United Nations Group of Experts
on Geographical Names (UNGEGN)

GEOGRAPHICAL NAMES

INFORMATION BULLETIN

No. 69 NOVEMBER 2025

The use of artificial intelligence in the standardization of geographical names

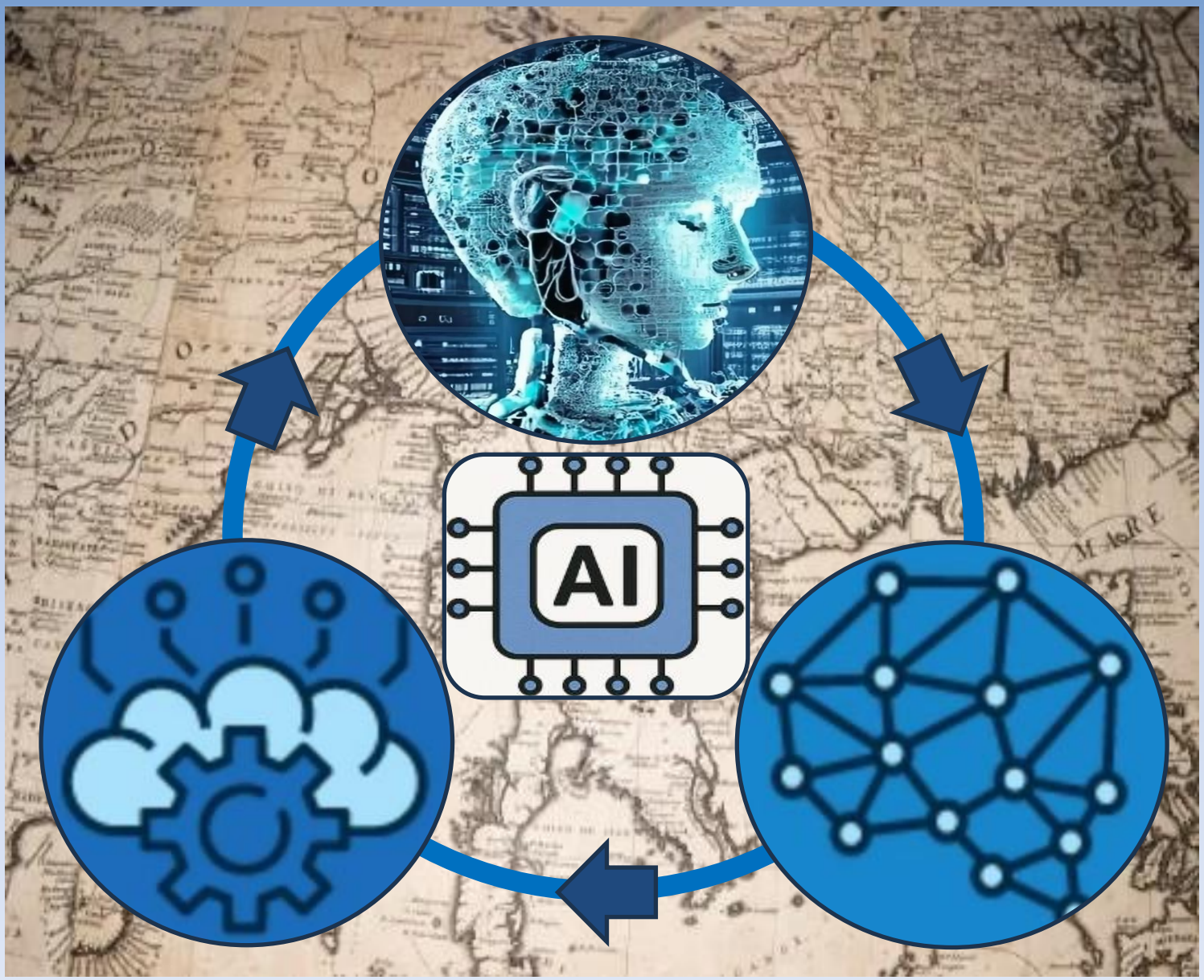


Table of Contents

MESSAGE FROM THE CHAIRPERSON

« Tenir Compte Des Progres Scientifiques »

Pierre Jaillard (France)
Président du GENUNG
***Mél**: pierre@jaillard.net*

Chers Collègues,

La dernière session du GENUNG, du 28 avril au 2 mai 2025, a été doublement significative. Au plan quantitatif, elle a réuni 244 experts de 64 pays pour discuter sur la base de 166 rapports, malgré la réduction de 10 % des moyens alloués cette année dans les circonstances budgétaires que connaissent les Nations unies comme beaucoup de nos États. Au plan qualitatif, de nombreuses interventions en séances plénières et discussions en événements parallèles ont de nouveau manifesté l'importance des enjeux techniques, économiques, sociaux et politiques des noms géographiques.

Ces enjeux, le thème de la session, et même la mission que le Conseil économique et social des Nations unies (ECOSOC) nous a donnée, appellent notre groupe d'experts à « tenir compte des progrès scientifiques réalisés en linguistique et dans les techniques de traitement et de production de données toponymiques » (Règlement intérieur, II, 4, a). Le faisons-nous assez ? En matière de « techniques de traitement et de production de données toponymiques », la dernière session et le présent *Bulletin* montrent notre attention aux derniers progrès réalisés. Mais en ce qui concerne les enjeux politiques ?

Les Nations unies ont naturellement tendance à envisager les questions politiques au prisme de leurs principes fondateurs : un système international fondé sur la juxtaposition territoriale de souverainetés nationales absolues et exclusives. Ce modèle ne permet toutefois pas de rendre compte du fonctionnement des exonymes, où des couches de souveraineté se superposent sur des plans différents : alors qu'un État se voit reconnaître la compétence juridique pour normaliser les noms de lieux de son territoire (endonymes), d'autres peuples peuvent employer dans leur langue des noms différents (exonymes).

Il est donc particulièrement significatif que ce soit un organe politique des Nations unies, l'ECOSOC, qui nous appelle à fonder scientifiquement nos décisions. C'est ce qui les rendra véritablement universelles et pérennes, au-delà du cadre traditionnel du système international. Car, comme l'a si bien exprimé, paradoxalement peut-être, une personnalité religieuse, le feu pape François, « la réalité est supérieure à l'idée ». Et la linguistique, comme notre propre expérience, nous apprend que l'usage prévaut toujours sur les normes extérieures dont on voudrait encadrer la langue.

Voilà une tâche enthousiasmante mais exigeante, dont il nous faut nous montrer dignes par la lucidité et par l'humilité.

"To be based on the achievements of science"

Pierre Jaillard (France)
Chair, UNGEGN
Email: pierre@jaillard.net

Dear Colleagues,

The recent session of UNGEGN, from 28 April to 2 May 2025, was doubly significant in two ways. In quantitative terms, it brought together 244 experts from 64 countries who engaged in discussions based on 166 reports submitted by our global community, despite a 10% reduction in resources, reflecting the budgetary constraints facing both the United Nations and many of our countries. Qualitatively, the many interventions during plenary sessions and discussions in side events once again demonstrated the technical, economic, social, and political importance of geographical names.

These issues, the theme of the session, and even the mission that the United Nations Economic and Social Council (ECOSOC) has given us, call on our group of experts to "be based on the achievements of science in relation to both language treatment and the technical means of processing and generating toponymic data" (Rules of Procedure, II, 4, a). Are we doing so sufficiently? With regard to the "technical means of processing and generating toponymic data", the recent session and this *Bulletin* reflect our attention to the latest progress. But what about political issues?

The United Nations naturally tends to see political issues through the lens of its founding principles: an international system based on the territorial juxtaposition of absolute and exclusive national sovereignties. However, this model is challenged by the way exonyms work, with overlapping layers of sovereignty at different levels: while a Country has the legal right to standardize place names within its borders (endonyms), other peoples may use different names (exonyms) in their own languages.

It is therefore particularly significant that a political body of the United Nations, ECOSOC, calls on us to ground our decisions in science. This is what will make them truly universal and enduring, extending beyond the traditional framework of the international system. For, as it was so aptly, perhaps paradoxically, expressed by a religious figure, the late Pope Francis, "reality is greater than ideas". And linguistics, like our own experience, teaches us that usage always prevails over the external norms that some may seek to impose on language.

It is an exciting but demanding task, and one we must approach with clarity of thought and humility, if we are to prove ourselves worthy of it.

MESSAGE FROM THE SECRETARIAT

Mark Iliffe
UNGEGN Secretariat
Email: mark.iliffe@un.org

Dear UNGEGN Experts,

The 69th issue of the UNGEGN Bulletin features articles on the theme "*The use of artificial intelligence in the standardization of geographical names*". It also discusses the outcomes of the 2025 UNGEGN session and covers the usual featured sections – from the countries, divisions, working groups and special projects, and other news items.

Featured Theme

In a world where change is happening faster than ever, this issue of the Bulletin takes a closer look at how new scientific and technological developments are shaping our work. One key focus is the growing connection between Artificial Intelligence (AI) and the global standardization of geographical names. AI offers exciting possibilities—it can help make geographical names more accurate, easier to access, and more efficiently managed. It also opens doors to tackling complex challenges related to language, history, and cultural diversity. In this issue, you'll find highlights that explore these opportunities and showcase innovative approaches from around the world. Some highlights include:

- **How AI is supporting the analysis and extraction of historical texts**: Through Natural Language Processing (NLP) tools, experts are extracting names from historical documents and link them to local context and heritage, which is crucial for standardization. It can systematically analyze historical gazetteers, maps, and administrative records that might otherwise be underutilized due to complexity. For instance, Claude Sonnet 4 was used to extract, standardize, and analyze toponymic data from J.W. McGuire's 1925 Geographic Dictionary of the Virgin Islands, handling multilingual content and complex scholarly citations.
- **Improving Multilingual Recognition and Romanization** that is facilitating the romanization of non-Latin scripts, aiding linguistic inclusivity. This is critical for languages with complex linguistic landscapes like Arabic, which present challenges due to its rich morphological system, pharyngealized consonants, and dialectal variations. In turn, this demonstrates demand for research into innovative methods for romanization, including machine transliteration.
- **Audio File Generation for Pronunciation**: AI can generate audio files for geographical names, particularly exonyms, to provide guidance on correct pronunciation, addressing a key challenge in cross-cultural communication. For example, the Secretariat of the Nomenclature Commission (SNK) of the Czech State Administration of Land Surveying and Cadastre (ČÚZK) successfully integrated AI-generated MP3 audio files for Czech exonyms into their "Names of the World" web application.
- **Efficiency and Consistency in the Standardization Process:** New Zealand demonstrates how AI is being investigated to analyze large numbers of public submissions received during geographical name proposals, aiming to ensure consistent and objective analysis, reduce timeframes and costs, and utilize new technology. A test with Claude AI on the "Petone to Pito One" renaming proposal showed its capability to identify numbers of submissions for/against, alternative suggestions, and main themes, with an expected accuracy over 90% compared to manual analysis.

But these are just a sample of the many diverse cases and national experiences on the theme discussed within this issue of the Bulletin.

Ongoing UNGEGN Activities

Revitalization of Working Groups
Since the 2025 session, each of UNGEGN's Working Groups has worked to develop Terms of Reference, Work Plans, and other administrative tools to support their operation. Member States and Experts in the area of

geographical names are invited to contact the Secretariat and the Convenors of UNGEGN's Working Groups to express their interest in participating in advancing UNGEGN's normative work.

Strengthening of Divisions

The next meeting of the Divisions is scheduled for 9 December 2025. Among the many topics of concern, this meeting is expected to discuss the revitalization of Divisions, guidelines for Division Creation and Dissolution, and the responsibilities of Divisional Chairs, all of which have been drafted with the support of Member States and the Secretariat. Member States are invited to indicate their interest in participating in this meeting and in their respective Divisions.

Communications and Engagement

The Group of Experts now has a new online presence on LinkedIn: https://www.linkedin.com/company/UNGEGN. We will be sharing more 'ephemeral' content on LinkedIn, including resources and the history of UNGEGN, as well as highlighting events in which UNGEGN Experts are participating.

We thank all our contributors to this issue, and to Andreas Hadjiraftis of Cyprus for designing the front page.

The *Information Bulletin of the United Nations Group of Experts on Geographical Names* (formerly UNGEGN Newsletter) aims to showcase and promote good practices from across the global geographical names community.

The 'Bulletin' is published by the UNGEGN Secretariat, United Nations Statistics Division, Department for Economic and Social Affairs, United Nations Secretariat. Contributions and reports are received from the Experts of the Group, its Linguistic/Geographical Divisions, and its Working Groups. Editorial guidance is provided by the UNGEGN Working Group on Publicity and Funding.

Contributions for the Bulletin can only be considered when they are made available digitally in Microsoft Word or a compatible format. They should be sent to the following email address: geoinfo_unsd@un.org

United Nations Group of Experts on Geographical Names Information Bulletin (ISSN 1014-798) is published by the United Nations Statistics Division, Department of Economic and Social Affairs. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Previous issues of the Bulletin (formerly Newsletter) can be found at https://unstats.un.org/unsd/ungegn/pubs

Secretariat of the Group of Experts on Geographical Names (UNGEGN)
Statistics Division
United Nations
New York, NY 10017, USA

The 2025 Session in Review

Mark Iliffe
UNGEGN Secretariat
Email: mark.iliffe@un.org

Participants of the 2025 session

The 2025 session of the United Nations Group of Experts on Geographical Names (UNGEGN) was convened on 28 April to 2 May 2025 on the theme "***Advancing geographical names standardization through inclusive, culturally-informed and evidence-based solutions to support sustainable development***". In all, 197 participants representing 56 Member States across UNGEGN's geographical and linguistic divisions and observers participated. Opening the session, Ms Bjørg Sandkjær, Assistant Secretary-General for Policy Coordination and Inter-Agency Affairs, DESA highlighted that less than five years remain to meet the 2030 Agenda for Sustainable Development, and emphasized that standardized geographical names are essential to achieving the SDGs, observing that "far from being mere labels, place names underpin data collection, analysis, and decision-making that shape sustainable outcomes worldwide".

Member States shared national progress in several "National Dialogues" segment, sharing innovative good practices in geographical names management with updates from Australia, Austria, Brazil, Brunei Darussalam, Canada, Cyprus, Czechia, Denmark, Estonia, Hungary, Iceland, India, Italy, Japan, Malaysia, Mexico, New Zealand, Norway, Poland, Republic of Korea, Russian Federation, Saudi Arabia, Slovenia, and Sweden. These contributions demonstrated the breadth of national approaches to standardization, preservation, and innovation. Further, the Committee of Experts on Global Geospatial Information Management (UN-GGIM) reaffirmed its strong collaboration with UNGEGN, in recognizance that geographical names as a Global Fundamental Geospatial Data Theme, acknowledging the work done as part of its "Collaboration Project: good practices of National Mapping Agencies and National Geographical Names Authorities" and welcoming further updates at the forthcoming fifteenth session of UN-GGIM in August this year. Further, the adoption of a Universal Unique Identifier for Cities marked a significant advance in harmonizing geospatial data infrastructure and updates on the World Geographical Names Database highlighted how Member States can provide their geographical names to the United Nations across languages and scripts.

Interactive panels throughout the session addressed the cultural, linguistic, and developmental dimensions of geographical names, underscoring their role in fostering inclusive, culturally informed, and evidence-based development approaches. Notably, a side event on "Geographical names as cultural heritage" featured a keynote by Ms Aluki Kotierk, Chair of the UN Permanent Forum on Indigenous Issues and ex-President of Nunavut Tunngavik Incorporated (NTI - the representative body of the Nunavut Inuit), who emphasized protecting names as living expressions of identity within and across borders and generations.

A highlight of the week was the special session on Artificial Intelligence (AI) and the future of geographical

names standardization, including discussions on generative AI applications for Arabic geographical names, illustrating how emerging technologies can support consistency, romanization, linguistic respect, and national capacity development. Side events and discussions reinforced the role of naming in recognizing minority and Indigenous languages and promoting participatory governance, while also examining tools such as the World Geographical Names Database and country-led innovations in legal and institutional frameworks.

From updates on national naming policies to deep dives into language recognition, cultural heritage, and the role of standardization in sustainable development, the 2025 session of UNGEGN showcased the richness and relevance of geographical names work. It also highlighted innovative tools, celebrated cross-country and interdisciplinary collaboration, and reaffirmed a shared global commitment to inclusive, culturally respectful, and evidence-based naming practices as part of a sustainable and equitable future. Guided by the mandates adopted in its closing session, UNGEGN now "gets back to work" and will meet again in person in May 2027 for its 2027 Session.

Following the formal end of the 2025 session, participants expressed their sincere thanks to Ms. Cecille Blake for her significant contributions since 2014. Mr Mark Iliffe, has assumed the role of UNGEGN Secretariat. Mark is an English national with over 15 years of experience in developing geospatial capacity in both developed and developing countries. In his previous role, he supported the UN Committee of Experts on Global Geospatial Information Management (UN-GGIM), serving as Secretary to the UN Expert Group on the Integration of Statistical and Geospatial Information, and also supported other matters related to global geospatial information management, including climate resilience and geospatial information for the Sustainable Development Goals. His email is mark.iliffe@un.org

Further information for the 2025 session, including documentation, webcasts, and other resources, can be found here: https://unstats.un.org/unsd/ungegn. Also, all staff are encouraged to 'follow' UNGEGN on LinkedIn here: https://www.linkedin.com/company/UNGEGN. This platform communicates the work, history and information on geographical names, and all are invited to participate.

THEME: The use of Artificial Intelligence (AI) in the Standardization of Geographical Names

The use of Artificial Intelligence (AI) in the Standardization of Geographical Names of Cyprus

Andreas Hadjiraftis
President of the Permanent Committee for the Standardization of Geographical Names of Cyprus
Chief Lands Officer, Survey/Cartography/Geodesy/Hydrography/Photogrammetry,
Department of Lands and Surveys, Ministry of Interior,
Cyprus
Email: ahadjiraftis@dls.moi.gov.cy

Introduction

Artificial intelligence (AI) refers to technologies that enable computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy. Applications and devices equipped with AI can see and identify objects. They can understand and respond to human language. It offers substantial economic benefits by creating new business models and job opportunities. Also, AI helps in increasing productivity, automating tasks, and reducing costs through more efficient operations and data-driven insights.

In this article, we attempt to analyze how AI could be used in the standardization process of the geographical names of Cyprus.

Historical Background

Cyprus, a small island in the Eastern Mediterranean, has one of the most historically layered toponymic landscapes in Europe. Over more than 10,000 years, waves of settlement and conquest, from Mycenaean Greeks to Romans, Byzantines, Lusignans, Venetians, British and others, have left their marks on the island's place names. Today, these toponyms reflect a unique blend of Greek, Latin, English, and other influences, often with multiple names and variants for the same locations.

The standardization of these geographical names is vital for consistent use in administration, mapping, education, diplomacy, and emergency services. However, the task is technically complex and politically sensitive. Artificial Intelligence, with its ability to process language, analyze data, and identify patterns, can now assist in harmonizing Cyprus's toponyms, while also respecting their historical, cultural, and linguistic significance.

Linguistic Complexity of Cypriot Geographical Names

Cypriot geographical names reflect millennia of cultural contact and political change. Ancient cities such as Alamís, Kítion (modern Lárnaka), and Amathoús, carried Hellenic names that survived Roman rule, albeit sometimes Latinized. During the medieval period, Lusignan and Venetian rulers introduced Western European naming conventions. Over time, Greek geographical names evolved further, while British rule (1878–1960) led to widespread use of English transliterations, often inconsistently applied.

As a result, many Cypriot locations have multiple names, both historic and modern. For example, the capital is officially Λευκωσία (Lefkosía) in Greek, but is internationally known as "Nicosia". Similarly, the city of Αμμόχωστος (Ammóchostos) is also known as "Famagusta", an exonym. Such variation complicates mapmaking, legal documents, navigation systems, and multilingual communication.

Efforts to standardize toponyms have been underway for decades, notably through the Republic of Cyprus's Permanent Committee for the Standardization of Geographical Names, in alignment with the UNGEGN (United Nations Group of Experts on Geographical Names). However, the manual nature of this work poses scalability and consistency challenges, which is where AI can contribute significantly.

Figure 1: Linguistic Complexity of Cypriot Geographical Names

ELOT Standard No. 743: A Foundation for Transliteration

Central to toponymic standardization in the Greek-speaking world is ELOT 743, the official Greek-to-Latin transliteration system developed by the Hellenic Organization for Standardization. Adopted by both the Greek and Cypriot governments and endorsed by the United Nations, this system ensures consistency when converting Greek-script geographical names into Latin script.

For example:

- Λευκωσία transliterates to Lefkosía
- Πάφος transliterates to Páfos
- Λεμεσός transliterates to Lemesós

Figure 2: AI and ELOT Standard 743 of *Hellenic Organization for Standardization*

ELOT 743 ensures that transliterations are accurate, reversible, and appropriate for official use. Despite this, older or widely used forms exonyms, such as "Nicosia", remain common in global usage, pointing to the need for broader enforcement and digital validation, a role well suited for AI.

The Role of Old Maps, Land Registers, and the Department of Lands and Surveys

Cyprus's toponymic legacy is documented, not only in contemporary use, but also in historic cartographic and administrative records. The Department of Lands and Surveys (DLS) of Cyprus maintains an invaluable archive of old maps, cadastral plans, and land registers that document the evolution of place names over time.

These records:

- Preserve historical toponyms used during different administrations (e.g. Greek, Roman, Byzantine, Lusignan, Venetian, British, and early independence),
- Document spelling and boundary changes through official surveying practices,
- Provide legal and spatial continuity for land ownership and place identification,
- Serve as authoritative references in disputes over land and name usage.

Incorporating these data sources into AI-driven systems enhances the reliability of name standardization by grounding it in verified historical and legal records. For example, AI can use old cadastral maps to cross-reference ancient toponyms with their modern equivalents, while spatial databases can help validate the geographical coordinates and contextual use of place names over time.

The DLS's digitization efforts, such as the Cyprus Integrated Land Information System, already provide structured geospatial data. AI can leverage these datasets to identify patterns, infer name changes, and fill gaps where manual records are incomplete.

AI Solutions for Toponymic Standardization in Cyprus

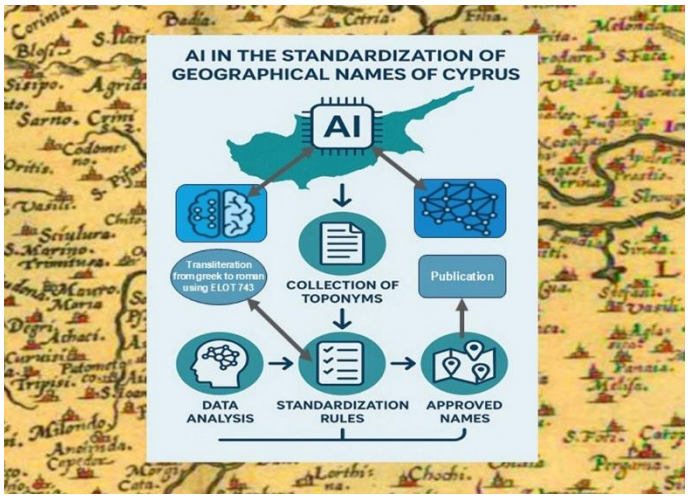


Figure 3: Use of AI in the Standardization of Geographical Names of Cyprus

AI tools can significantly support toponymic work in Cyprus in the following ways:

1. *Transliteration Engines Based on ELOT 743*

   AI can convert Greek-script names into ELOT-compliant Latin forms and flag inconsistencies in existing documents or databases.

2. *Geospatial Reconciliation Using DLS Data*

   AI can align historical names from DLS maps and registers with current geospatial layers, ensuring accurate historical continuity in naming.

3. *Automated Data Cleaning and Standardization*

   AI can detect duplicated or outdated toponyms in public records and suggest unified, standardized versions.

4. *Digitization and Interpretation of Historical Records*

   Through OCR (Optical Character Recognition) and AI-assisted text recognition, names in old manuscripts and maps can be digitized and linked to modern locations.

5. *Public-Facing Tools for Education and Access*

   AI-powered platforms could help users search and understand historical and authoritative place names, including variant forms and legal spellings.

Conclusion

Consistent geographical naming supports everything from emergency services to education, tourism, and diplomacy. In a politically complex environment like Cyprus, it also fosters transparency, mutual understanding, and administrative order. Standardization through AI, especially when enriched by official data from the DLS, ensures accuracy, cultural respect, and efficiency.

Cyprus's geographical names are a product of its unique history and multicultural heritage. Standardizing these names is not merely a technical task but a cultural responsibility. The ELOT 743 transliteration system, the rich cartographic and cadastral archives of the Department of Lands and Surveys, and modern AI technologies together offer a powerful framework to support this endeavor.

Artificial intelligence, grounded in official records, linguistic standards, and spatial analysis, can bridge the past with the present, ensuring that Cyprus's authoritative geographical names are preserved, respected, and presented accurately in the digital era.

Leveraging AI for Enhanced Accessibility: Generating Audio Files for Czech Exonyms

Michal Zídek
Secretariat of the Nomenclature Commission,
Czech State Administration of Land Surveying and Cadastre (ČUZK)
Czechia
Email: michal.zidek@cuzk.gov.cz

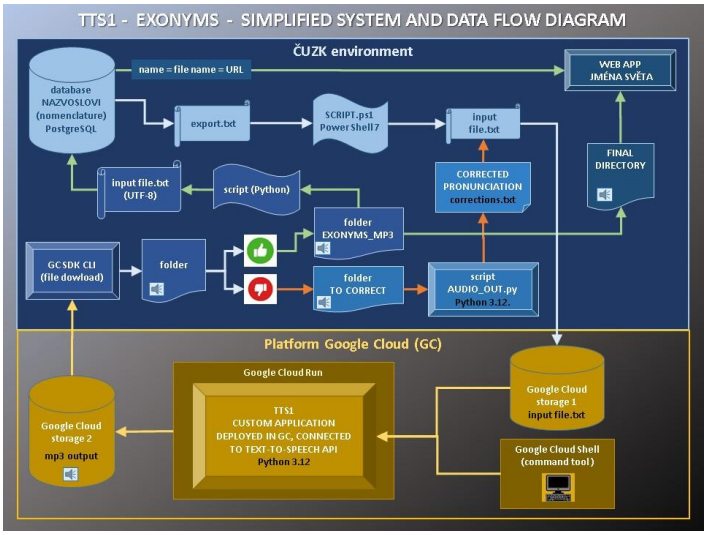
Resolution 11 of the Eighth UNGEGN Conference stated that the transliteration of geographical names from one script to another, such as romanization, does not normally provide a guide to the correct pronunciation of such names by persons who are unacquainted with the source language.[1] To address this, the past UNGEGN Working Group on Pronunciation's work plan aimed to "further the inclusion of pronunciation of individual geographical names in national gazetteers, through International Phonetic Alphabet (IPA) files and/or audio files" and to "provide support for the collection of audio files for the UNGEGN Geographical Names Database".[2] In alignment with these objectives, the Secretariat of the Nomenclature Commission (SNK) of the State Administration of Land Surveying and Cadastre (ČUZK) selected a collection of standardized Czech exonyms published in the List of Czech Exonyms as the starting set for leveraging an AI large language model (LLM) to generate audio files of Czech toponyms.

The SNK examined several methods for audio file generation. The traditional approach of creating the audio files in a recording studio with an expert native speaker was deemed impractical due to prohibitive costs, significant technical complexity and extended development timelines. Similarly, training a custom, on-site large language model (LLM) was dismissed due to implementation complexities and concerns regarding its effectiveness for exonyms. LLM models acquire predictive power regarding syntax, semantics, and ontologies inherent in human language corpora[3], but they also inherit inaccuracies and biases present in the data they are trained on[4]. Consequently, training an LLM on a relatively small dataset of exonyms to pronounce those same often-difficult words would not efficiently leverage the model's predictive capabilities. After examining various AI platforms, we chose Google Cloud Platform (GC) for its low cost, low error rate and capacity to leverage pre-built modules without requiring local, supporting AI infrastructure. Project emphasis was on cost efficiency, easy customization, simplicity of execution and quality of output.

We designed a custom Google Cloud (GC) project to leverage the GC Text-to-Speech API (TTS API) for converting simple text files into digital audio. The TTS API is a programming interface connecting to Google's cloud service, which utilizes advanced artificial intelligence to convert text into natural-sounding speech. The core of our project is a light-weight application TTS1 written in Python 3.12 and deployed in GC Cloud Run environment. This application is able to read a specified text file format as an input, interact with the TTS API, and generate the corresponding audio files.

```
Aara,Áára
Benghází,Benghhází
Chartúm,Charr-tuům
Cušima,Tsušima
Džadída,Džadýda
Freiberská Mulda,Frajberrská Mulda
Glognice,Gloggnice
Hadžar,Hadžarr
Kafarnaum,Kaffar naumm
Kesong,Kessong
Kjúšú,Kjuůšuů
Méghálaj,Meég-haálaj
```

By default, a TTS engine primarily relies on orthographic rules, utilizing an internal dictionary and a set of rules to map letter sequences (orthography) to sound sequences (phonemes) within a given language. We exploited this characteristic to address problematic exonyms – those that the TTS algorithm would not correctly pronounce from their original orthographic form. Our solution involves formatted text files where each row contains two comma-separated text strings. The first string is used exclusively to name the output audio file, ensuring consistency with our database. The second string, which the TTS processes for sound generation, is modified for problematic exonyms. While these two strings are identical for correctly pronounced exonyms, for unacceptable pronunciations, the second string is altered into a pseudo-orthographic form that guides the TTS towards the correct pronunciation. In addition, the application encodes metadata directly into the mp3 files, including the original name and the orthographic or the pseudo-orthographic form used to generate the sound.



Our workflow begins by exporting a list of exonyms from our PostgreSQL database, with each exonym on a separate row. A Python script then processes this file, duplicating each row's text string and inserting a comma between the two, before saving it as a UTF-8 encoded text file. This file is uploaded to GC storage, where it is available as an input for the application. The application is initiated via a curl command, using GC command tool Cloud Shell. It processes the text file, generating MP3 audio files that are saved to another GC storage, from where they are downloaded to a local computer. The downloaded audio files undergo an initial examination, and are categorized as either acceptable or unacceptable. Another Python script is used to extract the names of the unacceptable files stored in a folder into a new, double string input file. This file is then manually edited in a text editor, where the second string of each problematic exonym is modified to trigger the correct pronunciation. The revised input file is then re-processed by the TTS1 application, and the output is re-examined. This iterative cycle continues until all exonyms achieve correct pronunciation.

The code of the TTS1 application allows for the precise specification of voice, language, speed, and pitch. For this project, we selected the cs-CZ-Wavenet-B voice, applying a pitch of -2.8 and a slightly reduced speaking rate of 0.9. MP3 was chosen as the output format due to its universal compatibility, acceptable sound quality (at its highest setting), and minimal file size. The audio files were generated at a 48000Hz sampling rate and saved at the highest quality, ensuring optimal fidelity. Generated MP3 file size averages 14KB, with the entire exonym set occupying only 35MB. On the initial pass, the rate of unacceptably pronounced exonyms was 12%. While the initial generation of 2623 mp3 files took only ten minutes, the refinement process for the 308 rejected files required approximately one month. Remarkably, the monthly operational cost of this project within Google Cloud is approximately 0.5 USD.

The generated MP3 audio files for the Czech exonyms have been successfully integrated into the ČUZK web application *Jména světa* (Names of the World). They are now publicly accessible for listening on the application's Czech Exonym Index page: https://ags.cuzk.cz/jmenasveta/#en.

In a future development, we aim to incorporate SSML (Speech Synthesis Markup Language) for enhanced control over synthesized speech and IPA phonetic notation for improved pronunciation accuracy. Subsequently, we plan to expand the AI-generated audio collection to include toponyms from our publications Czech Names of Seas and International Territories and Names of States and their Territorial Parts. These will also be published in the Names of the World web application. Ultimately, we aim to apply these methods to generate audio for all 319,623 toponyms of the Czech Republic contained within our Geonames database.

References

[1] United Nations. (2003). *Eighth United Nations Conference on the Standardization of Geographical Names, Berlin, 27 August-5 September 2002*. (E/CONF.94/3). United Nations Publication, Sales No. E.03.I.14. New York: United Nations.

[2] UNGEGN. (2025). *Working Group on Pronunciation* - Retrieved from https://unstats.un.org/unsd/geoinfo/ungegn/wg9.html . The functions of this Working Group are now shared across the Working Groups for Geographical Names as Cultural Heritage and Geographical Names Data Management. See E/2017/80.

[3] Fathallah, N., et al. (2024). *NeOn-GPT: A Large Language Model-Powered Pipeline for Ontology Learning*. Extended Semantic Web Conference 2024, Hersonissos, Greece.

[4] Manning, C. D. (2022). Human Language Understanding & Reasoning. *Daedalus*. Available at https://www.amacad.org/daedalus.

Linking QR Codes to Historical Maps for Place Name Pronunciation and Background Information

Anahita Omidi, Rahele Alizadeh
Iranian National Committee on the Standardization of Geographical Names
Iran (Islamic Republic of)
Email: alizade.carto@gmail.com

Introduction

Quick Response (QR) codes are two-dimensional barcodes capable of encoding various types of information, such as text, URLs, or other data. Originally developed in 1994 by Denso Wave, a Japanese company, for tracking automotive parts. QR codes have since evolved into a versatile tool employed across multiple domains, including marketing, education, and tourism (Ashrafi et al. 2025). When scanned using a smartphone or a dedicated QR code reader, these codes can direct users to websites, display textual content, or trigger specific actions. In the context of historical maps, QR codes offer a novel method to augment user experience by integrating additional layers of information and interactivity. By strategically placing QR codes adjacent to place names or points of interest, users can access audio pronunciations or concise historical summaries, thereby enhancing accessibility and engagement with historical data. This article examines the benefits, practical implementation, and illustrative examples of this approach, supported by relevant scholarly references.

Benefits of QR Codes on Historical Maps

The incorporation of QR codes into historical maps yields several significant advantages. One primary benefit is their ability to facilitate the correct pronunciation of place names, which is often challenging due to linguistic evolution or regional variations. For instance, the English city of Worcester is pronounced "WUSS-ter," a detail not immediately apparent from its spelling. Similarly, Edinburgh is articulated as "ED-in-bur-uh," which may confound non-native speakers. QR codes can link to audio files that provide accurate pronunciations, thereby improving user comprehension and cultural appreciation.

Another key advantage is the provision of immediate access to historical information. Rather than requiring users to consult external resources, a QR code can deliver a succinct summary directly to their device. For example, a QR code on a map of ancient Rome might link to a brief overview of the Colosseum's construction and historical significance, offering contextual depth without interrupting the user's exploration. This feature is particularly valuable for tourists and students seeking efficient learning tools (Vuksanović et al. 2021).

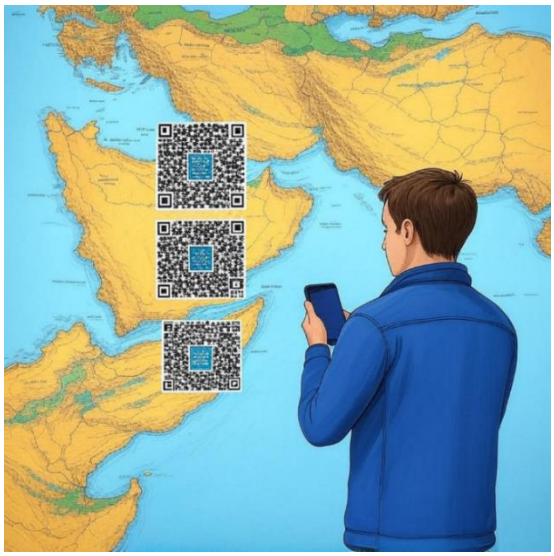
Additionally, QR codes enhance the interactivity and engagement of historical maps by integrating multimedia content, such as audio or video. This dynamic approach caters to diverse learning preferences, making history more accessible and appealing, especially to younger audiences. Furthermore, QR codes can preserve historical knowledge by digitizing and disseminating it widely, ensuring its availability beyond physical archives. They also support multilingual content, broadening their reach, and can be updated easily to reflect new findings without necessitating map reprints.

Practical Implementation

Implementing QR codes on historical maps requires a methodical approach to integrate digital enhancements effectively while preserving the map's primary function. The process begins with developing supplementary content tailored to diverse learning preferences, such as audio recordings of place name pronunciations verified by linguists or native speakers, concise textual summaries of historical contexts written in accessible language, and multimedia resources like videos or interactive web pages that provide immersive experiences, including virtual tours or historical reenactments. This content must be securely hosted online through reliable platforms, such as institutional websites or cloud storage, to ensure stable QR code linkage. Rigorous validation by historians or subject

matter experts is essential to maintain factual accuracy and contextual relevance. Once the content is finalized, QR codes are generated using accessible tools like QR Code Generator or GoQR.me and strategically positioned on the map near relevant geographical landmarks or points of interest.

Design considerations prioritize visibility, ensuring codes do not obscure map features; appropriate sizing to balance scannability and aesthetic harmony; and the inclusion of brief user guidance to assist individuals unfamiliar with QR technology. For instance, a historical battlefield map might feature QR codes at locations such as Little Round Top, directing users to audio narratives detailing pivotal events. To ensure functionality, codes should undergo testing across multiple devices and various environmental conditions, with contingency plans in place for low-connectivity areas, such as offering downloadable offline content packages. This multifaceted strategy ensures that technological integration enhances, rather than distracts from, the educational and historical value of the map.



Examples in Practice

The application of QR codes on historical maps is already evident in various contexts. Museums, such as the British Museum, employ QR codes to enrich visitor experiences by linking artifacts to audio guides or videos, a practice that has been shown to increase engagement (de Morais Sarmento, 2024). For example, a QR code beside an ancient vase might provide its historical context via an audio narration.

In educational settings, QR codes on historical maps may offer interactive learning opportunities. A map of ancient Greece could link to a quiz on Greek mythology or a video about the Peloponnesian War, enhancing student retention and interest.

Tourism also benefits from this technology. The Tourism Board of Rome has integrated QR codes into historical maps, enabling visitors to access virtual tours of landmarks such as the Roman Forum, complete with 3D reconstructions and commentary. These examples illustrate the versatility and effectiveness of QR codes in historical contexts.

Conclusion

The integration of QR codes into historical maps represents a significant advancement in historical education and exploration. By linking to audio pronunciations and concise historical summaries, this approach enhances user understanding and engagement while preserving and disseminating knowledge. Practical implementation is straightforward, requiring careful content creation, strategic placement of QR codes, and a user-focused design. Supported by real-world applications and scholarly evidence, QR codes have the potential to transform historical maps into interactive, accessible tools for learning and discovery.

References

Ashrafi, D. M., Iskender, A., Shahid, T. (2025). Bytes to bites: Investigating QR code menu use behavior and green satisfaction in the restaurantscapes through a hybrid PLS-SEM and machine learning approach. Journal of Foodservice Business Research, 1-47. https://doi.org/10.1080/15378020.2025.2466873

de Morais Sarmento, M. L. F. (2024). Design of Immersive Museum Experiences: Improving Visitor Engagement Through Co-Creation, Augmented Reality and Gamification (Master's thesis, Universidade do Porto (Portugal)).

Vuksanović, N., Demirović Bajrami, D., Petrović, M. D., Grigorieva, E. M. (2021). QR codes as a tool for receiving feedback about guests' satisfaction at destinations. Journal of Place Management and Development, 14(1), 19-31. https://doi.org/10.1108/JPMD-03-2020-0025

The Role of Artificial Intelligence in Enhancing the Efficiency and Reliability of Geographical Names Standardization: From Challenges to Opportunities

Mu'ammar Kamel Haddadin
Director General of the Royal Jordanian Geographic Centre
Chairman of the Arab Division of Experts in Geographical Names
Chairman of the National Committee on Geographical Names
Jordan
Email: rjgc@rjgc.gov.jo; contact@adegn.info; m.haddadin@rjgc.gov.jo

Introduction

In an age where technology is rapidly advancing and the need for spatial organization is growing, Artificial Intelligence (AI) emerges as a revolutionary tool for rethinking our approach to geographical names. These names are not merely labels on maps—they are vaults of collective memory, culture, history, and meaning. But do we truly need AI in this field? The answer is: yes.

How Can AI Contribute to the Standardization of Geographical Names?

- **Historical Text Analysis**
  Through Natural Language Processing (NLP), AI can extract names from historical documents and link them to local context and heritage —a crucial and often problematic issue in name standardization.
- **Unifying Multiple Names**
  AI can partially distinguish between linguistic variations of the same name across local dialects, unifying or accurately classifying them.
- **Site Recognition Using Satellite Imagery**
  Using computer vision algorithms, AI can identify unrecorded villages or neighborhoods and link them geographically to existing data.
- **Building Intelligent Databases**
  Smart models can connect a geographical name with its meaning, history, location, pronunciation, and changes over time.
- **Predicting Future Geographical Names**
  Based on linguistic and demographic trends, AI can propose potential future names.

Traditional Challenges in the Field of Geographical Names

- **Name Duplication**
  Many towns and villages share the same or very similar names (e.g., "Ramtha", "Ramtheh", "Hashimiyah", "Taybeh", "Husn"), which may differ only in diacritics—leading to confusion in documentation and analysis.
- **Historical Transformations**
  Place names often change due to political, religious, or social factors, making it difficult to connect current names with historical references.
- **Linguistic and Dialectical Diversity**
  Local dialects influence how names are pronounced and written.
- **Loss of Oral Heritage**
  Many names originated orally and were never documented, risking their disappearance over time.
- **Multiple Data Sources and Quality Variance**
  Conflicting or outdated data from local or national sources makes verification difficult.
- **Difficulty Analyzing Old Geographical Documents**
  Due to a lack of diacritics, handwritten formats, or image-based documents, advanced Optical Character Recognition (OCR) and high linguistic accuracy are needed.
- **Absence of Diacritics in Arabic**
  This causes visual repetition of words that differ in meaning and pronunciation.

The Role of the Royal Jordanian Geographic Center (RJGC)

The RJGC has played a pivotal role by:

- Providing foundational databases.
- Updating geographical names in coordination with official entities.
- Offering downloadable digital maps in formats like Shapefiles for various applications.
- Producing interactive maps of geographical names to aid in standardization efforts.

Tools That Can Be Used for Geographical Name Standardization

- Python + NLP (spaCy) – for extracting and linguistically analyzing names from documents.
- QGIS + GeoPandas – for spatial analysis and linking names to coordinates.
- Optical Character Recognition (OCR) tools.
- Data aggregation and organization software.

Challenges and Recommendations

1. Form multidisciplinary teams combining toponymy experts, data engineers, and linguists to ensure deep contextual analysis.
2. Create databases using unique identifiers (UIDs) to avoid duplication and track name changes over time.
3. Integrate AI tools institutionally, such as NLP for name classification, OCR for digitizing old documents, and GIS for verifying locations.
4. Develop open-source evaluation frameworks to objectively assess tool performance (e.g., extraction accuracy, processing time).
5. Promote transparency and open-source sharing of data and language models for collaborative development.
6. Digitize historical references to build a thematic, chronological, and linguistic digital library for geographic names.
7. Raise community awareness and encourage locals to help document names and proper pronunciations in regional dialects.

Conclusion

AI is not a luxury in the field of geographical names—it is a scientific, cultural, and security necessity. It acts as a bridge between heritage and the future, between meaning and the map.

Reference list

1- https://www.esri.com/ar-sa/geospatial-artificial-intelligence/overview.

2- https://geopandas.org/en/stable/.

3- https://spacy.io/models.

4- Wikipedia

The use of artificial intelligence in the standardization of geographical names

Wendy Shaw and Christopher Stephens
Secretariat, New Zealand Geographic Board Ngā Pou Taunaha o Aotearoa
New Zealand
Email address: wshaw@linz.govt.nz, cstephens@linz.govt.nz

Introduction

New Zealand's national report GEGN.2/2025/62[1] to the 2025 session of UNGEGN mentioned testing AI to analyze substantial numbers of submissions received during public consultation of geographical name proposals. This article goes into more detail about our intentions.

Why are we considering using AI?

We've decided to investigate using AI as part of our standardization processes for geographical names to:

- ensure consistent and objective analysis and accuracy,
- reduce timeframes and costs as current practice is not sustainable, and
- utilize new technology for improving processes.

If the results are of high quality and consistent when compared with the current manual analysis, it could mean an easy decision to transition to adopting AI for all future analysis for proposals with high numbers of submissions. However, provisos would be needed around informing people about how their information will be assessed, categorized and summarized. Preserving the security of the information received, and maintaining the integrity of the process, are priorities.

From Toitū Te Whenua Land Information New Zealand's Head of Architecture "I think there is real potential to use AI to assist the manual data analysis by providing a second opinion on the themes and insights. In the future, the potential benefit is to save time by having AI do the analysis and our experts review and check the results."

Processing submissions

New Zealand's national naming authority, Ngā Pou Taunaha o Aotearoa New Zealand Geographic Board[2], is required to publicly notify geographical name proposals and invite submissions with reasons for supporting or objecting to the proposals, and potentially to make alternative proposals. After the consultation timeframe of between 1-3 months, if no submissions - or only supporting submissions - are received, then the Board can make the final decision on the proposal. The Board may also make the final decision if it agrees with the objection(s). However, if the Board doesn't agree with any objection, then the final decision is made by the Minister for Land Information. The Board's report to the Minister must summarize all submissions to help the Minister make an informed decision.

How does the Board receive submissions?

Most submissions are made through an online form (Qualtrics) which auto-populates columns in an electronically generated spreadsheet. Other avenues to make submissions are via email[3] or by post. Submissions received by email and post are far less common.

The Board's Secretariat transcribes postal submissions and combines them with email submissions and Qualtrics submissions into a single spreadsheet to analyze all submissions.

Considerations and categories for objecting submissions

Because of its statutory requirements to uphold or reject, the Board is focusing on methods for more effectively/efficiently assessing objecting submissions. In October 2024 it agreed to apply common categories to objecting submissions based on a variety of considerations, such as community engagement numbers, submissions of substance that may also have uncommon reasons, high volumes of objections with the same reasons, and anything that might influence the Board's (or the Minister's) final decision.

Based on those considerations and responses from the past two decades, new high-level objection categories were developed, typically falling into these categories:

[1] https://unstats.un.org/unsd/ungegn/sessions/4th_session_2025/documents/GEGN.2_2025_62_CRP62_item4b.pdf

[2] https://www.linz.govt.nz/our-work/new-zealand-geographic-board/about-new-zealand-geographic-board

[3] NZGBsubmissions@linz.govt.nz

- Cost
- Wasting time
- Long-term use
- Confusion
- Lack of community support
- Māori views
- Anti-Māori sentiments
- Identity/connection/culture/history
- Pronunciation
- Language, macrons, spelling
- Consultation process inadequate
- Businesses, branded products
- No reasons provided
- Alternative name(s) proposed
- Other (including vitriolic)

How does the Board count submissions?

Our criteria for submissions are that:

- We accept only one submission per person, group, or organization. If submitters make multiple submissions, they will be combined into one submission.
- Individual submitters who sign a joint petition will be counted as individual submitters. If they have made a personal submission in addition to signing a joint petition, the content will be combined and counted as one submitter/submission.
- Where different people make different submissions from the same email or IP address, they are counted as separate submissions.
- *'threatening or offensive'* submissions are not considered.

The final decisions are not based on majority voting, but are influenced by community views, the reasons provided, legislative functions, and good naming practice.

What inputs would we provide to AI?

An initial consideration was testing the AI tool's capability to compile postal, email, and online submissions, accurately identify duplicate submissions, and merge multiple submissions from the same individuals.

However, completing these tasks requires comparison of personal information (names, phone numbers, addresses, IP addresses, emails). We will anonymize submissions to protect the privacy of submitters before any use of AI tools is permitted.

Therefore, the combined submission spreadsheets as prepared manually by the Board's Secretariat will be used (after anonymizing) as the input data source for the AI tool.

What results do we expect from AI?

The following results, which need to be comparable with a prior manual analysis, are expected from using the anonymized spreadsheets of all submissions:

1. The number of submissions for and against (and potentially those that are neutral).
2. Which objections made alternative name suggestions, and what they were.
3. Which objections gave no reasons or placing any text to meet the mandatory field filling requirement.
4. The number of submissions from locals, from outside the community, and from overseas.
5. An assessment against the high-level categories described above.
6. Identifying any submissions that warrant closer attention, i.e., with well-presented arguments and potentially additional reasons (in addition to the high-level objection categories listed above). Possibly based on the length of the submission or if it has attached document(s).
7. Any other things we would potentially like to know as outcomes.

We may experiment with producing different categories for submissions using AI summaries as we work through the results.

Test example: proposal to alter Petone to Pito One

A 2024 proposal to correct the spelling of the suburb of Petone to Pito One received 2,843 submissions (2,081 supporting and 762 objecting). The AI app 'Claude' developed by Anthropic [4] was used in a first-pass test against the submissions. For this first-pass test, 'Claude' was prompted to provide the results as described above, other than assessment against existing high-level categories. For comparison, the standard manual analysis for this proposal is in the *Submissions Report.*[5]

[4] https://www.anthropic.com/

[5] https://www.linz.govt.nz/sites/default/files/2024-12/NZGB_Pito-One_04a-Board-Submission-report_20241015.pdf

The exact prompt was:

> *Analyse the attached public submissions on the proposal to rename the suburb of Petone. and answer the following questions. · the numbers of submissions **for** and **against** (and potentially **neutral**) · which objections made **alternative name** suggestions and **what they are** · which objections gave **no reasons** or used non-sensical characters to meet the mandatory field filling requirement · the numbers of submissions from **locals**, from **outside** the community and from **overseas** · any submissions that **warrant closer attention**, ie with well-presented arguments and potentially additional reasons (in addition to what the Board decided and/or in addition to the objection categories). Possibly based on the length of the submission or if it has an attached document · potentially other things we would like to know as outcomes.*

Claude's key observations from the test anonymized-submissions spreadsheet were:

1. Strong Support: Nearly three-quarters of submissions supported the name change to Pito One.
2. Local Interest: The vast majority of submissions came from locals (71%), showing strong community engagement.
3. Alternative Approaches: Among objections, there's notable support for compromise solutions such as:
   a. Dual naming (keeping both names in some format)
   b. Partial renaming (applying Pito One to specific locations within the area)
4. Substantive Feedback: Nearly 10% of submissions contain detailed reasoning that may provide valuable insights beyond simple for/against positions.
5. Consistent Submission Method: Most submissions (96%) came through the Qualtrics online form, suggesting a coordinated outreach effort.

'Claude' was then requested to provide a summary of the *'main themes in Pito One renaming submissions'*, producing these results (noting that some submissions contain multiple categories, so percentages do not add up to 100%):

• Support for Cultural Recognition | Support for recognizing and restoring the original Māori name	2,081	73.0%
• Preservation of Established Name | Preference to retain Petone due to historical usage and recognition	772	27.0%
• Dual Naming Compromise | Support for using both names in some format (hyphenated, slash, etc.)	10+	0.4%
• Location-Specific Naming | Suggestions to rename specific areas (beach, park, pā site) rather than the entire suburb	11+	0.4%
• Spelling Variations | Alternative spelling suggestions (Pitōne, Pitoone, etc.)	6+	0.2%
• Economic/Practical Concerns | Concerns about costs of changing signage, addresses, business names, etc.	35-50	1.2-1.8%
• Historical Arguments | Detailed historical reasoning about settlement patterns and naming	40-60	1.4-2.1%
• No Substantive Reason | Submissions without meaningful justification	42	1.5%
• Local Identity Connection | Strong personal/community connection to current name	50-70	1.8-2.5%
• Te Tiriti/Partnership Principles | References to treaty principles and partnership obligations	30-45	1.1-1.6%

Next steps

These preliminary 'Claude' AI results will be refined further by experimenting with the input conditions and questions, including requesting results in terms of our existing thematic categories. The final results will be compared with the results of the manual analysis already completed for the Petone to Pito One proposal. We will also consider running the refined prompts against two other recent geographical name proposals: for altering the suburb name of Takanini to Takaanini (245 submissions) and altering the village name of National Park to Waimarino (2,089 submissions).

The threshold to accept the use of 'Claude' AI as part of our process to analyze submissions has yet to be determined, but the accuracy of the results of the test is likely to be over 90% in comparison to the manual summary. It should be noted that once refined, the results from AI may in fact be more accurate than manual processing, which inherently includes human error and inconsistencies. Therefore, the success criteria would be for the AI to be as good as or better than manual processing, achieving the results in less time. With all information, the Board will decide on the suitability of permanently adopting a robust and refined AI process for assisting with processing large volumes of submissions in the future. Such adoption would need to cover the legal, privacy, security and ethical considerations.

Summary

The New Zealand government has encouraged agencies to use AI and has issued guidelines on digital.govt.nz[6]. Within our own government agency, Toitū Te Whenua Land Information New Zealand provides educational resources to staff on using AI: AI use policy, using AI with Microsoft Copilot, adding AI features to existing applications, and an online course for using AI tools responsibly. It will be very important to declare to submitters how the Board will be using AI (if adopted), and to reassure them that it will complement and not compromise the past practice of manual assessment of their submissions. This can be justified over time, with more proposals having successfully applied the AI analysis. We expect the resulting accuracy rates to be beyond minimum thresholds or that manual interventions or audits become less necessary as confidence rises. More substantive submissions, especially those bringing new information to the Board, will still be identified for close attention. Numbers of submissions are used as a judgement of community support for the proposal. Board decision-making is led by good information.

[6] https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/artificial-intelligence/public-service-artificial-intelligence-framework

AI-Assisted Evidence-Based Toponymic Standardization: Digitizing Historical Geographical Names Dictionaries

Peder Gammeltoft
Expert of Norway
University of Bergen
Norway
Email: peder.gammeltoft@uib.no

Introduction

The recent 2025 session of UNGEGN emphasized the importance of evidence-based approaches to geographical name standardization, particularly in regions with complex colonial histories and multilingual heritage. Recent advances in artificial intelligence offer new methodologies for processing historical toponymic sources, enabling systematic analysis of naming patterns and supporting informed standardization decisions. This article presents a case study of AI-assisted analysis applied to J.W. McGuire's *Geographic Dictionary of the Virgin Islands of the United States* (1925), demonstrating how digital humanities approaches can enhance UNGEGN's mission of promoting standardized geographical nomenclature.

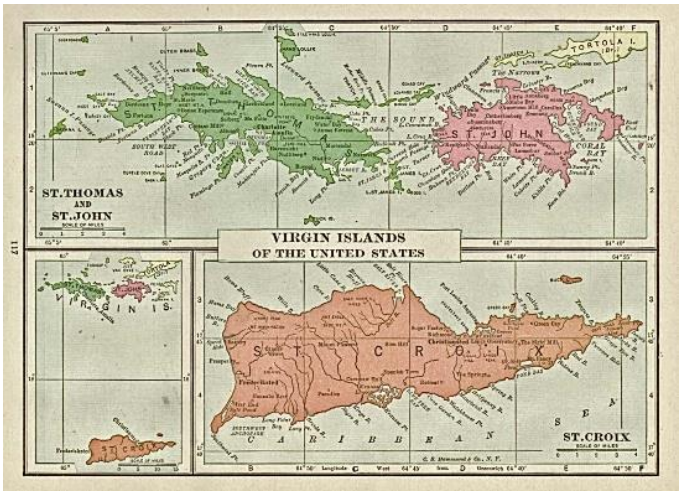


Figure 1: Historical map of the US Virgin Islands from Putnam's Handy Volume Atlas of the World (1920).

The Virgin Islands present a paradigmatic case for evidence-based geographical name standardization. The archipelago's political transitions, from early French and Dutch settlement through Danish colonial administration (1672-1917) to its current American territorial status, have created layers of toponymic heritage that call for careful documentation and analysis. McGuire's dictionary, compiled shortly after the U.S. purchase, represents a historical snapshot capturing the islands' multilingual nomenclature.

For this article, a subsection of 34 distinct geographical features across St. Thomas, St. John, and St. Croix (out of a total of 3,400 entries), recording not only contemporary names but also historical variants with source attributions. This methodological approach aligns with UNGEGN Resolution I/4 on the importance of historical documentation in geographical name standardization, providing the evidence base necessary for informed toponymic decisions.

Figure 2: Map section from the north coast of St. John showing the AI converted coordinates.

Methodology: AI-Enhanced Source Processing

The digitization process employed Claude Sonnet 4 to systematically extract, standardize, and analyze toponymic data from scanned and OCR-processed historical text. The Sonnet 4 model demonstrated its strength in handling multilingual content and complex scholarly apparatus. The analysis proceeded through five distinct phases. The first phase concentrated solely on error correction and source validation. A prompt was given to systematically correct OCR errors while preserving original scholarly citations and cross-references. This phase ensured a data integrity

essential for evidence-based analysis. Then followed the second phase of structural analysis and normalization, which involved converting narrative descriptions into structured data fields. These fields included: primary geographical names, feature classifications, precise locations, and comprehensive source attributions in one table. The individual source attributions were organized in one field in an array. Concentrating on the recorded name variants, phase three focused on the systematic extraction of all recorded name variants with language identification, source attribution, and historical context. This phase produced 82 documented name variants across Danish, French, Spanish, Dutch, and English linguistic traditions. Once performed, phase four was initiated, separating the historical and linguistic documentation into a separate table, based on the original documentation array field and the ID field. The original table was retained, but with the documentation field separated out, thus creating a relational database enabling cross-referencing between geographical features and their historical variants, supporting comprehensive toponymic analysis.

During the work with this, it became clear that some of the entries had coordinates in the description expressed in degrees, minutes, and seconds. To address this, a fifth-phase prompt was given to convert historical coordinate notations into the WGS84 standard, enabling integration with contemporary GIS systems. This proved surprisingly successful, as the coordinate conversion yielded relatively correct locations, albeit with a systematic error of 250-300 meters in a northern direction.

Results: Evidence-Based Toponymic Analysis

The systematic analysis revealed several patterns relevant to UNGEGN standardization principles:

- **Administrative Linguistic Stratification**: The data demonstrates clear linguistic layers corresponding to different administrative periods. French estate names from seventeenth-century settlement (Andrieu, Andrin) reflect early colonial toponymy. Danish official terminology dominates eighteenth- and nineteenth-century records (Amalienborg, Annashaab), while English adaptations appear concurrently with Danish administration and following the American acquisition.

- **Source Attribution Patterns**: McGuire's systematic citation of sources – including Høst's 1776 historical account, Lapointe's 1671 maps, and various Danish official publications – provides the evidence base essential for informed standardization decisions. The AI successfully preserved these attributions, enabling verification of toponymic claims.

- **Orthographic Issues**: Multiple spelling variants for identical features (Annaly/Annally, Annaberg/Anneberg) reflect inconsistent romanization practices across sources. The systematic documentation of these variants supports evidence-based decisions about preferred orthographic forms.

Implications for UNGEGN Practice

This methodology offers several advantages for UNGEGN's evidence-based standardization mission, such as systematic source processing with multilingual stratification: AI capabilities enable comprehensive analysis of historical gazetteers, maps, and administrative records that might otherwise remain underutilized due to processing complexity. Advanced language models can process historical sources across multiple languages simultaneously, essential for regions with complex linguistic or colonial histories. The structured *human-in-the-loop* approach ensures comprehensive source attribution and cross-referencing, meeting UNGEGN requirements for evidence-based toponymic work.

Based on this case study, some avenues emerge for future UNGEGN applications, such as pilot programs to consider AI-assisted analysis for processing historical toponymic sources in regions with complex linguistic heritage, particularly former colonial territories with multilingual nomenclatures. More crucially, though, is the development of standardized protocols for AI-assisted toponymic analysis would ensure consistency across different national and regional applications, alongside capacity-building programs to enable national toponymic authorities to apply these methodologies to their own historical sources and standardization challenges.

Conclusion

The AI-assisted analysis of McGuire's 1925 Virgin Islands gazetteer demonstrates the potential for enhancing evidence-based geographical name standardization through digital humanities approaches. By systematically processing complex historical sources while preserving scholarly rigor, these methodologies can support UNGEGN's mission of promoting standardized geographical nomenclature based on sound historical and cultural evidence.

As UNGEGN continues addressing toponymic challenges in multilingual and culturally diverse regions, AI-assisted analysis, combined with human supervision, offers valuable tools for processing the historical sources essential to informed standardization decisions. This article shows the potential of broader applications across regions where colonial histories, linguistic diversity, and incomplete documentation create challenges for evidence-based toponymic standardization. A digital humanities human-supervised AI approaches is beneficial in supporting UNGEGN's global mission of geographical names standardization while respecting the cultural and historical significance embedded in geographical nomenclature.

Resources

Claude Sonnet 4:
https://www.anthropic.com/claude/sonnet

J.W. McGuire (1925) *Geographic Dictionary of the Virgin Islands of the United States*. Washington. https://www.columbia.edu/cu/libraries/inside/projects/ebooks/prd/testing/Islands/html/pages/FID3.html

Main table derived from dictionary:
https://claude.ai/public/artifacts/4af9db85-f98a-4edc-b716-bcc4cdea7749

Historical sources table:
https://claude.ai/public/artifacts/d8c2cc89-5e7b-4c7d-8d23-88d078c674ae

AI-Driven Standardisation of Geographical Names: Innovations, Challenges, and Strategic Alignment with UNGEGN Goals

Nuha Hamed Al Subhi
Sultan Qaboos Naval Academy
PhD Student, Information Studies Department, Sultan Qaboos University, Muscat, Oman
Email: n-alsubhi@hotmail.com, s95666@student.squ.edu.om

Introduction

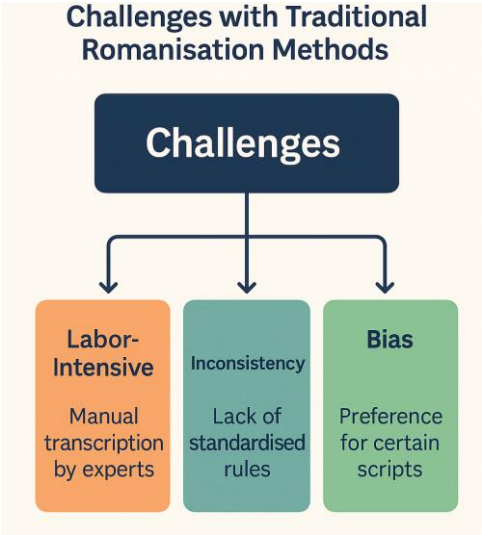
The standardization of geographical names is essential for effective global communication, governance, and technological integration while also preserving cultural identity. Accurate and consistent place names can reduce geopolitical tensions and support multilingual inclusivity, primarily by using endonyms over exonyms to maintain cultural authenticity (Kladnik, 2007; Purba, 2023; Mikesy, 2016). The United Nations Group of Experts on Geographical Names (UNGEGN), active since 1959, promotes best practices and global coordination in this field (Perko et al., 2017; Wolnicz-Pawłowska, 2017).

Incorporating AI technologies aligns with UNGEGN's Strategic Plan (2021–2029), particularly in fostering innovation and interoperability. Machine learning and natural language processing enhance decision-making, data management, and cross-sectoral integration (Gupta et al., 2024; Aldoseri et al., 2024; Nikolinakos, 2023). However, this progress must be guided by ethical frameworks to mitigate risks such as data bias, misrepresentation, and inadequate regulation (Yadav et al., 2024; Maroju, 2024). This article explores the role of AI in modernizing geographical name standardization, evaluates challenges and ethical considerations, and offers strategic recommendations aligned with UNGEGN goals.

AI Applications in Geographical Name Standardization

1. **Multilingual Recognition and Romanization**
   AI, through NLP tools like UROMAN, facilitates the romanization of non-Latin scripts, aiding linguistic inclusivity (Purkayastha et al., 2023; Xhelili et al., 2024). Performance increases when phonemic cues and informal romanized text are integrated (Nguyen et al., 2024; Husain et al., 2024). Despite progress, structural differences between languages present challenges. AI models trained in morphologically rich languages often underperform due to the scarcity of datasets, necessitating targeted data augmentation.

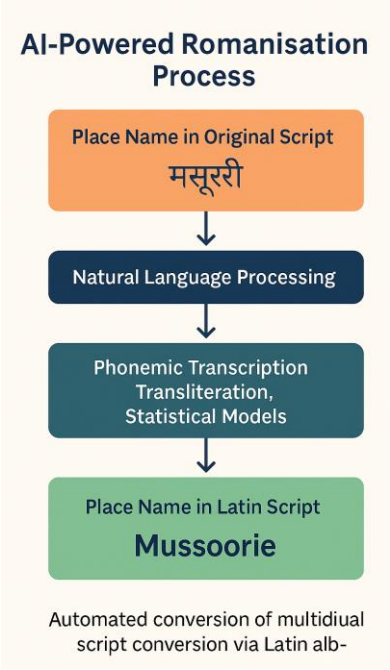
2. **Data Quality in Historical Gazetteers**
   AI improves gazetteer accuracy by correcting OCR errors, resolving duplicates, and clarifying toponyms. Techniques like character-level features and clustering algorithms have demonstrated high effectiveness (Dannélls & Virk, 2021). Still, model accuracy depends on data diversity (DeLozier et al., 2015; Martins, 2011; Pellegrino et al., 2021). Hybrid human-AI annotation methods can enhance training sets, especially in historical contexts where toponym forms shift over time.



3. **Cultural Sensitivity and Indigenous Knowledge**
   Incorporating Indigenous Knowledge Systems (IKS) promotes data sovereignty and cultural respect (Gao, 2022). Tools using culture-matching algorithms align outputs with community values (Tao et al., 2023; Liu, 2024). Real-world examples include AI tools for language preservation and Indigenous financial inclusion (Perera et al., 2024; Jangra et al., 2024). Community involvement and respect for local data ownership are crucial (Ray, 2024; Wang et al., 2024). Case studies from New Zealand and Canada illustrate how co-designed AI initiatives can enhance linguistic revival while aligning with local governance principles.

4. **Bias Mitigation and Ethical Design**
Bias from underrepresentation in training data can lead to systemic inequalities. Initiatives like the Child Growth Monitor show how context-rich datasets ground AI in reality (Güven et al., 2025). Frameworks for fairness and transparency enhance accountability (Mishra et al., 2024; Helm et al., 2023). Pre-deployment auditing and explainability tools, such as LIME and SHAP, can further mitigate opaque decision-making processes.

5. **Infrastructure, Governance, and Education**
Scalable AI requires investment in digital infrastructure, legal frameworks, and human capital. Decentralized platforms and blockchain enhance data security and interoperability (Domalis et al., 2021; Kulothungan, 2025). Legal harmonization and AI training, such as in healthcare, further support responsible innovation (Nurudeen et al., 2024; Ramírez & Islam, 2024). Government-academia partnerships can foster AI fluency among local policymakers and facilitate the development of regional standards.



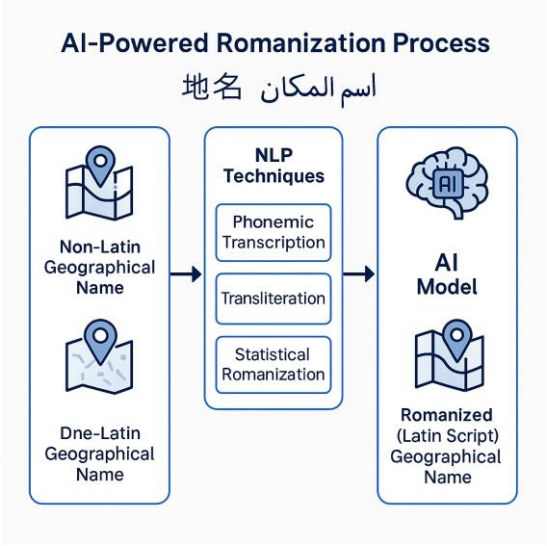
Automated conversion of multidiual script conversion via Latin alb-

AI and UNGEGN Strategic Alignment

1. **Technical Governance and Transparency**
AI enhances technical governance by automating standards enforcement and improving data traceability. AI model registries promote transparency (McKernon et al., 2024), and proposals like the IAIO support international regulatory coherence (Trager et al., 2023). Ethical AI development requires participation from Indigenous and local communities (Constantinides et al., 2023; Laux et al., 2024). Tools like datasheets for datasets and model cards can support UNGEGN's efforts to create transparent naming standards.

2. **Cross-Border Collaboration**
Projects like Harmo-Data, Clinnova, and the Humboldt Project demonstrate AI's potential for cross-border data harmonization (Barborič et al., 2019; Alekseenko et al., 2024). These efforts enhance semantic and geometric alignment, yet legal and privacy concerns persist (Gajda, 2023; Alabi et al., 2023). International working groups and regional data trusts could help address governance asymmetries and encourage mutual data sharing.

3. **Integration with Geospatial Systems**
Interfaces like OGDI facilitate integration between AI, gazetteers, and GIS platforms (Clément et al., 1999). Indonesia's AI-enhanced SDI is an example of how this integration improves decision-making (Nugroho & Supangkat, 2021). However, real-time processing demands and system inconsistencies remain obstacles. Blockchain-AI hybrids offer solutions but increase complexity (Vysotskyi & Vysotskyi, 2024). Investing in API standardization and edge computing may offer scalable pathways to harmonization.



Challenges and Ethical Considerations

Bias in AI systems, particularly concerning underrepresented languages, often stems from non-diverse datasets (Bella et al., 2024). Community engagement and fairness-aware machine learning are necessary for equitable outcomes (Mishra et al., 2024). Ethical guidelines from bodies like the IEEE and the EU help enforce standards of transparency and accountability (Oyeniran et al., 2022).

Global cooperation is essential for harmonizing regulations and protecting data sovereignty, especially in developing

nations facing infrastructural gaps (Folorunso et al., 2024). Investments in AI education and public-private partnerships can build local capacity and foster innovation. Policymakers must also monitor the social impact of AI to prevent deepening inequalities. Additionally, environmental sustainability must be considered; large AI models can contribute to carbon emissions if deployed inefficiently.

Recommendations and Conclusion

AI's role in toponymy must be governed by ethical principles that prioritize innovation without compromising cultural integrity. Embedding Indigenous knowledge, guided by frameworks like UNESCO's AI ethics recommendations, ensures context-aware development (Tiribelli et al., 2023; Morandín-Ahuerma, 2023). Collaboration between AI developers, geographers, and cultural experts fosters tools that preserve diversity (Pasupuleti, 2024; Yu, 2024). Ethics education grounded in traditional values can cultivate culturally competent AI professionals. UNGEGN is well-positioned to lead in promoting the adoption of ethical AI by encouraging inclusive governance and transparent monitoring systems (Tiribelli et al., 2023; Pasupuleti, 2024). Ultimately, a flexible and inclusive ethical framework is vital to reflect the world's cultural diversity and ensure AI in toponymy serves all communities relatively. As AI systems become increasingly embedded in global geospatial governance, ensuring they align with both technical standards and human values will be crucial for achieving a lasting impact.

References

Alabi, A. A., Mustapha, S. D., Ike, C. C., & Ige, A. B. (2023). Resolving cross-border privacy and security misalignments with a unified harmonisation framework for the U.S. and Canada. *Open Access Research Journal of Multidisciplinary Studies*. https://doi.org/10.53022/oarjms.2023.6.1.0036

Aldoseri, A., Al-Khalifa, K. N., & Hamouda, A. M. (2024). AI-Powered Innovation in Digital Transformation: Key Pillars and Industry Impact. Sustainability. https://doi.org/10.3390/su16051790

Alekseenko, J., Stieltjes, B., Bach, M., Boerries, M., Opitz, O., Karargyris, A., & Padoy, N. (2024). Clinnova Federated Learning Proof *of Concept: Key Takeaways from a Cross-border Collaboration*. https://doi.org/10.48550/arxiv.2410.02443

Barborič, B., Cefalo, R., Chiarandini, A., Zorzi, S. D., Tomaž, P., Puhar, M., Zanardo, M., Sluga, T., Tommasi, A., & Trivelloni, U. (2019). *Harmo-Data Project - Cross Border Spatial Data Harmonisation Using Inspire Model*. https://doi.org/10.2902/1725-0463.2019.14.art2

Bella, G., Helm, P., Koch, G., & Giunchiglia, F. (2024). *Tackling Language Modelling Bias in Support of Linguistic Diversity*. https://doi.org/10.1145/3630106.3658925

Clément, G., Larouche, C., Morin, P., & Gouin, D. (1999). *Interoperating Geographic Information Systems Using the Open Geospatial Datastore Interface (OGDI)*. https://doi.org/10.1007/978-1-4615-5189-8_23

Constantinides, M., Bogucka, E. P., Quercia, D., Kallio, S., & Tahaei, M. (2023). A Method for Generating Dynamic Responsible AI Guidelines for Collaborative Action. *arXiv.Org*. https://doi.org/10.48550/arxiv.2307.15158

Dannélls, D., & Virk, S. M. (2021). A Supervised Machine Learning Approach for Post-OCR Error Detection for Historical Text. *Language and Technology Conference*. https://doi.org/10.3384/ECP184170

DeLozier, G., Baldridge, J., & London, L. (2015). Gazetteer-independent toponym resolution using geographic word profiles. *National Conference on Artificial Intelligence*.https://doi.org/10.1609/aaai.v29i1.9531

Domalis, G., Karacapilidis, N., Tsakalidis, D., & Giannaros, A. (2021). *A Trustable and Interoperable Decentralised Solution for Citizen-Centric and Cross-Border eGovernance: A Conceptual Approach*. https://doi.org/10.1007/978-3-030-84789-0_19

Folorunso, A., Olanipekun, K., Adewumi, T., & Samuel, B. J. (2024). A policy framework on AI usage in developing countries and its impact. *Global Journal of Engineering and Technology Advances*. https://doi.org/10.30574/gjeta.2024.21.1.0192

Gajda, A. (2023). Wykorzystywanie systemów sztucznej inteligencji (SI) na zewnętrznych granicach Unii Europejskiej. Możliwości, zagrożenia i wyzwania. *Krytyka Prawa*. https://doi.org/10.7206/kp.2080-1084.574

Gao, G. (2022). Bias detection by using name disparity tables across protected groups. *Journal of Responsible Technology*. https://doi.org/10.1016/j.jrt.2021.100020

Gupta, R., Biswas, D. K., Ronty, A. P., Kapali, S., & Khan, M. S. (2024). Empowering Digital Transformation: The Strategic Role of Artificial Intelligence in Enterprise Innovation. *European Journal of Theoretical and Applied Sciences*. https://doi.org/10.59324/ejtas.2024.2(6).16

Gupta, S. (2024). Building Trustworthy AI: Proactive Guardrails and Reactive Moderation for Scalable Governance. *International Journal of Innovative Research in Science, Engineering and Technology*. https://doi.org/10.15680/ijirset.2024.1312181

Güven, Ç., Alishahi, A., Brighton, H., Nápoles, G., Olier, J. S., Šafář, M., Postma, E., Shterionov, D., Sisto, M. D., & Vanmassenhove, E. (2025). *AI in Support of Diversity and Inclusion*. https://doi.org/10.48550/arxiv.2501.09534

Helm, P. L., Bella, G., Koch, G., & Giunchiglia, F. (2023). Diversity and Language Technology: How Techno-Linguistic Bias Can Cause Epistemic Injustice. *arXiv.Org*. https://doi.org/10.48550/arxiv.2307.13714

Husain, J. A., Dabre, R. A., Kumar, A., Puduppully, R., & Kunchukuttan, A. (2024). RomanSetu: Efficiently unlocking multilingual capabilities of Large Language Models via Romanisation. *arXiv.Org*. https://doi.org/10.48550/arxiv.2401.14280

Jangra, G., Jangra, M., Irfan, M., & Verma, C. (2024). The Intersection of Indigenous Knowledge and Artificial Intelligence in Financial Technology. *Advances in Finance, Accounting, and Economics Book Series*. https://doi.org/10.4018/979-8-3693-7924-0.ch005

Kladnik, D. (2007). *Pogledi na podomačevanje tujih zemljepisnih imen*. https://doi.org/10.3986/9789612545253

Kulothungan, V. (2025). *A Blockchain-Enabled Approach to Cross-Border Compliance and Trust*. https://doi.org/10.48550/arxiv.2501.09182

Laux, J., Wachter, S., & Mittelstadt, B. (2024). *Three pathways for standardisation and ethical disclosure by default under the European Union Artificial Intelligence Act*. https://doi.org/10.1016/j.clsr.2024.105957

Liu, Z. Y. (2024). Cultural Bias in Large Language Models: A Comprehensive Analysis and Mitigation Strategies. *Journal of Transcultural Communication*. https://doi.org/10.1515/jtc-2023-0019

Maroju, P. K. (2024). *Advancing Synergy of Computing and Artificial Intelligence with Innovations, Challenges, and Future Prospects*. https://doi.org/10.69888/ftsin.2024.000153

Martins, B. (2011). *A supervised machine learning approach for duplicate detection over gazetteer records*. https://doi.org/10.1007/978-3-642-20630-6_3

McKernon, E., Glasser, G., Cheng, D., & Hadfield, G. K. (2024). *AI Model Registries: A Foundational Tool for AI Governance*. https://doi.org/10.48550/arxiv.2410.09645

Mikesy, G. (2016). Földrajzi nevek a közigazgatásban. *Névtani Értesítő*. https://doi.org/10.29178/nevtert.2016.8

Mishra, I., Kashyap, V., Yadav, N. K., & Pahwa, R. (2024). Harmonising Intelligence: A Holistic Approach to Bias Mitigation in Artificial Intelligence (AI). *Deleted Journal*. https://doi.org/10.47392/irjaeh.2024.0270

Morandín-Ahuerma, F. (2023). *Ten UNESCO Recommendations on the Ethics of Artificial Intelligence*. https://doi.org/10.31219/osf.io/csyux

Nguyen, H., Mahajan, K., Yadav, V., Yu, P. S., Hashemi, M., & Maheshwary, R. (2024). Prompting with Phonemes: Enhancing LLM Multilinguality for non-Latin Script Languages. *arXiv.Org*. https://doi.org/10.48550/arxiv.2411.02398

Nikolinakos, N. T. (2023). Boosting Excellence in AI: The 2021 Coordinated Plan with Member States. In *Law, governance and technology series* (pp. 309–325). https://doi.org/10.1007/978-3-031-27953-9_7

Nugroho, Y. S., & Supangkat, S. H. (2021). *Spatial Data Infrastructure Integrated with Geospatial Artificial Intelligence: A Systematic Literature Review*. https://doi.org/10.1109/ICISS53185.2021.9533208

Nurudeen, M. O., Latilo, A., Imosemi, H. O., & Imosemi, Q. A. (2024). *Comparative legal frameworks for regulating artificial intelligence: A model for harmonising AI laws in Latin America and Africa*. https://doi.org/10.58175/gjrms.2024.2.1.0038

Oyeniran, O. C., Adewusi, A. O., Adeleke, A. G., Akwawa, L. A., & Azubuko, C. F. (2022). Ethical AI: Addressing bias in machine learning models and software applications. *Computer Science & IT Research Journal*. https://doi.org/10.51594/csitrj.v3i3.1559

Pasupuleti, M. K. (2024). *Ethical AI Governance: A Global Blueprint*. https://doi.org/10.62311/nesx/97832

Pellegrino, M. A., Postiglione, L., & Scarano, V. (2021). *Detecting Data Accuracy Issues in Textual Geographical Data by a Clustering-based Approach*. https://doi.org/10.1145/3430984.3431031

Perera, M., Vidanaarachchi, R., Chandrashekeran, S., Kennedy, M., Kennedy, B., & Halgamuge, S. (2024). *Exploring the Intersection of Indigenous Knowledge and AI: A Systematic Review and Future Directions*. https://doi.org/10.31235/osf.io/6hrqj

Perko, D., Jordan, P., & Komac, B. (2017). Exonyms and other geographical names. *Acta Geographica Slovenica*. https://doi.org/10.3986/AGS.4891

Purba, N. P. (2023). The Nexus Between Geographical Naming, Place, and the Politics of Power. *Key Challenges in Geography*. https://doi.org/10.1007/978-3-031-21510-0_28

Purkayastha, S., Ruder, S., Pfeiffer, J., Gurevych, I., & Vulić, I. (2023). Romanisation-based Large-scale Adaptation of Multilingual Language Models. *arXiv.Org*. https://doi.org/10.48550/arXiv.2304.08865

Ramírez, J. G. C., & Islam, Md. mafiqul. (2024). *Navigating the Terrain: Scaling Challenges and Opportunities in AI/ML Infrastructure*. https://doi.org/10.60087/jaigs.v2i1.p228

Ray, S. (2024). Exploring the Role of Artificial Intelligence in Language Documentation and Endangered Language Preservation. *Journal of Propulsion Technology*. https://doi.org/10.52783/tjjpt.v45.i02.6122

Tao, Y., Viberg, O., Baker, R. S., & Kizilcec, R. F. (2023). Auditing and Mitigating Cultural Bias in Large Language Models (LLMs). *arXiv.Org*. https://doi.org/10.48550/arxiv.2311.14096

Tiribelli, S., Paolanti, M., Frontoni, E., & Giovanola, B. (2023). Design of an Ethical Framework for Artificial Intelligence in Cultural Heritage. *Ethics*. https://doi.org/10.1109/ETHICS57328.2023.10155020

Trager, R. F., Harack, B., Reuel, A., Carnegie, A., Heim, L., Ho, L., Kreps, S. E., Lall, R., Larter, O., h'Eigeartaigh, S. 'O, Staffell, S., & Villalobos, J. J. (2023). International Governance of Civilian AI: A Jurisdictional Certification Approach. *arXiv.Org*. https://doi.org/10.48550/arxiv.2308.15514

Vysotskyi, A. O., & Vysotskyi, O. (2024). Decentralising spatial data: the convergence of Geographic Information Systems and Web 3.0 technologies. *Journal of Geology, Geography and Geoecology*. https://doi.org/10.15421/112377

Wang, G., Bennamoun, H., Kwok, W. H., Marriott, R., Walker, R., & Kotz, J. (2024). Codesigning a Clinical Prediction Model for Aboriginal Perinatal Mental Health Using Glassbox AI and Aboriginal Wisdom and Lived Experience. *Studies in Health Technology and Informatics*. https://doi.org/10.3233/shti240925

Wolnicz-Pawłowska, E. (2017). The fiftieth anniversary of the first United Nations Conference on the Standardisation of Geographical Names. *Polish Cartographical Review*. https://doi.org/10.1515/PCR-2017-0010

Xhelili, O., Liu, Y., & Schütze, H. (2024). *Breaking the Script Barrier in Multilingual Pre-Trained Language Models with Transliteration-Based Post-Training Alignment*. https://doi.org/10.48550/arxiv.2406.19759

Yadav, M., Chandel, A., & Sabnis, R. (2024). Accelerating the Development of New Products and Services With AI-Enabled Innovation Management. *Advances in Marketing, Customer Relationship Management, and e-Services Book Series*. https://doi.org/10.4018/979-8-3373-0219-5.ch006

Yu, C. (2024). *The Garden of Forking Paths: Options for Non-Western Societies in the Age of AI*. https://doi.org/10.31219/osf.io/bk

The Use of Artificial Intelligence in the Standardization of Undersea Feature Names

Fatma Mubarak Al Dairi
Oman National Hydrographic Office
Oman
Email: FATMA.ALDAIRI@mod.gov.om

The Concept and the Role of Artificial Intelligence in Achieving the United Nations Sustainable Development Goals

Artificial Intelligence (AI) represents a qualitative shift in software and smart applications that rely on computers and modern technology to study and analyze data and algorithms in all economic, social, scientific, and other fields. There is no doubt that the integration of AI and sustainable development goals significantly enhances the achievement of these goals.

AI is distinguished by its ability to modernize the way data is processed and analyzed, providing countries and institutions with insights into the processes they create. Countries around the world have come together to enhance efforts to achieve the United Nations' Sustainable Development Goals (SDGs). There is no doubt that AI technologies play a fundamental role in exploring the marine environment and changing underwater ecosystems. AI contributes to the analysis of underwater data and modeling of data, allowing freedom to analyze these submerged marine environments and uncover the unknown in the terrain of these marine environments. This perspective explores the potential of AI to address environmental, social, and ethical challenges related to the SDGs.

Ocean Digitization and the Seabed 2030 Project

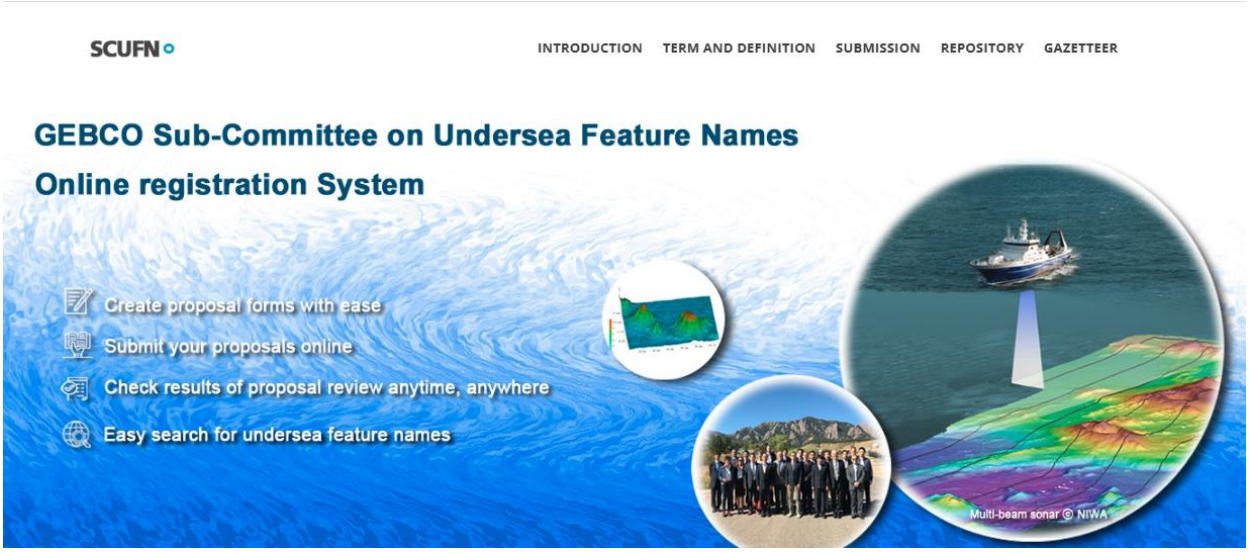
The Seabed 2030 Project, a joint initiative between The Nippon Foundation and the General Bathymetric Chart of the Oceans (GEBCO), represents a monumental effort to create the first definitive map of the world's ocean floor by the year 2030. This ambitious project was launched during the United Nations Ocean Conference in 2017 and aligns closely with Sustainable Development Goal 14 (SDG 14), which focuses on the conservation and sustainable use of marine resources. The project aims to compile all available bathymetric data into a unified, high-resolution map that will be freely accessible to researchers, policymakers, and the public. Given that only about 25% of the ocean floor has been mapped in detail, Seabed 2030 seeks to address this critical knowledge gap, enabling advancements in marine science, climate modeling, and maritime safety.

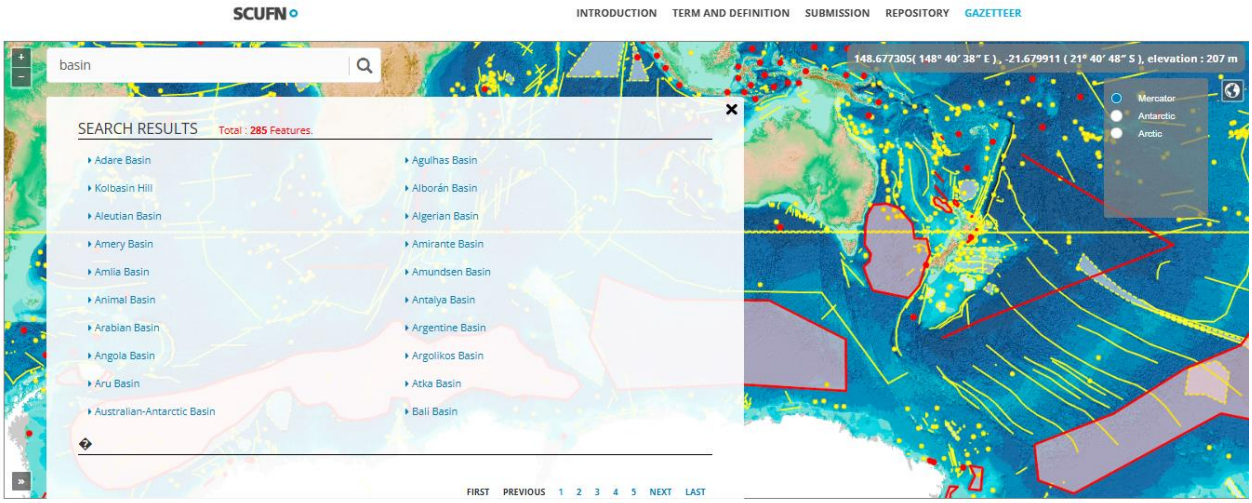
A key driver of the Seabed 2030 initiative is the integration of cutting-edge technologies, including Geospatial Artificial Intelligence (GeoAI), autonomous underwater vehicles (AUVs), and satellite-derived bathymetry. The International Hydrographic Organization (IHO) has emphasized the transformative potential of GeoAI in processing vast amounts of hydrographic data, improving accuracy, and automating the interpretation of seafloor features. Machine learning algorithms can analyze sonar data, predict undersea topography, and even identify previously unknown geological formations. Additionally, autonomous systems such as uncrewed surface vessels (USVs) and deep-sea drones are being deployed to explore remote and hazardous regions, significantly accelerating data collection.

Beyond technological innovation, the success of Seabed 2030 relies on global collaboration. Governments, research institutions, and private entities are contributing existing datasets to share depth measurements. Open-access platforms like GEBCO's global grid ensure transparency and foster interdisciplinary research. Once completed, the map will revolutionize our understanding of ocean ecosystems, improve tsunami and storm surge predictions, and support sustainable blue economy initiatives from offshore energy to deep-sea mining.

Undersea Feature Names: The Korean Initiative for a Digital SCUFN Portal

The General Bathymetric Chart of the Oceans (GEBCO) Sub-Committee on Undersea Feature Names (SCUFN) plays a crucial role as the international authority responsible for standardizing the names of seabed formations, including trenches, ridges, seamounts, and abyssal plains. As a joint body of the International Hydrographic Organization (IHO) and UNESCO's Intergovernmental Oceanographic Commission (IOC), SCUFN follows strict guidelines outlined in the IHO Publication B-6, which defines the criteria and procedures for proposing and approving undersea feature names. These standardized names are essential for scientific research, maritime navigation, and global ocean governance.





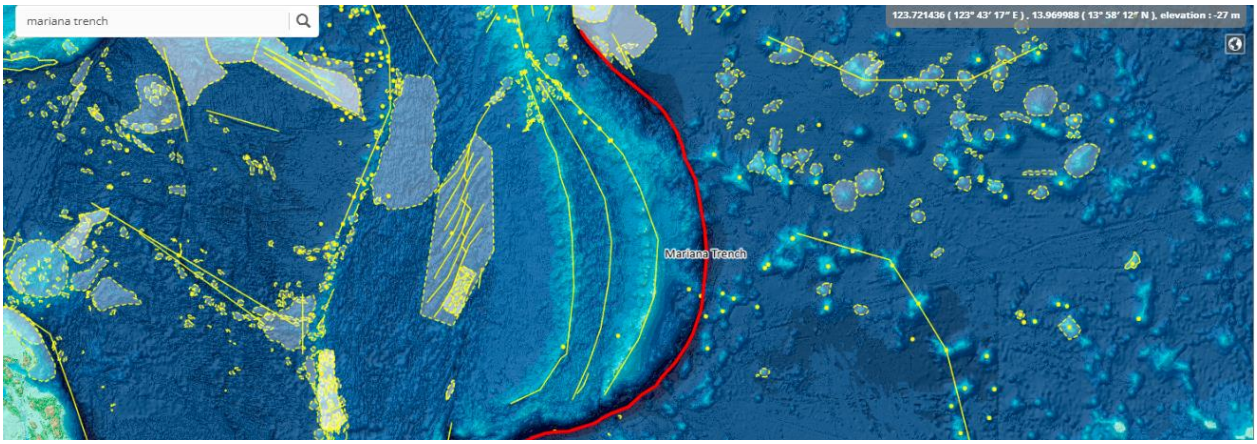
Recognizing the need for a more efficient and transparent naming process, the Korea Hydrographic and Oceanographic Agency (KHOA) has taken a leading role in modernizing SCUFN's operations by developing the SCUFN Operation Web Services (OWS). This electronic portal streamlines the submission and review of proposed undersea feature names, replacing traditional paper-based methods with a digital platform. The system enhances accessibility for researchers, hydrographic offices, and maritime organizations worldwide, ensuring faster processing and greater consistency in naming decisions.

The SCUFN OWS represents a significant step forward in oceanographic data management, aligning with broader initiatives such as the Seabed 2030 Project to enhance the global understanding of the seafloor. By integrating digital tools into the naming process, KHOA's initiative supports international collaboration, reduces administrative delays, and promotes the adoption of standardized undersea feature names in scientific literature and nautical charts. This effort underscores Korea's commitment to advancing marine geospatial information systems and contributing to global oceanographic research.

Capabilities of GeoAI for Hydrographic Data

AI plays a crucial role in standardizing workflows for undersea feature naming. The IHO has outlined emerging directions for GeoAI in hydrography, including integration with autonomous survey platforms for real-time analysis, the development of specialized foundation models for marine geomorphology, and enhanced visualization tools, powered by AI-generated metadata, further improve accessibility and interpretation of undersea feature data.

Supporting this effort is the GEBCO Gazetteer of Undersea Feature Names, the authoritative global database containing over 5,000 named features. Each entry includes detailed metadata such as geographic coordinates, feature type classification, physical dimensions, discovery information,



and name origins. The gazetteer is accessible through interactive 2D and 3D map interfaces, as well as downloadable formats like spreadsheets and shapefiles, ensuring researchers and policymakers can easily retrieve and utilize this critical information. As AI continues to evolve, its integration into hydrographic practices promises to enhance the accuracy, efficiency, and inclusivity of undersea feature naming ultimately contributing to a more comprehensive understanding of the world's oceans.

Advances in AI and machine learning (ML) are opening new possibilities for improving the efficiency and accuracy of undersea feature naming. One key application is automated feature detection, where AI can analyze high-resolution bathymetric data from sonar and LiDAR systems to identify uncharted seamounts, trenches, and ridges. These systems can classify newly discovered features according to international standards advanced by organizations such as the SCUFN. By flagging scientifically or navigationally significant formations, AI can help hydrographic offices prioritize which features require official naming.

Another promising area is cross-lingual and cultural processing, where AI can assist in translating existing names between languages while preserving their historical and etymological context. Additionally, AI can ensure compliance with naming guidelines. Quality control is another critical application, where AI algorithms can automatically verify that naming proposals meet SCUFN requirements. This includes verifying metadata completeness, validating geographic coordinates, and confirming that feature classifications (e.g., "seamount" vs. "knoll") align with their corresponding bathymetric profiles. However, challenges remain, including limited access to some bathymetric datasets.

Data Source: Korea Hydrographic and Oceanographic Agency (KHOA), Republic of Korea

References

1. IHO. (2025, 7 7). GEBCO – Seabed 2030 – Objectives – Grid Resolutions. Retrieved from IHO: https://iho.int/uploads/user/InterRegional%20Coordination/GEBCO/SCUFN/MISC/GEBCO_Seabed2023_Grid_Resolutions.pdf
2. International Hydrographic Organization (IHO). (2021). Standardization of Undersea Feature Names (IHO Publication B-6). https://iho.int/en/publications
3. International Hydrographic Organization (IHO). (2022). Cookbook for generic terms of undersea feature names Retrieved from IHO: (https://iho.int/uploads/user/Inter-Regional%20Coordination/GEBCO/SCUFN/MISC/Cookbook_for_Generic_Feature_terms_Oct_2022_v1.2.pdf
4. Lainjo, B. (2024). The Role of Artificial Intelligence in Achieving the United United Nations Sustainable Development Goals. Journal of Sustainable Development, 17(5), 1-30.
5. SCUFN. (2025, 7 7). GEBCO SUB-COMMITTEE ON UNDERSEA FEATURE NAMES. Retrieved from SCUFN: https://scufn.ops-webservices.kr
6. Verma, S., & Sharma, M. Role of Digitalization in a Sustainable Blue Economy. International Journal of Teaching, Learning and Education, 4(3), 618013.
7. Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Nerini, F. F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 11(1), 233. DOI:10.1038/s41467-019-14108-y

Designing AI Generative Agents for Arabic Geographical Names

Fatmah Baothman
Associate Professor, CEO of FI-Holding, Chairwoman of GENIE AI
Saudi Arabia
Email: Fatmah@ai4ia.tech
and
Talal AlShafaey
Assistant to the President for Sector Survey and Geospatial Activities
Saudi Arabia
Email: *t.alshafaey@geosa.gov.sa*

1. Introduction: Linguistic Complexity and Cultural Sensitivity

The Arabic language presents a highly complex linguistic landscape [1], leading to inconsistent transliteration. Arabic encompasses a rich morphological system, has pharyngealized consonants, geminate (doubled) consonants, vowels with three variations (i.e., short, long, and nunations (doubled)), sukuun (neither consonant nor vowel) [2], multiple alif scripts, three diacritics that change the meaning, six syllable structures, phonological processes that operate bi-directionally. In addition, dialectal variation may result in different meanings. Furthermore, 36% of consonant pairs occur in both orders, where "the glottal stop has the highest rate of occurrence, while the glide 'j' has the lowest. Each consonant, such as 'n,' can pair with up to 27 others, or 28 if geminates are included" [**3**]. This demonstrates the urgency for standardizing Arabic toponyms by establishing a direct mapping with the International Phonetic Alphabet (IPA). The author proposes an AI-based knowledge representation framework [4] to deal with linguistic ambiguities, designing a collaborative multi-agent system (MAS) to resolve transliteration inconsistencies in geographical names across Arabic dialects.

2. The Solution: A For Arabic Names Transcription Standardization

The proposed solution is to create a standardized IPA transcription mechanism [5] [6] supported by intelligent MAS. The roadmap (*Fig.1*) presents a comprehensive approach, initiated by data collection from various sources, adapting General Purpose (GP) theory, simplifying syllable structures and stress rules, performing acoustic analysis, measuring the durational (sukuun), while studying co-articulation effects. The goal is to achieve a direct mapping of Arabic patterns to standardized IPA-based transcription with certainty. This was followed by designing the language agent model, connecting to multi-agents [3], comprising tasks, goals, communications, and iterative statistical measures to validate accuracy. A case study using Google Maps shows these agents efficiently handle dialect variations and dynamic updates, achieving 92% accuracy [7], like الروسان (al-Rūsān vs. ar-Rawsān). The following transcription guidelines address key challenges:

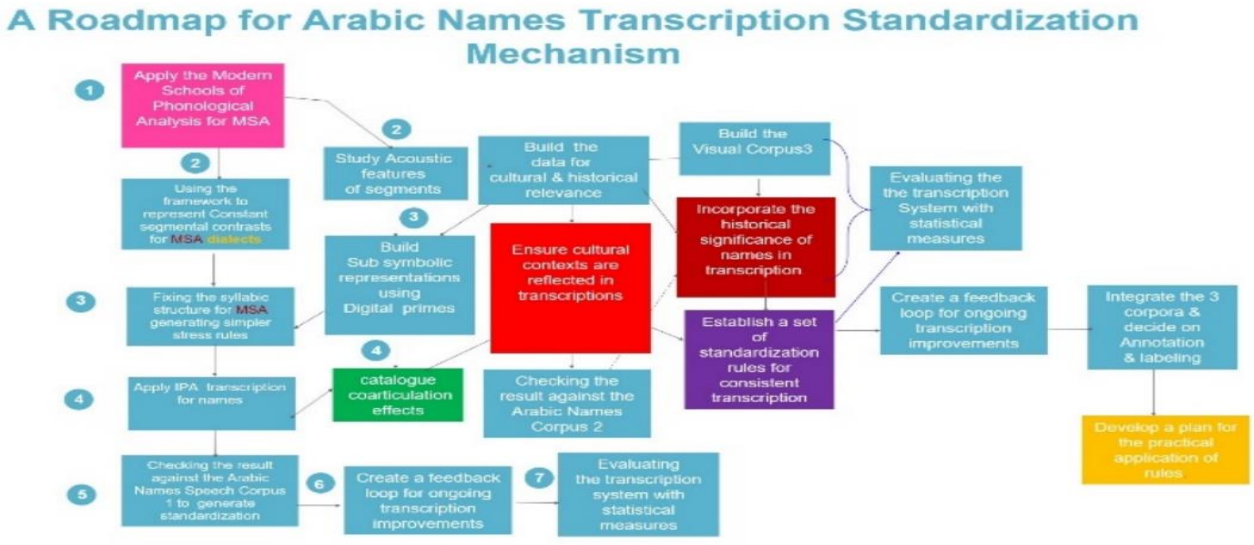


Figure 1 Roadmap

2.1 Syllable Structure Simplification: without altering language integrity (*Fig.2*), **each name, including** suffixes, infixes, and prefixes, is represented in CV format **[8].**

2.2 Management of Phonological Processes: capturing features of phonological processes [9] and acoustic behaviors (Fig.3)**.**

2.3 Treatment of Gemination: representing variabilities in consonants transcription [10] (*Fig.3*).

2.4 Agent Architecture [11]: extracting acoustic cues and defining phonological events.

2.5 Handling of Dialectal Variations: considering sukuun duration in regional pronunciation while maintaining standardization [3] [10] [12].

2.6 Consistent One-to-One IPA Mapping: ensuring uniform representations of Arabic in transcriptions. **[13]**

2.7 Continuous Validation: using advanced statistical measures.

These guidelines remove ambiguity and establish consistent naming conventions for different databases.

3. Leveraging AI Generative Agents for Arabic Names

To operationalize this roadmap, a MAS architecture using AI generative agents [14] (*Fig.4*) is designed to sense, learn, and act independently to achieve goals; the semantic AI ensures context-aware disambiguation; agentic AI validates against standards [15]; and actionable AI provides real-time transliterations [16] [17]. Together, these elements enable generative AI to create culturally valid name variants. Such a unified framework enables its components for perception, reasoning, decision-making, and communication capabilities:

3.1 Environmental Interaction Layer:

- **Other Agents/Environments**: agents interact with multiple environments and other agents through behaviors and commands.
- **Input Channel**: serves as an interface for receiving external information.
- **Actuators**: execute actions in the environment based on decisions [18].
- **Sensors:** collect data from the environment and feed it to the perception unit.

3.2 Processing Core:

- **Perception Unit**: processes sensory symbolic input and recognizes patterns using the signature store for IPA labels.
- **Knowledge Representation**: organizes and stores processed information in memory [19].

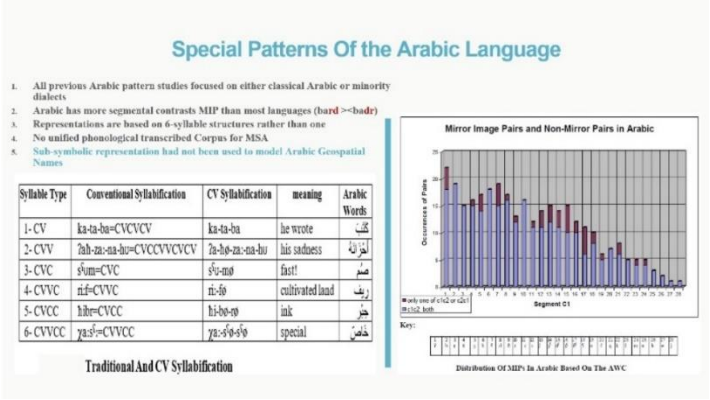


Figure 2 Arabic pattern

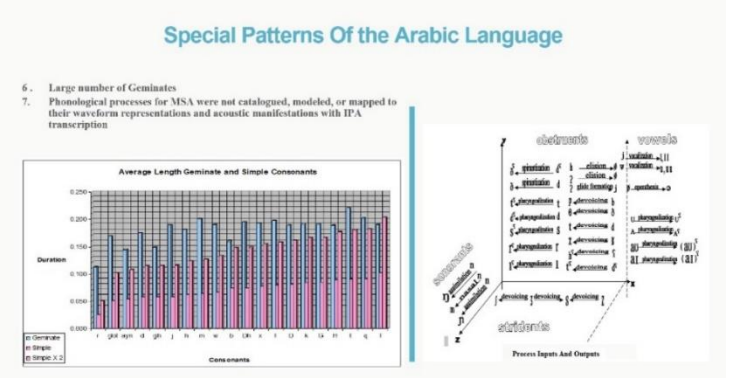


Figure 3 Arabic pattern

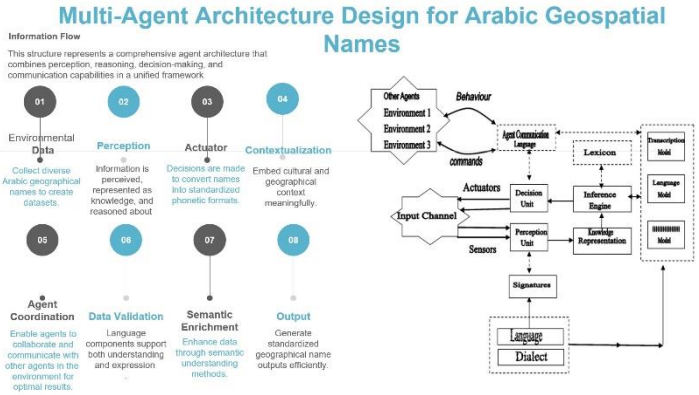


Figure 4 Multi-agent

- **Inference Engine**: analyzes knowledge and its relationships with language components.
- **Decision Unit**: makes action choices based on inference results.

3.3 Communication Elements:

- **Agent Communication Language**: enables structured interaction with other agents.

- **Lexicon**: contains vocabulary, frequencies, word pairs, and terminology for communication.
- The Signatures Entity: identifies patterns or protocols used for recognition and training.

3.4 Language Processing

- **Language and Dialect** Modules: influence signatures and models.
- **Models**: include transcription, language, and other models supporting understanding and generation.

This sophisticated hybridization approach offers the combination of symbolic and sub-symbolic AI to simulate human-like intelligence and possesses advanced capabilities in managing Arabic toponymic data.

4. Use Case: Standardizing Geospatial Names with AI

This case study demonstrates how AI resolves transliteration inconsistencies in Arabic geographical names [20]. It uses a high-level MAS [21] to balance linguistic accuracy and cultural nuance *(Fig.5)*. The core workflow for name processing is:

1. The input channel receives Arabic script (e.g., "الروسان(with geographical coordinates.
2. Semantic & Phonological Analysis:
   a. Semantic Agent: identifies signature features of the place name.
   b. Phonological Agent: applies phonological rules (e.g., /l/ → /r/ assimilation in "al-" before /r/; dialectal and diphthong /aw/ usage).
3. Conflict Resolution: prioritizes context-appropriate forms (e.g., selects "ar-Rawsān" for specific dialects).
4. Standardized Output: generates IPA transcription (e.g., /ʔar.raw.sa:n/) with metadata tagging variants (e.g., "Dialect: ar-Rawsān; Standard: al-Rūsān").

The system features an interactive chatbot that uses **a collaborative multi-agent architecture** [22] to deliver standardized IPA translation.

1. **Context Detection Agent**: analyzes context and determines appropriate transliteration based on geography, history, or morphology.
2. **Batch Processing Agent:** handles bulk operations simultaneously and efficiently**.**
3. **Semantic Analysis Agent:** provides etymological/linguistic insights and meaning**.**
4. **Name Standardization Agent**: ensures cross-database consistency [23].
5. **Interactive Chatbot**: allows users to query names (e.g., "طريق" for "road") and receive contextual transliterations (e.g., *ṭarīq* for "path," *masār* for "route," khaṭṭ for "line," minṭaqa for "area/district," and mamārr for "passage") (*Fig.6*).

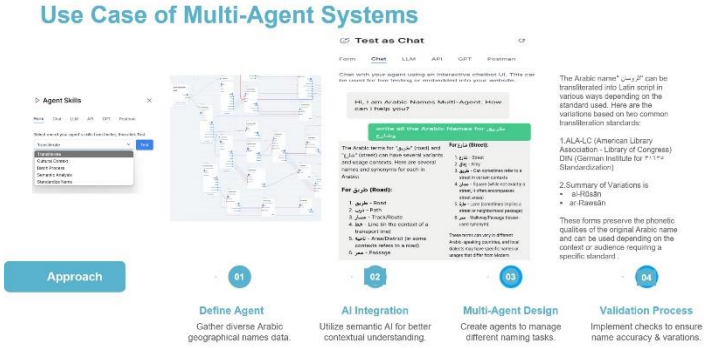


Figure 5. Usercase

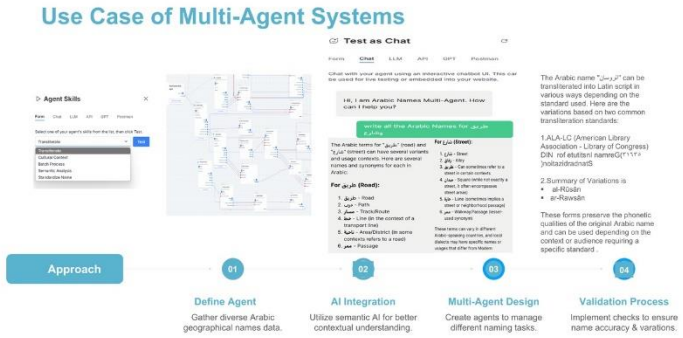


Figure 6. Usercase

The "rise of the syllabic learning curve is faster, 0.028 per iteration, compared to 0.0076 for the phonemic system" [6] (Fig.7). **The benefits** include managing complex linguistic and cultural variations, offering context-aware transliterations, standardized IPA transliteration, integrated databases [22], and a scalable architecture for continuous improvement. The model can be integrated into ALBIS Studio —GENIE AI's no-code platform.

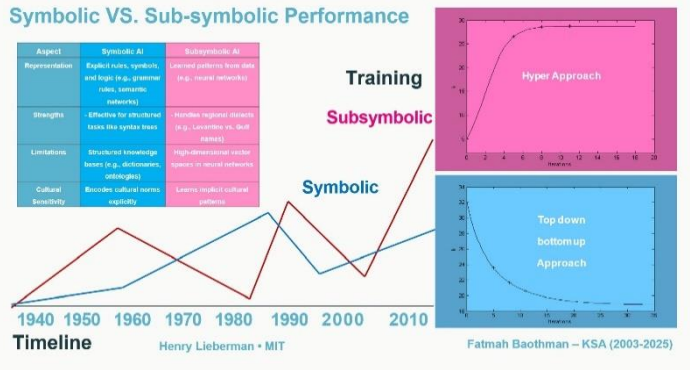


Figure 7. Symbolic vs subsymbolic

5. Framework for Multi-Agent System Collaboration

The framework centers around a MAS [24] (*Fig.8*) with a spherical core hub representing the collaboration among agents to improve performance metrics. The **agent design modules** focus on multi-agent architecture to generate Arabic place names [25], allowing scalability to include dialects, managing complex decisions in IPA standardization for the corpus [26]. The **collaboration strategy** module defines interaction patterns among agents to establish the protocols and methodologies for agents to communicate and coordinate to achieve optimal results. The data management section stores geographical name data. This framework illustrates an eight-step operational cycle surrounding the central processing hub, explained in Table 1 below:

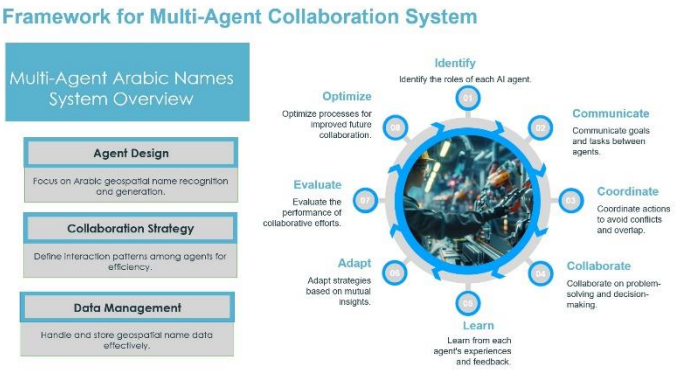


Figure 8. Multi-agent collaboration

Step 1	Identify	Assign the roles of each AI agent within the system.
Step 2	Communicate	Communicating goals and tasks across the agent network.
Step 3	Coordinate	Organize actions to prevent conflicts between agents.
Step 4	Collaborate	Enable agents to collaborate on problem-solving and decision-making for transcription standardization.
Step 5	Learn	Encourage agents to learn from each other, enhancing collective intelligence.
Step 6	Adapt	Adjust strategies based on shared insights and evolving requirements
Step 7	Evaluate	Assess AI frameworks' performance using evaluation metrics.
Step 8	Optimize	Refine processes for continuous collaborations.

6. Conclusion and Future Trends

This AI-driven system ensures improved standardized IPA representation of Arabic geographical names. It addresses challenges through phonological and acoustic analysis, multi-agent collaboration, deploying a hybrid approach, and continuous validation. Future research should focus on standardized IPA Arabic name benchmarks, generative AI, multi-agent hubs, centralized quantum computing, and multi-database integration. Further, augmented reality (AR) for real-time toponymic visualization and quantum encryption for data security are key areas of innovation.

References

[1] S. Hussein, A. Hamza, F. Hasan, "Addressing the Complexities of the Arabic Language with AI Techniques: Challenges and Future Solutions," Journal of Scientific Development for Studies and Research, pp. 128–150, Jun. 2025, doi: 10.61212/jsd/375.

[2] M. Ingleby and F. A. Baothman, "Empty Nuclei in Arabic Speech Patterns and the Diacritic sukuun," John Benjamins, 2006, pp. 83–94. doi: 10.1075/cilt.266.06ing.

[3] F. Baothman, "Phonology-based automatic speech recognition for Arabic," 2003, Doctoral dissertation, University of Huddersfield.

[4] J. Qin, "Knowledge Organization and Representation under the AI Lens," Journal of Data and Information Science, vol. 5, no. 1, pp. 3–17, Feb. 2020, doi: 10.2478/jdis-2020-0002.

[5] H. Saadane, A. Rossi, M. Guidere, and C. Fluhr, "Transcription of Arabic names into Latin," Mar. 2012, pp. 857–866. doi: 10.1109/setit.2012.6482028.

[6] C. Brierley, E. Atwell, M. Sawalha, and B. Heselwood, "A Verified Arabic-IPA Mapping for Arabic Transcription Technology, Informed by Quranic Recitation, Traditional Arabic Linguistics, and Modern Phonetics," Journal of Semitic Studies, vol. 61, no. 1, pp. 157–186, Mar. 2016, doi: 10.1093/jss/fgv035.

[7] P. Gentzel, R. Schlagowski, and J. Wimmer, "Doing Google Maps," Digital Culture & Society, vol. 7, no. 2, pp. 159–184, Dec. 2021, doi: 10.14361/dcs-2021-070208.

[8] M. Ingleby and F. Baothman, "Syllabic Markov models of Arabic HMMs of spoken Arabic using CV units," Oct. 2014, pp. 254–259. doi: 10.1109/cist.2014.7016628.

[9] Z. Imani and A. A. Ahangar, "On the phonological processes in two varieties of Arabic," Onomázein, vol. 64, pp. 137-148, 2024. DOI: 10.7764/onomazein.64.07.

[10] I. Zangar, D. Jouvet, V. Colotte, and Z. Mnasri, "Duration modelling and evaluation for Arabic statistical parametric speech synthesis," Multimedia Tools and Applications, vol. 80, no. 6, pp. 8331–8353, Nov. 2020, doi: 10.1007/s11042-020-09901-7.

[11] Z. Yu and Z. Li, "Architecture description language based on object-oriented petri nets for multi-agent systems," Mar. 2005, vol. 890, pp. 256–260. doi: 10.1109/icnsc.2005.1461197.

[12] M. Elmahdy, S. Abdennadher, W. Minker, and R. Gruhn, "Modern standard Arabic based multilingual approach for dialectal Arabic speech recognition," Oct. 2009, vol. 1, pp. 169–174. doi: 10.1109/snlp.2009.5340923.

[13] C. Brierley, E. Atwell, M. Sawalha, and B. Heselwood, "A Verified Arabic-IPA Mapping for Arabic Transcription Technology, Informed by Quranic Recitation, Traditional Arabic Linguistics, and Modern Phonetics," Journal of Semitic Studies, vol. 61, no. 1, pp. 157–186, Mar. 2016, doi: 10.1093/jss/fgv035.

[14] A. Halevy, C. Yang, J. Fan, L. Cao, N. Tang, and Y. Luo, "VerifAI: Verified Generative AI." Jul. 06, 2023. doi: 10.48550/arxiv.2307.02796.

[15] J. Schneider, "Generative to Agentic AI: Survey, Conceptualization, and Challenges." Apr. 26, 2025. doi: 10.48550/arxiv.2504.18875.

[16] S. Subramani et al., "Mining Actionable Knowledge Using Reordering Based Diversified Actionable Decision Trees," vol. 10041, springer, 2016, pp. 553–560. doi: 10.1007/978-3-319-48740-3_41.

[17] N. Habash, A. Soudi, and T. Buckwalter, "On Arabic Transliteration," springer netherlands, 2007, pp. 15–22. doi: 10.1007/978-1-4020-6046-5_2.

[18] K. Berns, B. Schürmann, and A. Köpper, "Actuators," springer, 2021, pp. 255–272. doi: 10.1007/978-3-030-65157-2_9.

[19] Gelfond M, Kahl Y., "Knowledge Representation, Reasoning, and the Design of Intelligent Agents: The Answer-Set Programming Approach." Cambridge University Press, vol. 2014. ISSN 9781139342124, DOI:10.1017/CBO9781139342124.

[20] K. C. Ryding, "Arabic transcription/transliteration/romanization," pp. 145–148, Mar. 2014, doi: 10.1017/cbo9781139151016.014.

[21] T. C. Son et al., "Action Languages for Multi-Agent Domains," Artif. Intell., vol. 302, p. 103601, 2022. DOI: 10.1016/j.artint.2021.103601.

[22] C. Zhang, X. Yang, and J. Xi, "An architecture for intelligent collaborative systems based on multi-agent," Apr. 2008, pp. 367–372. doi: 10.1109/cscwd.2008.4537008.

[23] ArabLEX Consortium, "Database of Arab Names (DAN)," Language Res. J., vol. 12, no. 3, pp. 45-62, 2022. DOI: 10.17605/OSF.IO/9Z8KQ.

[24] Y. Boulid, A. Souhar, and M. E. Elkettani, "Multi-agent Systems for Arabic Handwriting Recognition," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, p. 31, Jan. 2017, doi: 10.9781/ijimai.2017.03.012.

[25] Q.-S. Li, L. Diao, M. Zhang, H. Chu, and M. Li, "Collaboration strategy for software dynamic evolution of multi-agent system," Journal of Central South University, vol. 22, no. 7, pp. 2629–2637, Jul. 2015, doi: 10.1007/s11771-015-2793-2.

[26] M. Sawalha and E. Atwell, "Linguistically Informed and Corpus Informed Morphological Analysis of Arabic," Comput. Linguist., vol. 48, no. 2, pp. 301-318, 2022. DOI: 10.1162/COLI_a_00432.

Artificial Intelligence and other innovative tools for the romanization of geographical names

Catherine Cheetham
Convenor of the Working Group on Romanization Systems
United Kingdom Permanent Committee on Geographical Names
United Kingdom of Great Britain and Northern Ireland
Email: ccheetham@pcgn.org.uk

Romanization is the process of converting text from a non-Roman writing system into the Roman script. It is primarily used for products intended for a language community outside that of the local or national language – e.g. on mapping and road signs. And when romanization is undertaken through a systematic process, i.e. through consistent application of a scientific system, this contributes to standardization, whereby the same romanized results will be achieved by different users.

UNGEGN is concerned with the process of romanization through its Working Group on Romanization Systems (WGRS), the basic mandate of which is to consider and reach agreement on a single romanization system for each non-Roman writing system for use by the United Nations and the global community. These systems are for application to geographical names and should be proposed by a (donor) country. New systems are referred to WGRS for consultation with the proposing country, and subsequently to UNGEGN for review and endorsement through the Group of Experts' decisions. These decisions are then passed to ECOSOC to adopt. Following ECOSOC's adoption of these decisions, the new systems become a globally agreed United Nations standard.

In addition to this core role, and in line with UNGEGN's Strategic Plan 2021-2029 (item 1-iii-10), WGRS has *undertaken to evaluate and assess innovative methods to be used for romanization, e.g. machine transliteration*. It is notable that at the time of writing in mid-2025, evidence of such innovative methods has increased dramatically in recent years (even months). In previous UNGEGN plenary sessions (held biannually in 2021 and 2023) there was little to report on such innovative methods, a 'Transcriptor' tool developed in the Kingdom of the Netherlands, described in their 2021 report, *GEGN.2/2021/CRP.120*, being an exception.

At the 2025 UNGEGN session, by contrast, there were some 10 reports describing different aspects meeting this action. These ranged from Optical Character Recognition in the capture and analysis of names, to an algorithm developed for automating Arabic romanization that will be assessed for its transferability to other writing systems.

So, it appears that romanization is poised for significant transformation due to the increasing efficiency and accessibility of AI and innovative automated tools. These advances are likely to improve the accuracy, scalability, and accessibility of transliteration, particularly benefiting languages and scripts that have historically been underrepresented.

There already exist a number of open source and bespoke tools using AI-powered Optical Character Recognition (OCR) technologies, such as those integrated into *Google Lens*. These tools can detect text in one script and facilitate its translation or transliteration into another, streamlining the process of converting printed or handwritten text into digital formats. Moreover, some of the translation platforms, e.g. *Microsoft Translator* and *Google Translate*, have incorporated features that allow users to input text phonetically. These systems can convert the phonetic input into the desired script.

The potential benefits of these innovations include improved accuracy and context awareness: AI can understand context by learning from large linguistic datasets, reducing transliteration errors, particularly in languages with homonyms or multiple dialects.

Example case studies at UNGEGN May 2025

Oman, in *GEGN.2/2025/142/CRP.142* titled *"Addressing Inconsistencies in Romanization: Towards an Integrated Program for the Unified Arabic System for the Romanization of Names"* and in its presentation on this report, introduced an initiative that aims to standardize the representation of Arabic geographical names through addressing the fragmentation in current transliteration practices. The report advocated for the application of an integrated, software-based solution grounded in the United Nations-endorsed Unified Arabic Romanization System.

The prototype application incorporates algorithmic logic based on phonetic and grammatical rules from both Arabic and English. It aims to provide consistent romanized outputs and includes features such as manual input, bulk processing via Excel files, and a reverse transliteration function. These features are designed to enhance usability across various sectors.

Indonesia, in *GEGN.2/2025/126/CRP.126*, explored how AI technologies can automate the extraction of geographical names from various data sources (so-called geographical name data web scraping), potentially aiding in the creation and updating of romanized name databases.

Indonesia, also, in *GEGN.2/2025/130/CRP.130*, discussed the application of deep learning and computer vision techniques to recognize and process geographical names, which could include aspects of romanization.

India, in *GEGN.2/2025/137/CRP.137* and in its presentation for the National Dialogue, described planned developments to automate transliteration between Roman and Devanagari scripts with *Linguistic Expert Interface*.

Norway, in *GEGN.2/2025/77/CRP.77* and in its presentation for the National Dialogue, described the development and use of machine learning (ML) to identify geographical names in scanned maps. The project used *Google Cloud Vision Text Detection* to analyse maps of Norway. This focused on Roman-script materials, but this type of project could be expanded, perhaps initially using a staged approach of OCR before romanization analysis, for transliteration.

In summary, these and other countries are now actively exploring and implementing AI and machine learning technologies to enhance transliteration processes, preserve linguistic diversity, and integrate AI into their names standardisation processes. Importantly, several reports also acknowledge that as AI continues to advance, sustained human oversight is essential to guide its development responsibly. The Working Group on Romanization Systems remains keen to learn of other such advances in this field and welcomes contributions from experts through the UNGEGN Secretariat.

FROM THE COUNTRIES

Taxonomy of the twenty-five cities of Mozambique

Paulino Baptista Rafael Bata
Director of Central Services for Studies and Dissemination – INGEMO
Mozambique
Email: paulinobaptistabata@yahoo.com.br

1. Introduction

The main objective of this study is to classify the geographic names of a selection of twenty-five cities in Mozambique (Maputo, Matola, Xai-Xai, Chokwe, Chibuto, Inhambane, Maxixe, Vilankulo, Beira, Dondo, Chimoio, Manica, Tete, Moatize, Quelimane, Mocuba, Gurué, Nampula, Nacala, Island of Mozambique, Angoche, Pemba, Montepuez, Lichinga, and Cuamba) according to the taxonomic proposal by Dick (1990a). For better classification, each name is analyzed according to their meaning and pronunciation in its original language.

Data was collected through interviews to learn about the origin of each geographical name.

The structure of the study is divided into 5 parts:

1. Introduction;
2. Literature Review;
3. Presentation and Analysis of the Data;
4. Conclusion and
5. Bibliography.

2. Literature Review

2.1. Operational concepts

2.1.1. **Geographic name** - is the name of a place and may include names of access routes, geographical features (such as seas, mountains, etc.) (Furtado, 1959; Crystal, 1987; and Seabra, (n/d)).

2.1.2. **Cities of Mozambique** - according to the Ministry of State Administration (MAE, 2014), the concept of a city is defined by the functions performed by the place, the situs, taking into account levels of economic, social, cultural, and political-administrative development. It also considers the degree of urban land use, since it includes municipalities, villages, and the headquarters of administrative posts and localities established by law.

2.2. Taxonomic Proposal by Dick (1990a)

The taxonomic proposal by Dick (1990a) is divided into taxa of physical nature (11) and those of human nature (16).

The first category is divided as follows:

1. Astrotoponyms (refer to celestial bodies in general);
2. Cardinotoponyms (refer to geographical positions in general);
3. Chromotoponyms (refer to the color scale);
4. Dimensiotoponyms (refer to the dimensions of geographical features);
5. Phytotoponyms (refer to plants);
6. Geomorphotoponyms (refer to topographic forms);
7. Hydrotoponyms (refer to hydrographic features in general);
8. Litotoponyms (refer to minerals and soil composition);
9. Meteorotoponyms (refer to atmospheric phenomena);
10. Morfotoponyms (refer to geometric shapes); and
11. Zootoponyms (refer to animals).

The second category, in turn, is divided as follows:

1. Animotoponyms (refer to human psyche);
2. Anthropotoponyms (refer to first names, family surnames, diminutives, nicknames, etc.);
3. Axiotoponyms (refer to titles in general);
4. Corotoponyms (refer to names of cities, countries, etc.);
5. Chronotoponyms (refer to chronological indicators, such as the adjectives new, old, etc.);
6. Dirrematoponyms (refer to crystallized expressions);
7. Ecotoponyms (refer to dwellings in general);

8. Ergotoponyms (refer to elements of human material culture).
9. Etnotoponyms (refer to ethnic groups);
10. Hierotoponyms (refer to sacred names of different beliefs, places of worship, religious figures, religious associations, and dates related to these facts);

    10.1. Hagiotoponyms (refer to names of saints in Roman hagiology);

    10.2. Mitotoponyms (refer to mythological entities);
11. Historiotoponyms (refer to the country's history, its historical figures, and important dates);
12. Hodotoponyms (refer to roads and rural or urban routes of communication);
13. Numerotoponyms (refer to numbers in general);
14. Poliotoponyms (refer to aspects related to population centers or urban settlements);
15. Sociotoponyms (refer to professional activities, workplaces, and gatherings of a group);
16. Somatoponyms (refer to parts of the human or animal body).

3. History and Taxonomy of the 25 Geographic Names of Cities in Mozambique

The first name is "Portugueseized" and the second is the original form in a local language.

a) Maputo/KaMaputsu - from the Rhonga language, it comes from *Mathuthu* in the Ciswati language, meaning "land of lord Mathuthu." It is an Anthropotoponym.

b) Matola/KaMatsolo - from the Rhonga language, it comes from the anthroponym Matsolo, which means "knees" in the Xirhonga language. KaMatsolo means "zone of lord Matsolo." It is an Anthropotoponym.

c) Xai-Xai/Ncayicayi – comes from the word tsayitsayi which means "to wink" in the Xichangana language. It is an Anthropotoponym.

d) Chokwe/Kaxokwe - comes from the name of the first evangelist of the Free Methodist Church, Xokwe, in the Citshwa language. It is an Anthropotoponym.

e) Chibuto/Ximbutsu - comes from the word *xibuthu* which means "place of concentration" in the Changana language. It is a Poliotoponym.

f) Inhambane/Nyambani - comes from the expression nyumbani, meaning "inside the house" in the Gitonga language. It is a Dirrematoponym.

g) Maxixe/Matshitshi - comes from the name of the Masisi family who dominated the area before the Portuguese presence. It is an Anthropotoponym.

h) Vilankulo/Kavilankulu - derived from the name of the local king, Mr. Vilankulu, meaning "boiled until it grows" in the Citshwa language. It is an Anthropotoponym.

i) Beira – Portuguese for "an area that is on the edge of the sea". It is a Hidrotoponym.

j) Dondo/Kudondo - originates from the word n'tondo, the name of a fruit tree. It is a Phytotoponym.

k) Chimoio/Cimoyo - derived from the word moyo "heart", with Cimoyo meaning "little heart" in the Ciwute language. It is a Somatoponym.

l) Manica/Manyika - comes from the word Nyika meaning "land" in the Cimanyika language. It is a Litotoponym.

m) Tete/Ntete - comes from ntete, meaning "reeds" in the Nyungwe language. It is a Fitotoponym.

n) Moatize/Mwatidzi - comes from the word Mwatubvi, meaning "feces" in the Nyungwe language. It is a Sociotoponym.

o) Quelimane/Kalimani - is an expression in the Echuwabu language meaning " to cultivate". It is a Sociotoponym.

p) Mocuba/Mukhupa - comes from the word nikhuba, meaning "porridge" in the Elomwe language. It is a Historiotoponym.

q) Gurué/Kuruwe - comes from the expression kuruwe, meaning "the imitation of a quail's song" in the Elomwe language. It is a Historiotoponym.

r) Nampula/Wamphula - comes from the anthroponym Namphula "lord of the big nose", in the Makhuwa language. It is an Anthropotoponym.

s) Nacala/Onakhala - comes from the expression nakhala, meaning "are we really going to stay?" in the Emakhuwa language. It is a Historiotoponym.

t) Island of Mozambique - comes from the Arabic expression Mussa al biki, meaning "Mussa, the Sultan's tailor ". It is an Anthropotoponym.

u) Angoche - derived from the expression *akuwa axi*, meaning "they will fall" in the Ekoti language. It is a Historiotoponym.

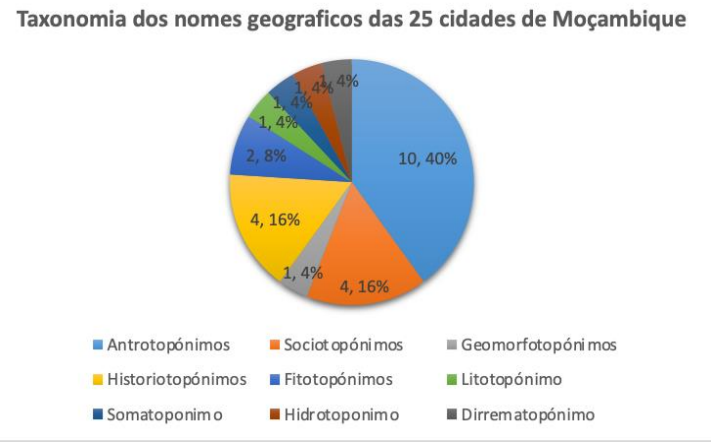
v) Pemba/Uphemba - comes from the word kuphempa, meaning "to float" in the Makhuwa language. It is a Sociotoponym.

w) Montepuez/N'thipwehi - comes from the name N'thipwi, meaning "lord of charms" in the Emakhuwa language. It is an Anthropotoponym.

x) Lichinga/Kunciinga - a Portuguese adaptation of Kwiciinga derived from the word *diciinga*, meaning "that which forms a barrier" in the Yawu language. It is a Geomorphotoponym.

y) Cuamba/Mukwapa – comes from the name of a king Mukhapa, who wore cloths called mukhwapa in the Emakhuwa language. It is an anthropotoponym.

3.1. Percentage graph of tax rates of city names in Mozambique



4. Conclusion

The aim of this study was to analyze the classification of geographical names in a selection of 25 cities in Mozambique, according to Dick's taxonomic proposal (1990a). The study led us to conclude that the geographical names of these cities originate from both human motivation and physical features, with the human-related factors having a higher percentage. Particularly anthroponymic names, which constitute 40% of the names, followed by sociotoponyms and historiotoponyms at 16%. Meanwhile, names derived from physical features have a much lower percentage. Overall, the study concludes that, taxonomically, city names in Mozambique are predominantly anthroponymic.

5. Bibliography

CRYSTAL, D. (1987). The Cambridge Encyclopedia of Language. Cambridge, England: Cambridge University.

SEABRA, M. (n.d.). Reference and Onomastics. UFMG. http://www.filologia.org.br/ileel/artigos/artigo_442. Accessed on 06/20/2014.DICK, M. V. P. A. (1990a). The Toponymic Motivation and Brazilian Reality. São Paulo.

FURTADO, C. (1959). Economic Formation of Brazil. São Paulo: Companhia Editora Nacional/ Publiufolha.

FROM THE DIVISIONS

The Portuguese-Speaking Division Holds Its First Seminar On Geographical Names

Paulino Baptista Bata, *Director of Central Services of Education and Dissemination,*
National Institute of Geographical Names of Mozambique (INGEMO, IP)
Mozambique
Email: paulinobaptistabata@yahoo.com.br
and
Ana Cristina Resende, *Head of Geographical Names,*
Brazilian Institute of Geography and Statistics
Brazil
Email: ana.resende@ibge.gov.br

The Portuguese-speaking Division (DPLP) held its First Seminar "*Geographical Names and Socio-Economic Development: the experience of Mozambique, Cabo Verde, and Brazil*" on May 30, 2025, in a virtual format via the Google Meet platform.

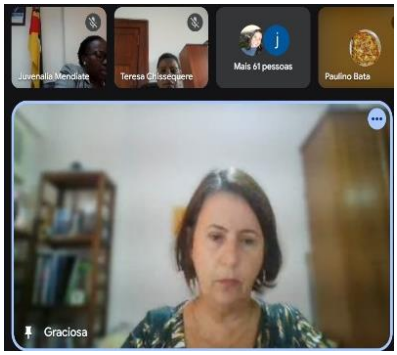
The seminar, part of the Division's activities planned for 2025, featured the following speakers: Dr. Graciosa Moreira (Brazil), Deputy Manager of Geographical Names at the Brazilian Institute of Geography and Statistics (IBGE); Prof. Dr. José Maria Semedo (Cabo Verde), university professor and researcher specialized in African studies and member of the Toponymy Commission of Cabo Verde; and MSc Teresa Chissequere (Mozambique), Director of the Municipal Office for Urban Development and member of the Municipal Toponymy Commission in the Municipality of Maputo. The seminar was moderated by the Divisional Chair, Ana Cristina Resende (Brazil).

The seminar, conducted entirely in Portuguese, was open to the public and attended by over 150 participants from the three countries currently active in the Division, connected through 75 devices. Below is the list of participating institutions by country:

Mozambique: National Institute of Geographical Names (INGEMO, IP), represented by its Director General, Deputy Director General, and all toponymy staff; State Representation Services in the Provinces; Provincial Executive Councils; Municipal Council of the City of Maputo; National Center for Cartography and Remote Sensing (CENACARTA); National Directorate for Territorial Planning (DNOT); National Institute of Statistics (INE); and the District Administrator of Mutarara.

Brazil: Brazilian Institute of Geography and Statistics (IBGE), represented notably by the Deputy Coordinator of Cartography and numerous staff members; Ministry of Planning and Budget; Bahia State Court of Municipal Accounts; University of São Paulo; Federal University of Rio de Janeiro; Fluminense Federal University; National Institute of Industrial Property (INPI); State Secretariats of Education of São Paulo and Rio de Janeiro; Directorate of the Brazilian Army's Geographic Service.

Cape Verde: National Institute for Territorial Management (INGT), represented notably by its President.



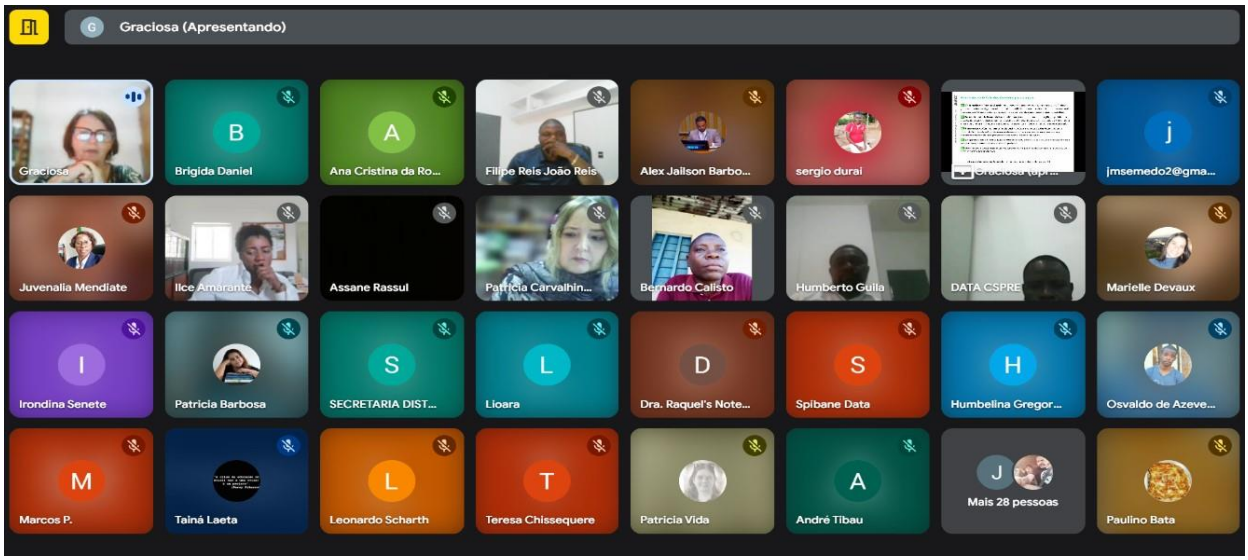
"Speakers: José Semedo, Graciosa Moreira, and Teresa Chissequere, and moderator Ana Cristina Resende"

Also noteworthy was the participation of a representative from the IPGH's "José Joaquín Hungría Morell" Geographical Names Committee and member of the UNGEGN Latin America Division.

The event was jointly organized by the members of the Division and hosted by INGEMO, IP, in its role as Divisional Secretariat.

identifying the geographical influence on the quality or characteristics of the product or service.

Geographical Indications bring benefits such as product valorization; greater visibility and market access; tourism promotion; strengthening of producer organizations; enhancement of cultural heritage; and, in the above-mentioned case, gender empowerment.



"Participants in the First Seminar on Geographical Names of the Portuguese-speaking Division"

The event aimed to discuss the experiences of Lusophone countries regarding the importance of standardizing geographical names for socio-economic development.

The first speaker, Dr. Graciosa Moreira (Brazil), addressed the topic: "Geographical Names in Geographical Indications: Identity and Tradition."

She explained that in Brazil, a geographical indication (GI) is used to identify the origin of products or services when a location has become known for a specific product, or when a product's characteristic is linked to its geographic origin. As an example, she used the GI "Panelas de Barro de Goiabeiras" (Goiabeiras Clay Pots), made by female artisans in Goiabeiras, a neighborhood in the State of Espírito Santo. This designation is regulated by Law No. 9.279/96 – Industrial Property Law.

She also explained that, according to Normative Instruction (IN) No. 25/2013 from the National Institute of Industrial Property (INPI), the registration of a GI requires: the geographical name; description of the product or service; documents proving the applicant's legitimacy; GI usage regulation; documents proving the existence of control mechanisms; evidence that the name has become known as a production or service provision center; and elements

The second speaker, Prof. Dr. José Maria Semedo, focused on: "Geographical Names and Economic and Social Development."

He explained that in the archipelago of Cabo Verde, the naming of islands and geographic features follows a colonial-era nautical tradition, mostly of European religious origin. African-origin names, which are often in the minority, are frequently linked to nature and cultural expressions (e.g., *tabanca*, *batuque*).

Streams often bear the names of their first owners, patron saints, or dominant activities in the early days of settlement.

He showed how the dates of feature naming and the toponymic changes over time reflect different stages in the country's socioeconomic development.

The third speaker, Eng. Teresa Chissequere, focused on the post-independence naming process in Mozambique and related legislation. She highlighted Decree-Law No. 10/76, of March 13, and Order No. 267/76, of November 16, which initiated the renaming of colonial place names (e.g., Lourenço Marques to Maputo); Municipal Toponymy Regulation, which established the Municipal Toponymy Commission and standardizes all initiatives related to assigning and altering place names at the municipal level in Maputo (approved through Resolution No. 42/AM/2006, of

June 15); Basic Law of Local Authorities – specifically item (s) of paragraph 3, Article 45 of Law No. 02/97, of February 18, as revised by Law No. 6/2018, of August 3; Decree-Law No. 1/2014, of May 22, which regulates national-level place name assignment and alteration procedures.

After its independence, Mozambique began listing and renaming administrative units, places, infrastructure, and public roads, especially those linked to exploitation and discrimination, in line with the thoughts, principles, and values of the revolutionary period, favoring pre-colonial names (e.g., Av. Pinheiro Chagas renamed to Eduardo Mondlane).

The relationship between geographical names and socio-economic development is considered complex but can be observed through the identification of places of economic activity, transportation patterns, and local culture.

Currently, the assignment of geographical names in Mozambique prioritizes proposals of native or sociologically appropriate names and their respective Bantu spellings. In Maputo, the former numerical district designations (1, 2, 3, 4, 5, Catembe, Inhaca) were replaced with names that turned them into territorial references, reflecting collective memory: KaMpfumu, Nhlamankulu, KaMaxakeni, KaMavota, KaMubukwana, KaTembe, and KaNyaka.

Prof. Dr. Patrícia Carvalhinhos and Deputy Director of INGEMO, IP, Juvenália Mendiate participating in the seminar

The Mozambican speaker also highlighted the key role of standardized toponymy in defining political-administrative units and basic urban structures (districts and their neighborhoods) and other basic units for the formation of urban spaces (blocks) delimited by roads, avoiding ambiguity and reducing political tension between districts. For example, in Maputo, toponymy supports traffic management, road upgrades, delivery services, and public services (firefighters, ambulances), ensuring each place has a unique, identifiable name.

The well-attended seminar served as a platform for knowledge exchange and capacity building in toponymy, highlighting the interplay between humans, the environment, language, and culture.

International Conference - Toponyms of historical linguistic minorities in Italy, Trento, 2 - 4 July 2025

Gianluca Casagrande
Adviser, Italian Geographical Society and European University of Rome, Italy
Chair of the Romano-Hellenic Division
Email: Gianluca.Casagrande@unier.it
and
Elena Dai Prà
Chair, International Geographical Union Commission on Toponymy,
jointly with the International Cartographic Association - University of Trento - Italy
Email: elena.daipra@unitn.it

From 2 to 4 July 2025, the Autonomous Region of Trentino-Alto Adige/Südtirol hosted the International Conference *Toponyms of historical linguistic minorities in Italy*, promoted by the Romano-Hellenic Division (RHD) of the United Nations Group of Experts on Geographical Names (UNGEGN), and the Joint ICA/IGU Commission on Toponymy, and organized by the Department of Humanities at the University of Trento and by the Autonomous Region of Trentino-Alto Adige/Südtirol. The conference adopted a multidisciplinary approach, engaging experts from various fields in reflecting on how place names are inherently linked to identity, historical memory, and the safeguarding of local cultures.

The introductory session was conducted by Elena Dai Prà (University of Trento), Chair of the Joint ICA/IGU Commission on Toponymy, and Luca Guglielmi, Councillor for Linguistic Minorities of the Autonomous Region of Trentino-Alto Adige/Südtirol.

Gianluca Casagrande (European University of Rome), Chair of the Romano-Hellenic Division of UNGEGN, presented the status quaestionis of national toponymic developments; Elena dell'Agnese (University of Milano-Bicocca), representing SoGeI, underlined how historical toponymy could serve as a bridge towards peace and mutual understanding.

Introductory speeches were given by Andrea Cantile (IGMI and University of Florence), Honorary Chair of the Romano-Hellenic Division of UNGEGN, and Valeria Piergigli (University of Siena). Both scholars discussed the issue of standardising minority toponymy - a crucial matter for ensuring consistency, recognisability, and the protection of place names, preventing linguistic fragmentation and fostering accurate representation of identities in official documentation. As Cantile pointed out, the matter of standardisation remains far from being fully regulated and resolved, due to a fragmented legislation and its uneven application. Councillor Guglielmi called for renewed dialogue between central institutions and local and regional bodies.

The session on Thursday 2 July, entitled La toponomastica minoritaria dell'Italia meridionale, was chaired by Cosimo Palagiano (Sapienza University of Rome) and offered an overview of the dynamics of toponymy and identity across Southern Italy. Patrizia Miggiano (Pegaso Telematic University) analysed the "toponymic threads" of Apulia. Leonardo Mercatanti (University of Palermo) explored the Arbëreshë and Gallo-Italic communities in Sicily. Valentina Campesi (Museum Ladin, Südtiroler Landesmuseen) presented a comparative study of toponymic policies implemented in Calabria for the Arbëreshë, Occitan, and Grecanic communities. Silvia Siniscalchi (University of Salerno) focused on Basilicata, where Arbëreshë toponyms narrate a fascinating interweaving of documented history, legends and landscape. Domenico Proietti (University of Campania Luigi Vanvitelli) offered a diachronic perspective on minority toponymy in Southern Italy, analysing the geographical distribution and future prospects for safeguarding these linguistic communities. Annalisa D'Ascenzo (Roma Tre University) extended the discussion to the northern part of the former Kingdom of Naples, where minority toponymy continues to reflect traditions despite social and linguistic changes. Giuseppe Scanu (University of Sassari), addressing the case of Sardinia, illustrated the role of toponymy with particular attention to bilingualism and policies aimed at recovering traditional place names.

In the afternoon, a first Round Table, entitled Le minoranze linguistiche storiche viste dall'esterno, was chaired by Gianluca Casagrande. Participants in the debate included Silvia Dal Negro (Free University of Bozen-Bolzano), Christiane Dunoyer (Centre d'Études Francoprovençales

René Willien), Vittorio Dell'Aquila (University of Milan), Franco Finco (University of Klagenfurt), Matjaž Geršic (Anton Melik Geographical Institute), Joan Tort Donada (University of Barcelona), Genc Lafe (University of Salento), and Francesco Giannachi (University of Salento). During this discussion, scholars offered an external perspective on Italy's historical linguistic minorities, analysing their representation, protection policies, and identity significance from geographical, linguistic, and cultural viewpoints.

Wednesday, 3 July opened with the second session, La toponomastica minoritaria dell'Italia settentrionale, chaired by Federica Ricci Garotti (University of Trento). On this occasion, focus was placed on the historical linguistic minorities of Northern Italy, in particular Patois, Franco-Provençal, Occitan, Ladin, Mòcheno, Cimbrian, Slovene, and Friulian. Numerous academics and experts in the field contributed to the discussions. Guido Lucarno (Università Cattolica del Sacro Cuore) delved into the toponymy of linguistic minorities in north-western Italy, concentrating on the complex cases of the Aosta Valley, Piedmont, and Liguria. David Colmano (Autonomous Province of Bolzano) and Johannes Ortner (independent researcher) presented a survey of place names in South Tyrol, from which a database was created, aimed at the production of toponymic maps. Lydia Flöss (Superintendence for Cultural Heritage and Activities - APT) explored the role of the Dizionario Toponomastico Trentino (Trentino Toponymic Dictionary) in collecting and systematising Cimbrian and Mòcheno toponymy within the rich geodatabase developed by the Autonomous Province of Trento. Viviana Ferrario (IUAV University of Venice) addressed the toponymy of linguistic minorities in north-eastern Italy, with particular attention to the Friuli-Venezia Giulia. Federico Vicario (University of Udine - Friulian Philological Society) continued tracing the steps of the evolution of Friulian toponymy over time. Guglielmo Cevolin (University of Udine) examined place names in the same region, focusing on the legal aspects of "linguistic rights".

A second Round Table, entitled La toponomastica delle lingue minoritarie dell'arco alpino, chaired by Elena Dai Prà. Participants of this discussion were representatives of associations, cultural institutes, and research centres focused on the minorities of Northern Italy: Michele Musso (Walser Cultural Association Augusta of Issime, Aosta Valley), Teresa Geninatti Chiolero (Regional Council of Piedmont for Historical Linguistic Minorities), Fiorenzo Nicolussi Castellan (Cimbrian Cultural Institute of Luserna / Kulturinstitut Lusérn), Claudia Marchesoni (Mòcheno Cultural Institute / Bersntoler Kulturinstitut), Sabrina Rasom (Ladin Cultural Institute Majon di Fascegn), Werner Pescosta (Ladin Institute Micurá de Rü), Denni Dorigo (Ladin Cultural Institute Cesa de Jan), Nicola Cassisi (Federation Ladina of Veneto), Andreja Kalc (Slovenian Research Institute of Trieste), and Laura Sgubin (Service for Minority and Regional Languages Abroad, Autonomous Region of Friuli-Venezia Giulia). The Round Table provided an engaging exchange among the various experts, aiming to foster the sharing of suggestions and opinions, based on the speakers' experiences of research and intercultural exchange, regarding the need to standardise the issue of toponymy in diverse Italian and European minority linguistic contexts. As Elena Dai Prà recalled, scientific research in this field can serve as a bridge between society and institutions.

FROM THE WORKING GROUPS

Report from Working Group on Geographical Names as Cultural Heritage for UNGEGN Bulletin

Emily Lethbridge
Associate Research Professor, Árni Magnússon Institute for Icelandic Studies, Reykjavik
Iceland
Email: emily.lethbridge@arnastofnun.is
and
Nicole Halseth
Senior Policy Analyst, Canada Centre for Mapping and Earth Observation
Natural Resources Canada
Canada
Email: nicole.halseth@NRCan-RNCan.gc.ca

"Geographical names are much more than just words on paper"

"Geographical names are much more than just words on paper" was one of Ms. Aluki Kotierk's closing remarks in the keynote speech she delivered at the first of two side-events organised by the WGGNCH and the Focus Group on Indigenous Geographical Names at the 2025 UNGEGN session. These words held great meaning as Ms Kotierk had just been elected chair of the 24th session of the United Nations Permanent Forum on Indigenous Issues, which was being convened concurrent to the 2025 UNGEGN session. It was a great privilege to have Ms Kotierk share her thoughts on the cultural importance of geographical names with us.

Ms. Kotierk began by talking about the significance of the Indigenous name *Tallurutiup Imanga* and how this name for the inshore waters around the island of *Tallurutit* (*Devon Island*) encapsulates many aspects of Inuit worldview. The 17th-century colonial name *Lancaster Sound* has no meaning for Ms. Kotierk. She explained how *Tallurutiup Imanga*, on the other hand, communicates information about the area's topography (the literal meaning of the name is "the body of water that is close to the jaw-line with the facial tattoo on the chin") as well as about the ecological richness found there, and the cultural associations of the area's plentiful marine life – for example with the sea-goddess Nuliajuk whose hair must be combed to ensure good hunting.

She closed her talk with a powerful example of the symbolic power of commemorative naming. In 2022, the City Council of Iqaluit changed the street-name *Federal Road* to *Sivumugiaq*. The name means 'moving forward' and the orange colour of the street sign also conveys, in a visual way, the memory and living history of the system that sent children of Ms. Kotierk's parents' generation to residential schools.

Ms. Kotierk made some valuable practical recommendations to the Working Group as well. She encouraged the Working Group to strengthen its relationship with the United Nations Educational, Scientific and Cultural Organization (UNESCO) as the Secretariat for the Global Taskforce for the Decade of Indigenous Languages. In her view, exploring the Global Action Plan Output 6 ("Indigenous languages are sustained, as a vehicle of living heritage and biodiversity, whilst participation in - and access to - all forms of culture are enhanced for Indigenous Peoples") and Output 10 ("Public and private partnerships are firmly established to place on the global agenda a long-term commitment to the preservation, revitalization and promotion of Indigenous languages") would be beneficial, as well as engaging with the ad hoc groups which relate to Indigenous language transmission

Group photo of all speakers at Forum part I

and resilience building. UNGEGN is in a position to have a very positive impact on the maintenance and revitalization of Indigenous languages by using and promoting Indigenous geographical names.

This guidance, and more broadly the role of cultural heritage and Indigenous geographical names, was reflected throughout the discussions and decisions taken by Member States at the 2025 session. The decisions support the revitalization of the Working Group on Geographical

Photo from Forum part II

Names as Cultural Heritage and its workplan/Terms of Reference. They encourage members of the Focus Group to continue work related to the International Decade of Indigenous Languages, and call upon the Working Group and Member States to build relationships with other relevant entities of the United Nations system, such as the Permanent Forum on Indigenous Issues and the UNESCO. An expert from the Livonian Institute at the University of Latvia was appointed to act as a liaison officer to these bodies. Progress was also made regarding specific deliverables of the Working Group, both during and after the session. We thank all Member States for their participation, and all those involved in presenting at the two side events.

WGGNCH Global Survey – Update

The Working Group's Online Global Survey was launched in late April and promoted at the session. To date, we have had over 50 responses, which is fantastic! We are grateful to everyone who has taken the time to answer. If any representatives of Member States wish to submit further answers (short or long), the online form remains open: https://forms.office.com/r/uGxre4Ubhu.

All information supplied will be used to help us develop specific parts of our work plan. We will hold an online webinar in November (details to be circulated), open to everyone interested, at which we will present an analysis of the survey results.

Working modalities and participation

We have revised the structure of the Working Group, reducing the six Focus Groups named in UNGEGN's Strategic Plan and Programme of Work 2021-2029 to four. These are:

i. Engagement and communication (led by co-Convenors);
ii. Focus Group on Indigenous Geographical Names (led by Canada);
iii. Focus Group on Commemorative Naming, including urban naming (led in the interim by Australia);
iv. Focus Group on Legislative and Policy-driven aspects of geographical names as cultural heritage (leadership to be established).

We are developing a schedule that involves online Working Group meetings (recorded for those who cannot attend) every three months. The possibility of an in-person meeting in 2026 is also being explored.

Any representative interested in participating in any of the Working Group's activities (whether as part of the larger Group or one of the Focus Groups) is warmly encouraged to contact the co-convenors (emily.lethbridge@arnastofnun.is, rafe.benli@transport.vic.gov.au), the coordinator for the Focus Group on Indigenous Geographical Names (canadaungegn-canadagenung@nrcan-rncan.gc.ca) and the UNGEGN Secretariat.

Working Group and Focus Group members are now working to maintain this momentum and continue progress on geographical names as cultural heritage, including Indigenous geographical names, for the next UNGEGN session in 2027.

SPECIAL PROJECTS AND NEWS ITEM

Implementation of the UNGEGN Strategic Plan and Programme of Work 2021-2029: Insights from the 2025 session

Sungjae CHOO
Vice-Chair, UNGEGN
Republic of Korea
Email: sjchoo@khu.ac.kr

The 2025 session of the United Nations Group of Experts on Geographical Names (UNGEGN) welcomed a report on the progress and achievements of UNGEGN Strategic Plan and Programme of Work 2021-2029 (SP&PoW) for the years 2023-2025. Detailed in document GEGN.2/2025/161/CRP.161 and its appendices Add.1 and Add.2, the report highlighted several action items, out of 51, that demonstrated progress under each strategy, alongside challenges faced during implementation.

Progress achieved during the 2025 session

Significant advancements were made during the 2025 session, guided by the decisions adopted. Key action items and CRPs are outlined as follows:

Strategy 1. Technical expertise

- The concept for a universal unique identifier for cities within the World Geographical Names Database was adopted. The Secretariat was tasked with collaborating with Member States and UN entities to integrate authoritative geographical and toponymic data at national and city levels (1-i-5, CRP.21, Decision 2025/8).

Strategy 2. Relationships, links and connections

- A report on the collaborative project between UNGEGN and UN-GGIM is to be presented to the Committee of Experts in August 2025, informed by questionnaire analyses and other initiatives (2-i-2, CRP.143, Decision 2025/12).
- The Task Team for Africa will be revitalized in cooperation with the Economic Commission for Africa, adopting a systematic approach to support African Member States by identifying and prioritizing needs (2-i-3, Decision 2025/6).
- An expert from the Livonian Institute at the University of Latvia was appointed as a liaison officer to the Global Task Force for Making a Decade of Action for Indigenous Languages, the Permanent Forum on Indigenous Issues and the Group of Experts. (2-ii-6, 2-iii-8, Special Presentation 1, Decision 2025/16)

Strategy 3. Effective work programmes

- A legal opinion from the UN Office of Legal Affairs was requested regarding the status of "decisions" and "resolutions" as outcomes of the Group of Experts (3-ii-4, CRP.4, Decision 2025/2).
- The Working Group on Geographical Names as Cultural Heritage was revitalized, appointing Co-Convenors from Australia and Iceland and updating the WG's terms of reference and work plan (3-iii-7, CRP.31 and Add.1, Decision 2025/15).
- An analysis of the current strategic plan and programme of work will be initiated to inform the development of the next strategic plan to be discussed at the 2027 session (3-iii-10, CRP.161, Decision 2025/3).

Strategy 4. Culture, heritage and language recognition

- Inclusivity and culturally-informed approaches, as well as evidence-based decision making, will each be addressed in reports at future sessions (4-ii-7, Decision 2025/1).
- The Working Group on Geographical Names as Cultural Heritage initiated a global survey to gain a deeper understanding of the interest and interlinkages within Member States regarding

geographical names as cultural heritage (4-i-2, CRP.31 and Add.2, Decision 2025/15).

- The Working Group on Geographical Names as Cultural Heritage was tasked with building relationships with the Permanent Forum on Indigenous Issues and UNESCO (4-i-1, 4-ii-8, 2-i-4, CRP.94, Decision 2025/16).
- The Working Group on Geographical Names as Cultural Heritage organized two side events during the 2025 session, titled "Forum on Geographical Names as Cultural Heritage."

Strategy 5. Promotion and capacity building

- A panel discussion and special presentation were held during the 2025 session: "Toponymic Training 2030: Building Capacity for Tomorrow's Geographical Names Challenges" and "The e-learning course on the Norwegian Place Name Act and work on the standardization of geographical names" (5-iii-5, Decision 2025/19).
- The establishment of a trust fund was recommended to enhance training frameworks and address funding challenges (5-iv-7, CRP.14, Decision 2025/20).

It should be noted that meaningful progress was observed in addressing the challenges pointed out in the report (CRP.161); establishing a framework to support the standardization of geographical names in African countries and re-establishing the framework for Strategy 4, focusing on cultural, heritage, and linguistic aspects. Continued progress is anticipated during the inter-sessional period, with updates to be reported at the 2027 session.

Report on the Panel discussion "Advancing geographical names standardization through inclusive, culturally-informed and evidence-based solutions to support sustainable development"

28 April 2025, 5:10-6:10 p.m.
2025 session, UNGEGN, New York

Sungjae CHOO
Vice-Chair, UNGEGN
Republic of Korea
Email: sjchoo@khu.ac.kr
and
Allison Dollimore
PCGN, U.K.
United Kingdom of Great Britain and Northern Ireland
Email: adollimore@pcgn.org.uk

Introduction

The theme of the 2025 UNGEGN session, as reflected in the title of this panel discussion, was adopted from the ECOSOC theme for 2025, "Advancing sustainable, inclusive, science- and evidence-based solutions for the 2030 Agenda for Sustainable Development and its Sustainable Development Goals for leaving no one behind." The theme was also aligned with the UNGEGN Strategic Plan and Programme of Work 2021-2029, Strategy 1 (Technical expertise) and Strategy 4 (Culture, heritage and language recognition).

Building on this, the panel explored three key principles that had been central to UNGEGN's work for the last few decades:

- Inclusivity – ensuring that geographical names represent and respect all communities.
- Culturally-informed approaches – recognizing the historical, linguistic and societal significance of names.
- Evidence-based decision making – using research, data and best practices to guide standardization.

The session invited documents and subsequent discussions focusing on these three key principles. In this panel discussion, it was suggested to explore them in a different way, attempting to derive and elaborate the essence of each of these three principles by sharing experiences of Member States, international bodies, experts and private companies. The approach thus would be rather inductive, targeting a better understanding of inclusive, culturally-informed and evidence-based solutions to advance geographical names standardization and then further apply them to practices of geographical names management.

Given this background, the following thematic issues and questions were raised:

1. Inclusivity in geographical names standardization
    - What are inclusive solutions in standardizing geographical names?
    - How do standardized names affect marginalized or minority communities?
    - How does your organization ensure inclusive representation in geographical naming?

2. Culturally-informed approaches
    - Should geographical naming always be culturally-informed?
    - How do you distinguish culturally-informed geographical names?
    - Are culturally significant names always inherited from the past?
        - How does your institute handle names where different cultures overlap?

3. Evidence-based standardization
    - Should geographical naming always be evidence-based?
    - What types of evidence are considered in decision-making?
        - How does your institution verify and validate names?

4. Supporting sustainable development
    - How do these solutions support sustainable development?
    - Can you share examples where geographical names have promoted social cohesion or sustainability?
    - What policies contribute to the sustainability in geographical names standardization?

Discussion

Frédéric Giraut, Chair of the UNESCO Toponymy Inclusive Team, spoke about promoting minority and vernacular place names through what he called "participatory diagnostics" — gathering names directly from communities. He pointed out that while digital platforms offer new opportunities, they also bring risks, and that supplementing official name lists with community-based inventories is vital.

Dieter Laskowski, Technical Lead of Google Maps, spoke about the challenges Google faced balancing a good user experience with making space for inclusive naming. He explained how Google used a wide range of sources — from official data to direct user feedback — and how minority and historic names were preserved inside their database, even if they're not always immediately visible.

Emily Diana Lethbridge, Co-convenor of the UNGEGN Working Group on Geographical Names as Cultural Heritage, gave us a picture of how social and land use changes in Iceland were affecting geographical names, and how important it was to find a balance between protecting cultural heritage and embracing new naming trends. She also highlighted how collaboration between mapping authorities, the names committee, and academia was central to maintaining an evidence-based approach.

Ana Cristina da Rocha Berenger Resende, Geographic Names Manager of the IBGE, Brazil, shared how Brazil's IBGE was working to safeguard traditional and minority place names — not just by collecting names, but also the stories behind them. She talked about harmonizing names across maps of different scales, building tools for others to collect names, and raising awareness of names as part of cultural heritage.

Conclusion

The panel discussion offered a valuable platform for sharing the experiences of UNESCO, Google, Iceland and Brazil in the realm of geographical names management, with a focus on inclusivity and cultural-informed approaches. Participants reaffirmed the importance of three key orientations: inclusivity, culturally-informed strategies, and evidence-based decision-making. These themes were seen as vital for ongoing discussions and implementations, despite facing certain challenges and barriers. The consensus was that adhering to these orientations would ultimately contribute to sustainable development. This significant discussion has been officially documented in Decision 4/2025/1 of the session.

Additional resources on this panel discussion are available on the UNGEGN website.

(From the left) Sungjae Choo, Frédéric Giraut, Dieter Laskowski, Emily Diana Lethbridge, Ana Cristina da Rocha Berenger Resende, and Allison Dollimore © UN Web TV

In memoriam Ferjan Ormeling (1942-2025)

Jasper Hogerwerf
The Netherlands' Cadastre, Land Registry and Mapping Agency
Netherlands (Kingdom of the)
Email: Jasper.Hogerwerf@kadaster.nl

On 13 June Ferjan Ormeling passed away, at the age of 82. One of the most prominent members of the UNGEGN community of the past decades. An authoritative cartographer, a worldwide renowned scientist and a champion of geographical names standardization, but most of all a good friend and valued colleague.

Youth and studies

Ferdinand Jan Ormeling, born on 20 November 1942 in Utrecht, the Netherlands, was the son of two geographers, Ina ten Hoopen and Ferjan Ormeling Sr. His father would become a well-known professor of cartography and a leading member of both the International Cartographic Association (ICA) and UNGEGN. In many ways, Ferjan followed in his father's footsteps, beginning as early as 1949, when he and his family travelled to Indonesia to join Ormeling Sr. While his father set up the Geographical Institute of the Topographic Survey and became its director, Ferjan's parents divorced and he came to live with his mother and stepfather, first in Jakarta and then in Bandung. This period sparked a lifelong interest in Indonesia, where he remained a very welcome and highly esteemed guest.

In 1955, the family moved back to the Netherlands. After finishing high school, where Ferjan was taught Classics by the well-known Dutch poet Ida Gerhardt, he studied geography at the University of Groningen and attended lectures by the famous Dutch writer Willem Frederik Hermans. Ferjan, however, was only moderately impressed with Hermans' teaching. During his student years in Groningen, he worked on the editorial staff of De Bosatlas, the leading Dutch school atlas, of which his father had become editor-in-chief.

Early university work

Ferjan took Arabic as a minor and went to Utrecht for an additional minor in cartography under Professor Cornelis Koeman, who became a great source of inspiration. After graduating in 1969, he started working at Utrecht University. There, he gave lectures, edited atlases made by students, wrote lecture volumes and articles, and prepared exhibitions and conferences. He actively participated in the Royal Netherlands Geographical Society (KNAG)'s Cartographic Section, which became the Dutch Cartographic Association (NVK) in 1974, as well as the International Cartographic Association (ICA). He helped organize the International Cartographic Conference (ICC) in Amsterdam in 1967 and was involved in nearly every subsequent national and international cartographic conference over the following decades.

PhD research

In the 1970s, Ferjan became familiar with geographical names issues, possibly due to his father's involvement with UNGEGN. He devoted his PhD research to the representation of toponyms in linguistic minority languages on topographic maps, working under the supervision of Professor Koeman and Professor Dick Blok, an onomastician and then-director of the Meertens Institute. When Ferjan defended his thesis in 1983, Blok was serving as UNGEGN Chair. In his dissertation *Minority Toponyms on Maps: The Rendering of Linguistic Minority Toponyms on Topographic Maps of Western Europe,* Ferjan evaluated the rendering of toponyms in minority languages on maps compared with the resolutions of the United Nations on the standardization of geographical names and with the practical needs for topographic maps. He found that minority toponyms were often adapted to the majority language and thus differed from the locally used names. He concluded that acceptance of the monolingual rendering of toponyms in the minority language would be preferable.

Today, such an idea would probably raise little debate, but at the time, it was considered revolutionary and highly controversial within the UNGEGN community, where the dogma prevailed that standardization should take place in the national language and that every geographical feature should have only one official name. It led to an inevitable confrontation when Ferjan presented a paper based on his PhD thesis during the UNGEGN Session in Geneva in 1986, with some well-established delegates claiming that the acceptance of minority toponyms within national names standardization would "put the whole building of national standardization at risk". In retrospect, however, Ferjan was a pioneer in addressing the importance of respecting cultural and linguistic diversity in the standardization of

geographical names, recognising minority-language toponyms as an essential part of a country's cultural heritage - principles that UNGEGN would later go on to adopt in several United Nations resolutions.

Professor Ormeling

After Professor Koeman left Utrecht University in 1981, Ferjan took over as head of the Cartography section. Four years later, he was appointed full professor of Cartography. His inaugural lecture made national headlines because of his criticism of the poor quality – or complete absence - of mapping material in governmental policy reports, which, he argued, could lead to poor decision-making. He demonstrated how the maps could be improved and advocated for following the principles of Semiology of Graphics, as outlined by the French cartographer Jacques Bertin. He also identified early on the possibilities that digital techniques would offer to improve the field of cartography.

In addition to visual communication and toponymy, Ferjan's main research fields during his professorship included atlas cartography, education, data quality, cartographic infrastructure, and visualization as well as historical cartography.

Teaching was, without doubt, Ferjan's greatest passion and achievement. He taught the basics in cartography and GIS to generations of geography students in Utrecht, though it was only those who chose cartography as their doctoral specialization who came to know him better. As a passionate teacher combining thorough knowledge with humor, he was very popular among his cartography students. When cartography was discontinued as a specialization of the geography program in 2003, he became one of the founders and lecturers of the joint Master's program in *Geographical Information Management and Applications* (GIMA) of the universities of Delft, Utrecht, Wageningen, and Enschede (ITC).

International representation

Education was also an important focus of Ferjan's international activities. Within the ICA, he became a member of the Commission on Education and Training in 1980, and went on to co-chair it from 1987 to 1999. During this period, the Basic Cartography textbooks were published, which he co-authored with Roger Anson. He subsequently served as Secretary-General and Treasurer of ICA from 1999 to 2007, tirelessly promoting international collaboration on cartographic research and education.

Within UNGEGN Ferjan could fully utilize his teaching qualities at an international level thanks to a series of training courses in toponymy, initiated by his father as convenor of the Working Group on Training Courses in Toponymy. He gave his first lectures in 1984 during the second UNGEGN pilot course in Rabat, Morocco, and took part in the UNGEGN training course in Cipanas, Indonesia, in 1989. Ferjan was convenor of the Working Group on Training Courses in Toponymy from 1991 to 2017, making him the longest-serving officer in UNGEGN's history. Over the years, he was involved in more than 20 toponymy training courses for UNGEGN, the last of which took place in the Philippines in 2018.

Additionally, Ferjan served as Vice-Chair of UNGEGN from 2009 to 2017 and was particularly proud to have served as President of the United Nations Conference on the Standardization of Geographical Names (UNCSGN) in 2017. He contributed to several key UNGEGN publications, including the *Manual for the National Standardization of Geographical Names* in 2006 and the *Toponymy Training Manual* in 2017.

In representing the Netherlands in UNGEGN for almost 40 years, his focus and activities went beyond toponymic education. He worked on several updates of the Toponymic Guidelines of the Netherlands, supported the recognition and rendering of Frisian minority names, addressed the standardization and use of exonyms, and promoted the development and availability of toponymic databases.

National representation

In the Netherlands, Ferjan played a leading role in the Dutch Cartographic Association (NVK), being a board member from 1975, vice-chair from 1983 to 1995, and chair from 1995 to 1997. He became an honorary member in 2001, just before the association merged in 2003 with geodetic, land surveying, and other associations to form Geo Information Netherlands (GIN). Additionally, Ferjan was editor of the *Kartografisch Tijdschrift* ('Cartographic Magazine') from 1975 to 2003 (as editor-in-chief from 1983 to 1995) and of the subsequent *Geo Info* magazine from 2003 to 2020. He was also an editor of the historical cartographic journal *Caert Thresoor*.

Although the Netherlands, to Ferjan's great regret and frustration, has no national geographical names authority, there is a supranational Commission on Geographical Names (CAN) of the Dutch Language Union (*Nederlandse Taalunie*). Ferjan served as a member of its predecessor, the Commission on Exonyms, from 1988 and chaired the

commission from 1999 to 2018. In this role, he contributed to the compilation and maintenance of authoritative lists of country names and exonyms in the Dutch language. He was also a member - and later chair - of the Advisory Commission on Geographical Names in the Netherlands, primarily advising on new names for merged municipalities, until the CAN incorporated this commission in 2023.

Publications and recognition

Ferjan wrote and edited countless books, reports, articles, and other publications. Important releases include the co-edited scientific and thematic *Atlas of the Netherlands* (1984-1991, 20 volumes) and the widely translated textbook *Cartography: Visualization of geospatial data* (first Dutch edition in 1985, most recent English edition in 2020), co-authored with Professor Menno-Jan Kraak, that for decades has been teaching people around the world the basics of cartography. In his later years, his focus shifted towards historical cartography, with a specialization in the colonial cartography of the Netherlands, the history of atlas cartography, and historical toponymy. As a member of the Explokart research group, housed at the University of Amsterdam since 2013, he co-authored the Comprehensive Atlas of the Netherlands East Indies in 2003, the Comprehensive Atlas of the Dutch United East India Company in 2006, and the *Biography of the Bosatlas* (in Dutch) in 2005. Shortly before his death, Ferjan submitted the manuscript of his book *Mapping the Netherlands East Indies: A Cartographic Appraisal*, which will be published posthumously.

It is hardly a surprise that his outstanding scientific work has been widely recognized and honored, with the Carl Mannerfelt Gold Medal of the ICA in 2009, the Plancius Medal of the Royal Netherlands Geographical Society in 2017, the Mercator Medal of the German Cartographic Society in 2022, as well as with honorary doctorates from Eötvös Loránd University in Budapest in 2013 and Aristotle University in Thessaloniki in 2015.

Appreciation

Indeed, Ferjan was not only an exceptionally skilled scientist but also a gifted teacher who excelled in transferring his knowledge and ideas to others. He organized and led meetings and conferences smoothly and respectfully, while also voicing strong opinions with firm positions when necessary. His often modest and dignified demeanor contributed to his impressive appearance.

But most of all, Ferjan was a great colleague, a good friend, and, for his family, who affectionately called him Flip, a loving and caring father, brother, and partner. He was always genuinely interested in how people were doing and in the progress of their work, happy to help and offer guidance on issues, knowing how to inspire with his own knowledge and experience.

Ferjan educated many people around the world, and almost everyone currently involved in cartography and toponymy in the Netherlands has, in some way, learned from him. This, together with his scientific work, is part of his vast and lasting legacy. We can only look back with great admiration and gratitude at what he has meant to the fields of cartography and toponymy and, to some of us, personally.