

Masó, J. & Pons, X. (2026). Can the discrete global grid systems and global vocabularies contribute to the Digital Earth? / ¿Pueden los sistemas de malla global discreta y los vocabularios globales participar en la creación de una Tierra Digital? *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (Editorial), 37, I-V. <http://dx.doi.org/10.21138/GF.1000>

CAN THE DISCRETE GLOBAL GRID SYSTEMS AND GLOBAL VOCABULARIES CONTRIBUTE TO THE DIGITAL EARTH?

PUEDEN LOS SISTEMAS DE MALLA GLOBAL DISCRETA Y LOS VOCABULARIOS GLOBALES CONTRIBUIR A LA CREACIÓN DE UNA TIERRA DIGITAL?

^{1a}JOAN MASÓ , ^{2,1b}XAVIER PONS 

^{1,2}Grumets Research Group

¹CREAF. Universitat Autònoma de Barcelona. 08193, Bellaterra, Catalonia, Spain

²Departament de Geografia, Edifici B. Universitat Autònoma de Barcelona. 08193, Bellaterra, Catalonia, Spain

^ajoan.maso@ieee.org ^bxavier.pons@uab.cat

We are approaching the 29th anniversary of the visionary speech delivered by former U.S. Vice President Al Gore at the California Science Center in Los Angeles on 31-January-1998. He envisioned a multi-resolution, three-dimensional virtual globe that would allow citizens to seamlessly navigate spatial and temporal data about the planet. The dream has been realized only to a limited extent, through applications such as Google Maps or Microsoft Flight Simulator. After all this time, for GIS professionals the “Digital Earth” sounds as a natural concept to aspire to, while at the same time remaining somewhat science fiction in terms of a full representation of the physical, social and ecological variables. This Editorial does not aim to review all the reasons that have hindered the realization of the Digital Earth concept; instead, it focuses on a couple of specific approaches that may contribute to the solution. For a more comprehensive review of the Digital Earth implementation, see Annoni (2023).

For more than five centuries, a myriad of cartographic projections have been used to represent the Earth on a flat paper surface. As countries developed their own maps, they also created independent projections and reference systems. The reasons for this range from practical considerations (such as the need of equal-area *versus* conformal projections) to more technical factors (including topographic perspectives based on instruments such as the plumb line and the spirit level, reflecting the fact that early surveying and mapping were fundamentally local operations tied to local gravity fields). With the advent of the digital era, instead of removing the old projections, the spherical Mercator was revived and adapted to what is known as the “Popular Visualization Pseudo-Mercator” (although the US NGA (2017) officially discourages its usage beyond visualization purposes, and does not allow mapping beyond 85.051129° north and south). And latitude/longitude coordinates have proliferated, while some countries have abandoned traditional national datums in favor of global ones. One may also consider the vast and ever-growing range of formats available for representing raster and vector data. This situation may evolve with the explosion of Large Language Models; however, at present, only an elite of GIS experts can effectively navigate coordinate reference systems and geospatial formats.

At the same time, the amount of geospatially referenced data is continuously increasing (Remote Sensing imagery and products, automated in-situ sensor data, etc), challenging the capabilities of the average modern computer. This growing volume of data should bring us closer than ever to realization of the Digital Earth, if only we had the capacity to organize it in an agile and efficient manner. Established techniques for geospatial data storage and analysis may be reaching their limits and even acting as impediments rather than facilitators in harnessing the full potential of the Digital Earth (Gibb

2022). We need to rethink how we store, manage, share, and work with geospatial data. It is therefore not surprising that there is considerable interest in the cloud-optimized formats, which partition data into geospatial tiles or in arbitrary chunks, allowing access to subsets of data via web requests, as described in Maso and Pons (2022). While these efforts support a more distributed and flexible data access, they remain only a partial solution.

Perhaps, a part of the solution may lie in the adoption of a new model that considers the Earth system as a single entity, described through a global multi-resolution framework that predefines a discrete set of entities in a structured and regular manner. Discrete Global Grid Systems (DGGS) partition the entire planet into a hierarchy set of progressively finer resolution zones (or cells). In this approach, traditional coordinate reference systems and 2D coordinates are replaced by unique spatio-temporal identifiers assigned to each zone. Conventional raster square cells and vector geometric objects are replaced by conterminous polygonal shapes (*e.g.*, combinations of pentagons and hexagons) that define the extent of each zone. Each zone is associated with a set of attributes representing physical or socioeconomic variables of the Earth system. As zones identifiers are attached to particular resolution levels (that are hierarchically related to the other lower resolutions), the associated attributes inherently reflect a given level of positional uncertainty. Data overlay within a DGGS becomes straightforward, as zones sharing the same spatiotemporal identifier correspond to the same place and can be directly combined.

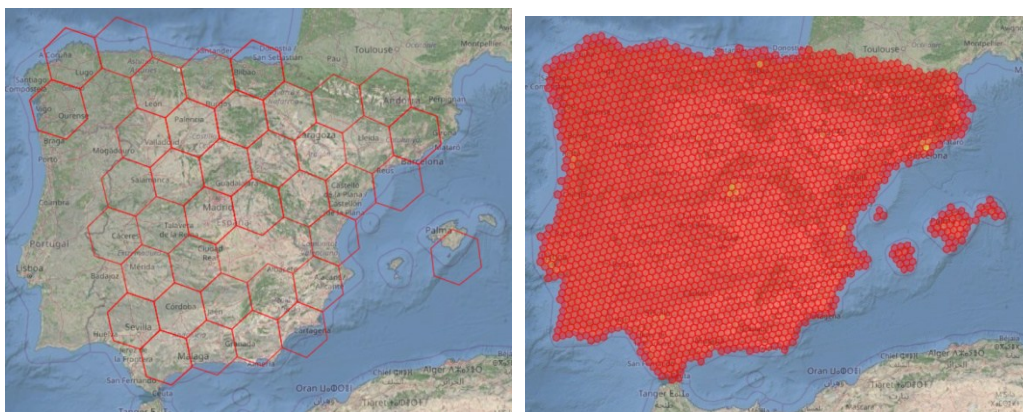


Figure: Left: Capitals of the Spanish provinces as DGGS Uber H3 at level 3. Right: Population on the Iberian Peninsula as DGGS Uber H3 at level 5 (created with TAPIS <https://github.com/grumets/tapis>. Raw data from <https://www.ign.es/web/ane-datos-geograficos> and https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/GHSL/GHS_POP_GLOBE_R2023A)

From an analytical perspective, an advantage of a global multi-resolution system is the ability to readily apply common operations such as grouping and joining tables, as well as to combine data through recursive algebraic and logical operations across all zones. However, operations of convolution remain non-trivial if the software does not recognize the topological relationships between neighboring zones or the hierarchical relationships across different resolution levels (Hojati 2022). Furthermore, plate tectonics continue to pose additional challenges. With these difficulties still to be overcome, it is not surprising that current research is focused on enhancing the spatial analytics of DGGS, and providing more geospatial functionality to the emerging systems using DGGS. In doing so, however, research focuses on the geometry of the zones while overlooking the thematic component.

In order to realize the dream of the Digital Earth at the level of precision now emerging, it is neither feasible to rely on centralized data storage nor to depend on a single institution taking care of all required datasets. The only viable approach is to enable decentralized access to dataset fragments in a seamless way. While the adoption of a DGGS allows for immediate combination of data within a common spatial framework, it also raises a fundamental question: what is the semantic meaning of the data being combined? In a fully distributed DGGS-based environment, it is therefore equally important to characterize the attributes associated with each zone by clearly defining the variables and properties they represent.

Ideally, these variables should be linked to globally resolvable identifiers within controlled vocabularies, such as QUDT URIs (Ray *et al.* 2011). By referencing shared concepts rather than relying

solely on textual descriptions—which are often ambiguous, especially for non-experts—it becomes clearer whether datasets produced in different parts of the world contribute to measuring or estimating the same phenomena, and how these concepts relate to one another. These conceptual definitions should be complemented by units of measurement, also expressed through globally identifiable vocabularies.

To date, Geofocus has not published contributions related to DGGS. With this Editorial the authors would like to encourage readers to experiment with DGGS and to elaborate contributions on this emerging topic.

Nos acercamos al 29º aniversario del visionario discurso pronunciado por el exvicepresidente estadounidense Al Gore en el Centro de Ciencias de California en Los Ángeles el 31 de enero de 1998. En él, imaginó un globo virtual tridimensional y multirresolución que permitiría a los ciudadanos navegar sin problemas por datos espaciales y temporales del planeta. Este sueño se ha materializado solo parcialmente, mediante aplicaciones como Google Maps o Microsoft Flight Simulator. Después de todo este tiempo, para los profesionales de los SIG, la «Tierra Digital» suena como un concepto natural al que aspirar, aunque sigue siendo en cierto modo ciencia ficción en lo que respecta a una representación completa de las variables físicas, sociales y ecológicas. Este Editorial no pretende revisar todas las razones que han obstaculizado la realización del concepto de Tierra Digital; en cambio, se centra en un par de enfoques específicos que pueden contribuir a la solución. Para una revisión más completa de la implementación de la Tierra Digital, véase Annoni (2023).

Durante más de cinco siglos, se han utilizado innumerables proyecciones cartográficas para representar la Tierra en una superficie plana de papel. A medida que los países desarrollaban sus propios mapas, también creaban proyecciones y sistemas de referencia independientes. Las razones para esto van desde consideraciones prácticas (como la necesidad de proyecciones equivalentes (de área comparable) frente a proyecciones conformes) hasta factores más técnicos (incluidas las perspectivas topográficas basadas en instrumentos como la plomada y el nivel de burbuja, lo que refleja el hecho de que los primeros levantamientos topográficos y cartografiados eran fundamentalmente operaciones locales vinculadas a campos gravitatorios locales). Con la llegada de la era digital, en lugar de eliminar las antiguas proyecciones, la Mercator esférica fue revivida y adaptada a lo que se conoce como el "Pseudo-Mercator de Visualización Popular" (aunque la NGA de EE.UU. (2017) desaconseja oficialmente su uso más allá de fines de visualización y no permite la cartografía más allá de 85,051129º norte y sur). Y las coordenadas de latitud/longitud se han multiplicado, mientras que algunos países han abandonado los dátums nacionales tradicionales en favor de los globales. También se puede considerar la vasta y cada vez mayor gama de formatos disponibles para representar datos ráster y vectoriales. Esta situación puede evolucionar con la explosión de los Modelos Extensos de Lenguaje (LLM); sin embargo, en la actualidad, solo una élite de expertos en SIG puede manejar eficazmente los sistemas de referencia de coordenadas y los formatos geoespaciales.

Al mismo tiempo, la cantidad de datos georreferenciados aumenta continuamente (imágenes y productos de teledetección, datos de sensores *in-situ* automatizados, etc), lo que supone un reto para las capacidades de los ordenadores modernos. Este creciente volumen de datos debería acercarnos más que nunca a la realización de la Tierra Digital, si tan solo tuviéramos la capacidad de organizarlo de forma ágil y eficiente. Las técnicas establecidas para el almacenamiento y análisis de datos geoespaciales podrían estar llegando a sus límites e incluso actuando como obstáculos en lugar de facilitadores para aprovechar todo el potencial de la Tierra Digital (Gibb 2022). Necesitamos replantearnos cómo almacenamos, gestionamos, compartimos y trabajamos con datos geoespaciales. Por lo tanto, no sorprende el considerable interés que despiertan los formatos optimizados para la nube, que dividen los datos en teselas geoespaciales o en fragmentos arbitrarios, permitiendo el acceso a subconjuntos de datos mediante solicitudes web, como se describe en Maso y Pons (2022). Si bien estos esfuerzos favorecen un acceso a los datos más distribuido y flexible, siguen siendo solo una solución parcial. Quizás, parte de la solución radique en la adopción de un nuevo modelo que considere el sistema terrestre como una entidad única, descrita mediante un marco global multirresolución que predefine un conjunto discreto de entidades de manera estructurada y regular. Los Sistemas de Malla Global Discreta (DGGS) dividen

todo el planeta en una jerarquía de zonas (o celdas) de resolución progresivamente más fina. En este enfoque, los sistemas de referencia de coordenadas tradicionales y las coordenadas 2D se reemplazan por identificadores espaciotemporales únicos asignados a cada zona. Las celdas cuadradas ráster convencionales y los objetos geométricos vectoriales se reemplazan por formas poligonales contiguas (por ejemplo, combinaciones de pentágonos y hexágonos) que definen la extensión de cada zona. Cada zona se asocia con un conjunto de atributos que representan variables físicas o socioeconómicas del sistema terrestre. Dado que los identificadores de zona se adjuntan a niveles de resolución particulares (que están relacionados jerárquicamente con las resoluciones inferiores), los atributos asociados reflejan inherentemente un determinado nivel de incertidumbre posicional. La superposición de datos dentro de un DGGS resulta sencilla, ya que las zonas que comparten el mismo identificador espaciotemporal corresponden al mismo lugar y pueden combinarse directamente.

Desde una perspectiva analítica, una ventaja de un sistema global multirresolución es la capacidad de aplicar fácilmente operaciones comunes como agrupar y unir tablas, así como combinar datos mediante operaciones algebraicas y lógicas recursivas en todas las zonas. Sin embargo, las operaciones de convolución siguen siendo complejas si el software no reconoce las relaciones topológicas entre zonas vecinas o las relaciones jerárquicas entre diferentes niveles de resolución (Hojati 2022). Además, la tectónica de placas continúa planteando desafíos adicionales. Ante estas dificultades aún por superar, no sorprende que la investigación actual se centre en mejorar el análisis espacial de los DGGS y en proporcionar mayor funcionalidad geoespacial a los sistemas emergentes que los utilizan. Al hacerlo, sin embargo, la investigación se centra en la geometría de las zonas, pasando por alto el componente temático.

Para hacer realidad el sueño de la Tierra Digital con el nivel de precisión que está surgiendo, no es viable depender del almacenamiento centralizado de datos ni de que una sola institución se encargue de todos los conjuntos de datos necesarios. El único enfoque viable es permitir el acceso descentralizado a los fragmentos de los conjuntos de datos de forma fluida. Si bien la adopción de un DGGS permite la combinación inmediata de datos dentro de un marco espacial común, también plantea una cuestión fundamental: ¿cuál es el significado semántico de los datos que se combinan? En un entorno totalmente distribuido basado en DGGS, resulta igualmente importante caracterizar los atributos asociados a cada zona definiendo claramente las variables y propiedades que representan.

Idealmente, estas variables deberían estar vinculadas a identificadores globalmente resolubles dentro de vocabularios controlados, como las URI QUDT (Ray *et al.* 2011). Al hacer referencia a conceptos compartidos en lugar de depender únicamente de descripciones textuales —que a menudo son ambiguas, especialmente para quienes no son expertos—, resulta más claro si los conjuntos de datos producidos en diferentes partes del mundo contribuyen a medir o estimar los mismos fenómenos, y cómo se relacionan estos conceptos entre sí. Estas definiciones conceptuales deberían complementarse con unidades de medida, también expresadas mediante vocabularios globalmente identificables.

Hasta la fecha, Geofocus no ha publicado contribuciones relacionadas con los DGGS. Con este Editorial, los autores desean animar a los lectores a experimentar con los DGGS y a elaborar contribuciones sobre este tema emergente.

References

- Alessandro Annoni, Stefano Nativi, Arzu Çöltekin, Cheryl Desha, Eugene Eremchenko, Caroline M. Gevaert, Gregory Giuliani, Min Chen, Luis Perez-Mora, Joseph Strobl & Stephanie Tumamos (2023) Digital earth: yesterday, today, and tomorrow, *International Journal of Digital Earth*, 16:1, 1022-1072, DOI: [10.1080/17538947.2023.2187467](https://doi.org/10.1080/17538947.2023.2187467)
- Gibb, Robert G., Matthew B.J. Purss, Zoheir Sabeur, Peter Strobl, and Tengpeng Qu. 2022. "Global Reference Grids for Big Earth Data." *Big Earth Data* 6 (3): 251–55. DOI: [10.1080/20964471.2022.2113037](https://doi.org/10.1080/20964471.2022.2113037).
- Majid Hojati, Colin Robertson, Steven Roberts & Chiranjib Chaudhuri (2022) GIScience research challenges for realizing discrete global grid systems as a Digital Earth, *Big Earth Data*, 6:3, 358-379, DOI: [10.1080/20964471.2021.2012912](https://doi.org/10.1080/20964471.2021.2012912)

Masó, J., & Pons, X. (2022). After years convincing people about geospatial services, we are going back to files, but this time we call them "Cloud optimized" / Después de años convenciendo a la gente sobre los servicios geoespaciales, volvemos a los archivos, pero esta vez los llamamos "Optimizados para la nube", GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Editorial), nº 30, p. 1–4. <https://doi.org/10.21138/GF.813>

Ray, S., *et al.* (2011). QUDT: Quantities, Units, Dimensions and Types Ontology. Retrieved from QUDT.org.

US NGA (2017) Map Projections for GEOINT Content, Products, and Applications (2017-12-13). NGA standardization document NGA.SIG.0028_1.0_MAPPROJ. Retrieved from