCountries
- GET /API/GetCountry/{ID}
- GET /API/GetAllCountries/{toponymTypeID}
- GET/API/GetCountry/{ID}/{toponymTypeID}
- GET /API/GetCountriesWithSearch/{pageSize}/{pageNumber}/{searchName}
- GET /API/GetCountryDetails/{m49Code}
- GET/API/GetCountryNameDetails/{m49Code}

#### Geographies

- GET/API/GetAllGeographies
- GET /API/GetGeography/{ID}
- GET /API/GetAllGeographies/{toponymTypeID}
- GET /API/GetGeography/{ID}/{toponymTypeID}
- GET /API/GetGeographiesWithSearch/{pageSize}/{pageNumber}/{searchName}
- GET /API/GetGeographyDetails/{id}
- GET /API/GetGeographyNameDetails/{id}

#### Languages

- GET /API/GetAllLanguages/{ID}
- GET /API/GetAllLanguages

#### Romanisation Systems

- GET /API/GetAllRomanisationSystems/{ID}
- GET /API/GetAllRomanisationSystems

#### Temporal

- GET /API/Temporal/GetTemporalCountryById/{id}
- GET /API/Temporal/GetAllTemporalCountries
- GET /API/Temporal/GetTemporalCityById/{id}
- GET /API/Temporal/GetAllTemporalCities
- GET /API/Temporal/GetTemporalGeographyById/{id}
- GET /API/Temporal/GetAllTemporalGeographies

## Visualisation

- GET /API/Visualisation/GetAllCities
- GET /API/Visualisation/GetAllCountries
- GET /API/Visualisation/GetAllGeographies
- GET /API/Visualisation/{toponymTypeID}/GetAllCountries
- GET /API/Visualisation/{toponymTypeID}/GetAllCities
- GET /API/Visualisation/{toponymTypeID}/GetAllGeographies