BSGLg, 85, 2025, 1-12
DOI: 10.25518/0770-7576.7644

Accepté le 7 janvier 2026

BUILDING TRUSTWORTHY DIGITAL TWINS OF TERRITORIES: TRUTH, METADATA, AND RELEVANCE

CONCEPTION DE JUMEAUX NUMÉRIQUES TERRITORIAUX DE CONFIANCE : ENJEUX DE VÉRITÉ, DE MÉTADONNÉES ET DE PERTINENCE

Bénédicte BUCHER, Mohamed CHACHOUA

Abstract
Adapting the digital twin paradigm to territories raises high expectations from municipalities, who seek tools to guide sustainable transformations. Beyond data integration, Digital Twins of Territories (DTTs) must engage humans through geovisualization and participatory interfaces, to explore scenarios and support public debate. This paper examines how to enhance the trustworthiness of DTTs, by grounding them in theories of truth and communication. We emphasise the crucial role of metadata and user feedbacks in ensuring that DTTs enable their users to reason, communicate and justify their representations of territories, while clarifying their objectives. Case studies on air pollution and urban densification illustrate how these principles can be applied in practice

Keywords
Digital Twin of Territory, truth, relevance, quality, metadata, user feedback

Résumé
L'adaptation du paradigme des jumeaux numériques aux territoires cristallise de nombreuses attentes de la part des municipalités, qui cherchent des outils pour guider leurs transitions écologiques. Au-delà de l'intégration de données, les Jumeaux Numérique de Territoires (JNT) doivent impliquer les humains, notamment grâce à la géovisualisation et des interfaces participatives, afin de favoriser le développement, l'échange et la confrontation d'arguments sur les dynamiques territoriales. Cet article examine comment renforcer la fiabilité des JNT en les fondant sur des théories de la vérité et de la communication. Nous soulignons le rôle crucial des métadonnées et des commentaires des utilisateurs pour garantir que les DTT puissent permettre à leurs utilisateurs de raisonner, communiquer et justifier leurs représentations des territoires, tout en clarifiant leurs objectifs. Des études de cas sur la pollution atmosphérique et la densification urbaine illustrent comment ces principes peuvent être appliqués dans la pratique.

Mots-clés
Jumeau Numérique de Territoire, véracité, pertinence, qualité, métadonnées, retour utilisateur

INTRODUCTION

Managing contemporary territories in order to engage them in multiple and interdependent transitions and make them resilient to crises is highly challenging. Among other decisions, societies must adapt infrastructures to future climate conditions, develop housing differently to limit soil sealing or adapt water usage to resource scarcity. Decisions should anticipate dramatically different climatic scenarios in the next century as well as consider the complexity of our world and interdependencies between domains. Municipalities must engage citizens in deciding the trajectory they want for their territory, through public debates adapted to their gaps of expertise, and fruitful controversies that bypass polarized opinions.

In that context, the paradigm of the Digital Twin crystallizes expectations. A digital twin is a "digital representation of a physical, real-world system (element or process) continuously adapted based on an ongoing two-fold exchange of information" which supports what-if scenarios in predictive maintenance processes or by autonomous systems (Callcut *et al.*, 2021). It emerged from industrial design (Grieves and Vickers, 2017) and has spread to other manufacturing domains (Kritzinger *et al.*, 2018), and non-manufacturing domains like the healthcare field (Bruynseels *et al*., 2018). (Batty, 2018) argues that a city model is a kind of digital twin of a city, not an identical twin but a digital representation that moves closer to reality. How to get the city digital twin close to the real city, including not only built environment but also social and

economic processes, is "a collective exploration" for society (ibid). Qanazi *et al*. (2025) investigate the different forms of what they call Urban Digital Twins (UDT) and emphasize the need for improved management of data quality, ownership and security. UDT should link technical modeling with citizen participation and data governance (ibid). With the emergence of digital twins of geographical entities, like bridge, electric grid, cities, countries develop national programs on the one hand to support the development of digital twin for the different relevant entities, including territories, and on the other hand to support the exchange of information between different digital twins of located things, based on location (Ellul *et al.*, 2022).

We consider the Digital Twin of a Territory (DTT) as the transposition of the Digital Twin paradigm to a contemporary territory which needs to engage in multiple transitions and become resilient to crises. DTTs are tools for societies to make informed choices and democratically act. It extends the UDT concept to every territory where decisions are needed, including rural municipalities, protected areas, a country, Europe. A UDT is a specific case of DTT. Hence, we define a DTT as a process where a digital artefact is connected with a real territory and where a bidirectional connection exists between both to update and improve the model, and update and improve the reality, as illustrated in Figure 1.

The digital artefact can vary to adapt to the specific priorities, contexts, incl. funding and capacities, and strategic objectives of each territory. The connexion is not made up of real-time data streams only. In one direction, manual processes are often needed to update and improve the digital artefact, like establishing partnerships with more data providers or experts' collaboration to improve the concepts used to describe a phenomenon. In the other way, human interaction with the model is needed to update and improve actions in the real world to reach the expected impacts.

Different generic DTT sub-processes are identified based on the Urban Digital Twin literature as well as on empirical findings from focus groups organised in France by the French national mapping agency, IGN, to apply the DT paradigm to French territories. These are illustrated in Figure 2. A DTT should support achieving a shared diagnosis, of a territory, identifying priority targets, exploring possible actions for the future and their impacts, debating to refine and democratically adopt a scenario, communicating to facilitate adoption of new projects, and monitoring the project efficiency. Many functionalities call for enhanced representation of the territory. The representation of the territory can take different forms during the different sub-processes. We compare the DTT artefacts to maps that tell different stories about reality (Monmonier, 1991), or, as (Wood, 2010) puts it, are different "*arguments about the world*" subject to rhetorical choices and biases. There are different DTT stories reading or telling: users receiving direct messages from the DTT (text or visual), human elaborating stories through serendipitous interfaces, human or machine communication, human sharing opinions and theories in public debates. This comparison is motivated by the fact that DTT power of visualization and simulation can both illuminate and distort stories.

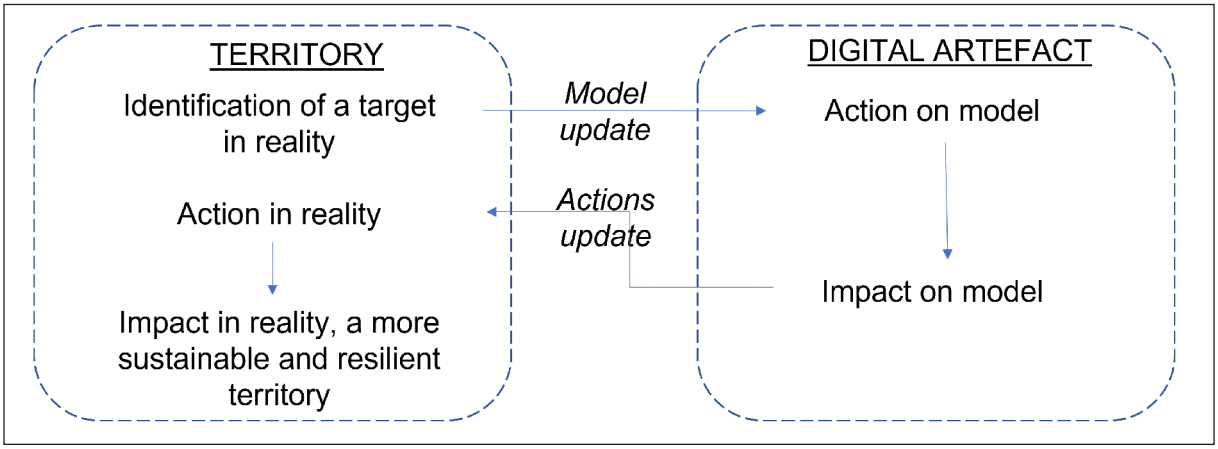

Figure 1. The digital twin of a territory is a process with a bidirectional connexion between a territory and a digital artefact that: (1) continuously updates and improve the model so that it better represents the targets, possible actions and their impacts (2) update and improve our actions so that their impacts meet the targets.

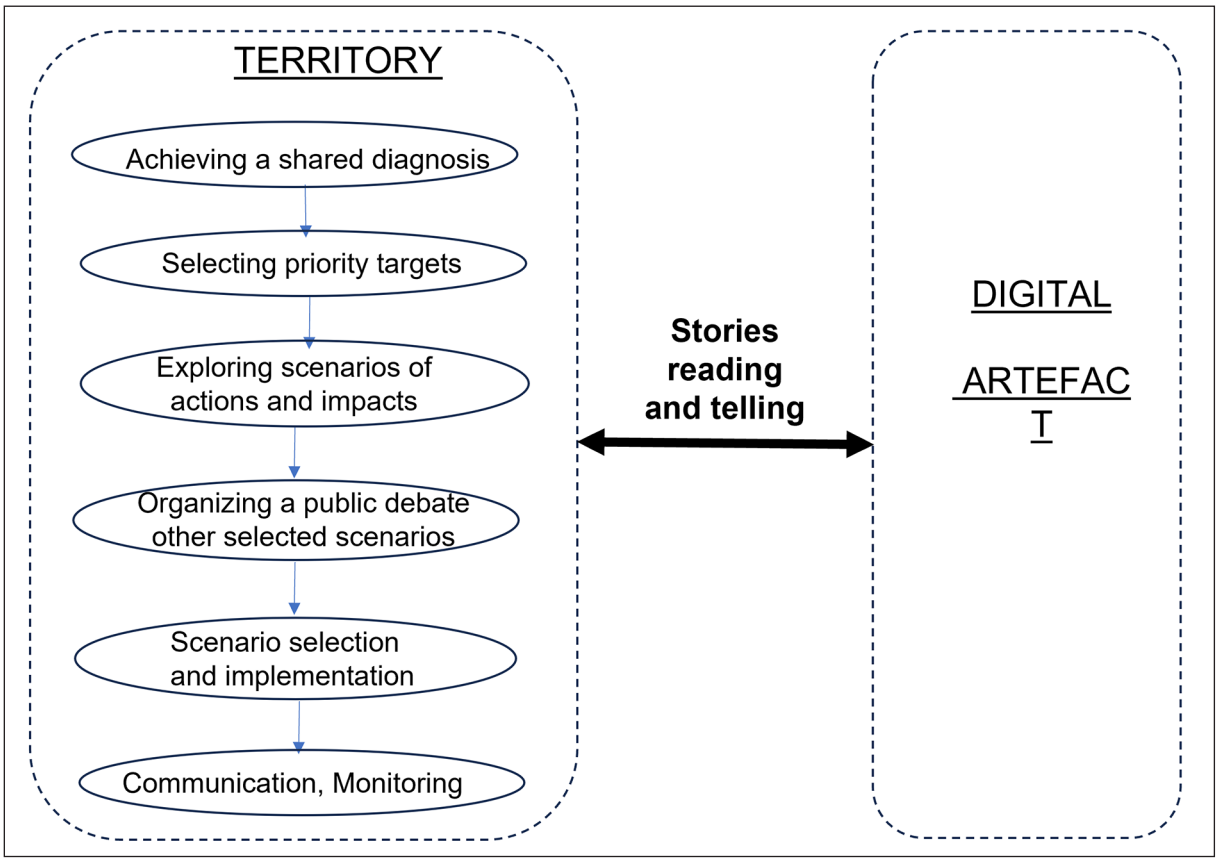

Figure 2. The DTT process to achieve ecological transition implies different sub-processes, many of which imply human interaction with the DTT artefact to read or share stories about the territory, about the diagnosis, targets, possible actions, and their impacts.

This article examines DTT's capacity to help its different users to tell and share trustworthy stories, needed for the processes of understanding a territory, engaging in open debates on scenarios for its future. It draws on contributions from information science (Shannon and Weaver, 1949), the epistemology of truth and factual narratives (Tarski, 1936; Floridi, 2011), and on geographic information science and cartography focused on the representation of space (Goodchild, 2020; Batty, 2018; Monmonier 1991).

Two research projects are used to illustrate this analysis. These projects do not cover the full DTT functionalities listed in Figure 2 and rather focus on the following ones where storytelling is crucial: achieving a shared diagnosis, identifying possible actions, and organizing public debates.

The research project POLLUSCOPE (Brahem *et al*., 2020; Mehanna, 2023) aims at improving the diagnosis of citizens' exposure to air pollution in Paris region and at identifying possible impacts on health and adaptation or mitigation actions, at the scale of each individual or at the scale of the territory. A pollution model is derived from a set of in situ and wearable sensor measurements of different pollutants and algorithms to process missing values and achieve a continuous model in space and in time (Brahem *et al*., 2020; Mehanna, 2023). The integration of mobility data and demographic data is used to reveal exposure peaks during daily commutes and the integration of building data to quantify indoor-outdoor contrasts. Health data are integrated to quantify the impact of air pollution.

The research project SUBDENSE aims at improving scientists' and practitioners' understanding of urban densification processes by describing buildings dynamics across different territories in France, Germany and the United Kingdom (Erhardt *et al.*, 2025). It uses building data surveyed at different times to detect changes and produce Building-Changes maps on the functional urban area of Strasbourg, Dortmund and Liverpool (Bucher *et al.,* 2025). Building change data requires less expertise to read than current dashboards and density maps

and has the potential to support more inclusive discussions on densification (ibid). The application is to continuously improve the digital material scientists and practitioners can use to elaborate trustworthy stories about urban transformation that produce more dwellings and their other impacts. SUBDENSE model is a sub-component for a DTT digital artefact to support a shared diagnosis of the territory with respect to densification, explore scenarios and debating.

I. TRUTH AND METADATA

A. Theoretical frameworks of truth applied to DTT stories

1. Adequacy

Alfred Tarski's theory of truth (1936, 1994) is a major philosophical foundation for research on truth applied to formalized language. According to it, the proposition "P", expressed in a formal language, is true if, and only if the model corresponds to reality. This adequacy, called T-convention, is to be assessed and processed in another language that Tarski calls the metalanguage.

In the context of a DTT, Tarski's Convention T applies when a ground truth can be established. In the POLLUSCOPE project, a proposition can be "The concentration of NO_2 on October 25, 2025, at 10:15 a.m. in Daniel's classroom is 85 µg/m³". The adequacy of this proposition to reality can be assessed if there can be ground truth, i.e. direct sensor measures in the classroom and at that time. Besides, there are always implicit spatial and temporal referencing frames. How precise is the date, and does "in Daniel's classroom" refer to every location inside the room or to locations where Daniel and the students breathe.

2. Consistency

Direct adequacy cannot be established when ground truth is lacking or when it is too expensive to evaluate the truth of all possible propositions in the model. For example, in the SUBDENSE project, an example of proposition in the model is "*there is specific BuildingChange of changetype construction in this location between 2011 and 2021"*. The adequacy to reality cannot be assessed directly as it involves the past.

In that case, internal consistency becomes a pragmatic indicator of validity. This was formalized by Tarski (1936) who defined truth as the satisfaction of statements within a model. Statements are used on the one hand, to generate more propositions based on true propositions, and on the other hand to detect possible anomalies when a consistency statement is violated.

In the POLLUSCOPE project, internal consistency can be evaluated through a set of statements like "if there is heavy traffic, then there is air pollution". These statements are used to detect anomalies in the model, such as the absence of pollution in an area with heavy traffic. Such anomalies can occur in models that rely on dynamic data exchanged with connected devices (IoT) or actuators, and that can be affected by sensor malfunctions, transmission errors, or environmental interferences (Zhou *et al*., 2015).

Internal consistency is also useful when a model relies on heterogeneous data which can generate artefacts. In the SUBDENSE project, BuildingChange data is computed based on the data at two timestamps, 2011 and 2021, as illustrated in Figure 3. In 2011, the scope of the french topographic building data included all buildings with area bigger than 25m². In 2013 it has evolved to include all buildings represented in cadastral data. Hence, a pragmatic rule has been introduced to express that some BuildingChanges may correspond to changes in the survey procedure, and not to changes in reality: "If a BuildingChange is of type construction and has an area smaller than 25m² then it is probably caused by a change in specification and not a change in reality".

3. Uncertainty and imprecision

The adequacy between a proposition and reality can be subject to uncertainty and imprecision. For example, in the POLLUSCOPE project, during a temporary outage of several sensors, data becomes missing or incomplete. The system cannot determine whether the air is polluted or not (i.e. True or False). As the mobility model uses GPS data from vehicles to estimate traffic congestion, it does not account for pedestrians, bicycles, or temporary events (e.g., protests or construction) that can impact traffic. As a result, it only reflects a partial view of reality. In SUBDENSE, the initial Building data that are used to detect change do not systematically have documented values for their

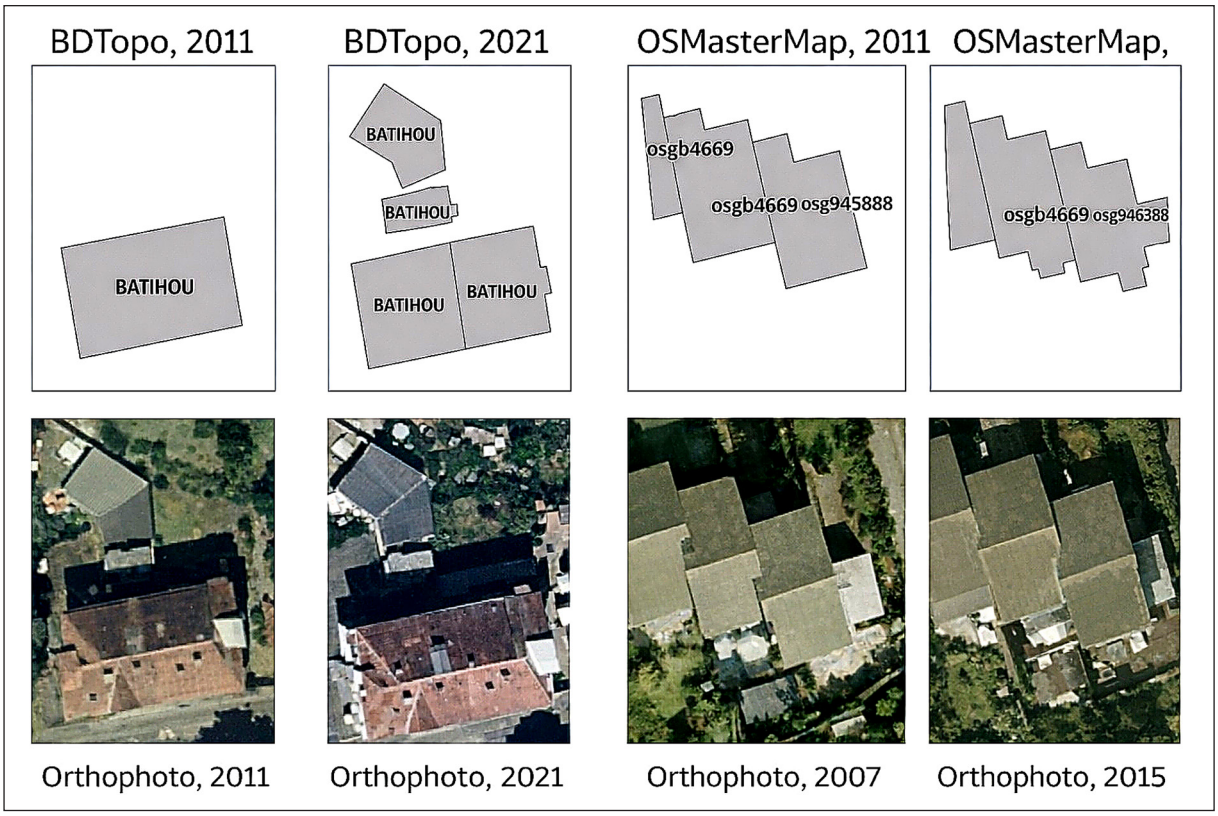

Figure 3. after (Bucher *et al*., 2025) Changes in building data can correspond to changes in specifications of BD-Topo® (left) or of OS MasterMap® (right). Source: © copyright IGN-2011, IGN-2021; © Crown copyright and database rights 2023 Ordnance Survey (100025252); © Getmapping Ltd

properties *height* and *buildingUse*. Consequently, the raw BuildingChange data derived from change detection are incomplete and do not capture all changes of height and changes of use.

More generally, models that describe territories often operate within informationally incomplete environments where different forms of imprecision and uncertainty must be considered to complement the truth value by adequacy or by consistency (Chachoua, 1998; Chachoua and Pacholczyk, 2006). At the level of the formalism in which data are structured, different models have been proposed for different situations of uncertainty and imprecision, such as belief function theory (Shafer, 1976; Smets and Kennes, 1994) or fuzzy logic (Zadeh, 1965; Dubois and Prade, 1988). There is no single theoretical framework sufficient in isolation to handle uncertainty and imprecision in real-world urban data (Fuller *et al*., 2020).

B. Metadata: a core asset for telling true stories on DTT

The previous section introduced Tarski's notion of metalanguage to manage and assess the truth of propositions elaborated in any formal model. In a DTT context, metadata is a useful asset to contribute to such metalanguage as summarized in Table 1. Metadata refers to any data, structured or not, about the data, like documentation, readme files, structured metadata, and possible log files. More generally, metadata are acknowledged to be an important element to assess what information can be inferred through data and adopt a critical perspective on data, by describing different categories of information about the data: the scope of data, modeling choices and provenance, modalities of observations, modalities of generation, algorithmic assumptions, scope of the model, modeling choices and provenance of the data, what is observed, what is generated, default parameters, data gaps, algorithmic assumptions (Batty, 2018; Leszczynski, 2020; Goodchild, 2007; Biljecki *et al*., 2015). Metadata models and methods have been studied for decades to manage and communicate the quality of geospatial data, distinguishing so-called internal quality from so-called external quality (Devillers *et al.*, 2010). The internal quality of geospatial data corresponds to how well it describes a given reality and is documented through three different kinds of metadata:

Identified causes	Truth problems	Impact on analyses and decisions	Metadata intensive solutions
Sensors used under non-compliant conditions, measurement noise. Errors during derivation of new data	Inaccuracy	Incorrect decisions, false alerts, loss of system reliability	Standardization of measurement protocols. Provenance metadata that can be looked up to detect the cause Quality control using inconsistency rules, reference data
Heterogeneous formats, non-harmonized units, integration of data with different precisions, schema divergences	Inconsistency	Impossible integration, incompatible systems, degraded reliability	Cross-checking methods between sensors to identify contradictory values; detection of artifacts via comparison with expected spatio-temporal gradients; multi-station consistency control using spatio-temporal correlation models. Documenting the precision of every data source and commenting the possible changes to precision brought by data processing,
Sensor hardware failures, network interruptions, untransmitted data.	Incompleteness	Estimation bias, spatial and/or temporal gaps, artefacts	Product specification with a commitment to reach completeness ratio, and quality control procedure (SUBDENSE) Multidimensional modeling of completeness (spatial, temporal, sensor quality); Reconstruction through time series. (POLLUSCOPE)

Table 1. Three main truth problems in DTT that can be addressed through metadata.

- the specifications –what part of reality the data intends to describe and how-,
- the measures of different gaps between the intended description and the actual data, a.k.a. quality criteria,
- the provenance information, description of observation modalities, and modelling steps.

Practically, designing explicit and detailed data specifications is a complex task, mainly undertaken by publicly mandated authorities whose missions comprehend the definition of such specifications to guide the production of a homogeneous data product on a given territory.

In the POLLUSCOPE project, different metadata are integrated and exploited to document the uncertainties of the pollution model (Brahem *et al*., 2020; Mehanna 2023); 1) GPS coordinates and timestamp to contextualize the measurement, 2) Sensor model and ID to trace the source, 3) Calibration date and method to assess reliability, 4) Ambient temperature and humidity to correct environmental biases and 5) battery level to detect potential sensor degradation. Each data point collected by a mobile sensor is associated with these different metadata, which are used to compute confidence levels for pollutant concentration measurements. Together with consistency rules, this proved useful to distinguish actual pollution peaks from sensor anomalies due to overheating or low power.

In the SUBDENSE project, the detailed specifications of the French topographic building data, BD-Topo, and other quality metadata, is used to evaluate if the creation of a building feature in the database can be interpreted as a probable construction of a building entity in reality or if it can be explained by the fact that between 2008 and 2010, smaller buildings were integrated in the product scope when merging French topographic data and French cadastral data (Bucher *et al.*, 2025). The specifications of the French and German product are also useful to avoid comparison biases: the building geometry

in French BDTopo is the gutters until 2013, and the building geometry in the German GeoBASIS product is the walls footprint.

To make metadata available for a DTT truth metalanguage, there is a need to manage metadata attached to the different data integrated in a DTT and to organize them to support transversal queries to assess the truth of different propositions P elaborated in a DTT story. This advocates firstly for keeping data and its metadata in the best format to manage them instead of aggregating them. Data that corresponds to different forms of uncertainties and imprecisions in urban data, or to different modalities of observing and surveying reality, should be processed close to their source and integrated only when it is necessary. This also calls for better metadata infrastructure interconnected with data infrastructure. In many cases, metadata is delivered as separate files from data. Engineers process the data and store the files elsewhere. As a result, there is no intuitive interface for a user to access metadata from the data. Advances in the automatic processing of textual data opens new perspectives to use metadata across different formats and including text files (Benjira *et al.*, 2025).

II. COMMUNICATION AND RELEVANCE

Human capacity to read and tell stories based on propositions elaborated in a DTT is a complementary stake to assessing the intrinsic truth of these propositions. It is analysed in this section with the help of the communication theory.

A. The triad of communication theory

The foundational communication theory of Shannon and Weaver (1949) identifies three complementary levels of communication that apply to any kind of communication:

- *Syntactic level*: establishment of a common code between sender and receiver to ensure accurate transmission of symbols, independent of meaning.
- *Semantic level*: shared understanding of symbols to avoid ambiguity and misinterpretation.
- *Pragmatic level*: efficiency of communication to produce the intended response or action in its context of use.

The usage of the triad is illustrated hereafter, on the generation of real-time personal notifications of pollution peaks to inhabitants, in the POLLUSCOPE project (See Figure 4 below). *Syntactic* communication requires the definition of standardized formats for alerts, including units of measurement and spatial references, to ensure interoperability and legibility across devices and interfaces. It refers to syntactic interoperability between the server and the smartphone, and to human readability. *Semantic* communication addresses the interpretation of the message by the recipient. An alert such as "NO_2 spike" must be unequivocally understood as a health hazard. Semantic clarity involves avoiding ambiguity and tailoring terminology to the needs of sensitive populations (e.g., individuals with respiratory conditions, children). *Pragmatic* communication refers to the effectiveness of the notifications to drive some users away from polluted areas. The alert should prompt the vulnerable user to modify his trajectory and move away from the polluted area. Pragmatic success is thus measured by the extent to which the system influences user behavior and reduces exposure. As illustrated in Figure 4, there is an increasing complexity in these communication levels, and the challenge lies at the pragmatic level: *transforming information into real protective action.*

Interoperability standards have emerged to support the communication of information between different machines or humans, in particular these promoted by the Open Geospatial Consortium that are dedicated to communication of geographic information and these promoted by the World Wide Web Consortium that are dedicated to communication on the Web. A pending issue in interoperability is that of semantics.

Communicating semantics level is addressed through the definition of ontologies. Core categories of knowledge have been identified in the literature on spatial information: location, neighborhood, field, object, network, event, granularity, accuracy, meaning, and value (Kuhn, 2012; Kokla and Guilbert, 2020). In the SUBDENSE project, the identification of concepts definitions for topographic entities, such as *Buildings*, is easy. Other concepts revealed ambiguous, like the notion of *Built up area,* that has various meanings depending on the user's expertise in quantitative geography. A *Vacant plot* was often used in questions or descriptions but revealed a polysemic expression that did not refer to a well-defined concept. Last, the main challenge was the identifica-

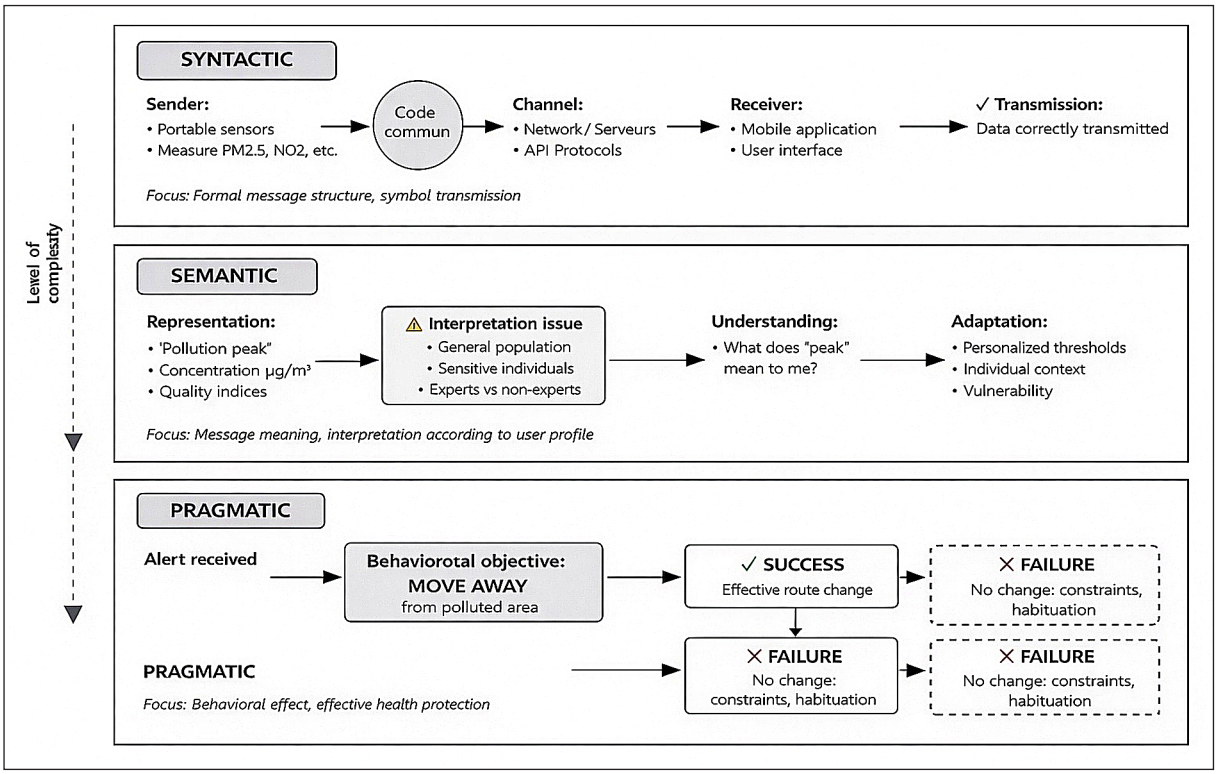

Figure 4. Application of Shannon's communication triad to the Polluscope approach, showing the three levels of communication in a DTT-based air pollution alert system.

tion of shared concepts to describe dynamics such as densification. More generally, discussing time and dynamics is challenging (Gangemi *et al*., 2002). The meta-ontology DOLCE extends upon philosophical principles to define a set of concepts useful to analyze statements in natural language that relate to time, change and processes (ibid).

The most used media to communicate information about space are 3D models and maps. At a syntactic level, map makers identify the minimum size for symbols, as well as how color contrasts can alter perceived colors. The semantic level needs to consider how the human eye interprets graphical relations as semantic relations, as well as connotations of specific colors and conventions (Tufte, 1997). To design maps of BuildingChanges in the SUBDENSE project, red was used for demolitions while green was used for constructions. A stripped symbolization was used so that both changes could remain visible even when overlaid, because in many cases a demolition is followed by the construction of new buildings with overlapping footprints.

B. Relevance

In a DTT context, the objective is not to determine and communicate absolute truth but rather to extract and prioritize the most *relevant information* to optimize the reliability of analysis and decision-making processes. Chachoua and Pacholczyk (2002) extend Shannon's information theory (Shannon, 1948; Shannon and Weaver, 1949) to introduce a cognitive layer that is indispensable for transforming technical data into information with high practical value. This layer is the notion of contextual relevance, defined by:

- the needs of an actor or decision-maker,
- the territorial and temporal situation,
- the operational or strategic objectives being pursued,
- the knowledge already available.

Chachoua and Pacholczyk (2002) define the informational relevance of a message *as the degree of ignorance reduction brought by that message to the state of knowledge during problem-solving*. This definition establishes a functional relationship between relevance and epistemic uncertainty reduction.

It is illustrated in the context of POLLUSCOPE project. The initial state of knowledge of asthmatic users who plans their daily route through an urban area is characterized by complete ignorance regarding the spatial and temporal distribution of air pollution along their potential route. The relevance

R of a communication is evaluated on a scale of 0 (irrelevant) to 1 (perfectly relevant).

1. Scenario 1

The system sends a generic alert stating "High pollution in the city today". This message has very low relevance ($R \approx 0.2$) because it only marginally reduces the user's ignorance. They know there is pollution somewhere, but have no precise spatial information to modify their behavior or adapt their route.

2. Scenario 2

The system sends a geolocated alert stating "NO_2 peak detected on Victor Hugo Avenue (your usual route) between 8am and 9am: 180 µg/m³ (exceeding the threshold of 100 µg/m³)". This message has high relevance ($R \approx 0.8$) because it significantly reduces ignorance: the user now knows the pollutant concerned, its concentration, its precise location, and the timing of the peak. This information enables them to make an informed decision (take an alternative route, postpone their trip).

3. Scenario 3

The system adds a contextualized recommendation: "NO_2 peak detected on your usual route (Victor Hugo Avenue, 180 µg/m³). Alternative route suggested via the municipal park (45 µg/m³), adding 5 minutes to your journey". This message reaches near-perfect relevance ($R \approx 0.95$) because it almost completely eliminates ignorance by providing not only the risk diagnosis, but also an immediately applicable operational solution. The user no longer needs to search for an alternative route themselves or estimate pollution levels elsewhere.

The progression from relevance $R = 0.2$ to $R = 0.95$ illustrates how progressive enrichment of the message (spatialization, temporalization, contextualization, prescription) reduces epistemic uncertainty and increases informational relevance for solving the concrete problem: "How can I avoid pollution exposure during my trip?"

There are two complementary challenges to model relevance in a DTT context.

The first challenge consists of evaluating the contribution of heterogeneous data sources to answer specific spatial queries. In the case where the data is integrated, the approach by Chachoua and Pacholczyk (2002) proposes a graded relevance function $P \in [0,1]$ quantifying an information's capacity to reduce the system's symbolic entropy. This formalization enables simultaneously handling multidimensional imperfections (geometric imprecision, thematic uncertainty, obsolescence), dynamically contextualizing relevance criteria according to territorial stakes, and ensuring epistemic traceability of informational choices. In the case where the data is in the DTT catalogue but is not yet integrated, it can be necessary to evaluate such a potential contribution prior to selecting and integrating the data. For example, in SUBDENSE project, the municipality wants to diagnose the elevations of buildings in the past decade on its territory. The DTT can investigate different options: detect elevation from topographic data and Digital Elevation Models, or use the national Building Permits database. In that case, there is a need for evaluating the relevance based on available metadata and prior to integrating the data.

The second challenge is to support the elicitation of information needs when it depends on user context and often implicitly pursued objectives. For example, a political decision-maker may have a legitimate goal of being re-elected, which will impact the information needed for public debates. Another municipality can wish to use a DTT research prototype to engage different generations, during dissemination events, which requires designing different interfaces adapted to digital natives on the one hand, and to more traditional map readers on the other hand. Next-generation geographic information systems integrate dimensions of cognitive, situational, and intentional relevance (Saracevic, 2007) through conversational interfaces and consolidate this information in knowledge graphs. Such approach leverages learning from collective interactions to progressively refine relevance models (Li *et al*., 2025; Zabala *et al*., 2021). The SUBDENSE project illustrates this approach by organizing multi-user feedback on the identification of relevant concepts of BuildingChanges, between users pursuing similar objectives (Bucher *et al*., 2025).

CONCLUSION AND PERSPECTIVES

In this paper, we argued that a key functionality of digital twins of territories is their capacity to build and share trustworthy stories about territories.

This paper investigated the application of theories of truth and theory of communication to this DTT context.

Tarski's truth theory is used to analyse the capacity to evaluate the truth value of propositions in a DTT. We presented the closely related notions of truth, imprecision, and uncertainty. Existing metadata could be reused to achieve a truth metalanguage for propositions expressed in Digital Territorial Twins (DTTs). To explicitly state uncertainties, data should be kept close to metadata and processed close to their source. There is a need for metadata infrastructures that support joint queries of metadata to evaluate the truth, uncertainties and imprecisions, of propositions elaborated with multiple data. Shannon's theory of communication as well as a later theory of information relevance ground our analysis of required conditions to ensure the effectiveness of interactions in uncertain environment. It is necessary to better integrate users' feedback to elaborate a relevance model for digital twins to better align with human decision processes.

As a conclusion, DTTs power to simulate urban phenomena, while valuable, entails significant risks, particularly when interpretive assumptions, modelling defaults, or outdated data are not made explicit. This paper emphasizes the importance of representing not only data, but also the epistemic status of data: its relevance, uncertainty, precision, and temporal validity. Ensuring the trustworthiness and utility of DTT requires systems that reason with, communicate about, and justify their representations of the territory. The reflection presented in this paper advocates for a metadata-centric ecosystem that supports the filtering, weighting, and contextual interpretation of data, based on how much uncertainty it reduces with respect to user well-defined objectives, and for gathering users feedback tto increase the DTT capacity to dialogue with users and to gain inspiration from other territories.

This work was partly funded under the Open Research Area 7 framework by the French National Research Agency (ANR) grant number 22-ORAR-0004.

ACKNOWLEDGMENTS

The authors wish to express their warm gratitude to editors and reviewers for their thorough and constructive reviews and suggestions. Their work and insight were most helpful to reframe our initial reflexion and improve the clarity of the paper.

BIBLIOGRAPHY

Batty, M. (2018). *Digital twins. Environment and Planning B: Urban Analytics and City Science*, 45(5), 817–820. https://doi.org/10.1177/2399808318796416

Biljecki, F., Ledoux, H. & Stoter, J. (2015). Improving the consistency and usability of 3D city models with semantic and geometric data validation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 111, 13–29.

Benjira, W., Atigui, F., Bucher, B., Grim-Yefsah, M. & Travers, N. (2025). Web Open Data to SDG Indicators: Towards an LLM-Augmented Knowledge Graph Solution. *In* Barhamgi, M., *et al*. *Web Information Systems Engineering* – WISE 2024 PhD Symposium, Demos and Workshops. WISE 2024. Lecture Notes in Computer Science, vol 15463. Springer, Singapore. https://doi.org/10.1007/978-981-96-1483-7_7

Brahem, M., *et al*. (2020). Data perspective on environmental mobile crowd sensing. *In Intelligent Environmental Data Monitoring for Pollution Management* (pp. 269-288). Academic Press.

Bruynseels, K., Santoni de Sio, F. & van den Hoven, J. (2018). Digital twins in health care: Ethical implications of an emerging engineering paradigm. *Frontiers in Genetics*, 9, 31.

Bucher, B., Hein, C., Raines, D. & Gouet Brunet, V. (2021). Towards Culture-Aware Smart and Sustainable Cities: Integrating Historical Sources in Spatial Information Infrastructures. *ISPRS Int. J. Geo-Inf.,* 10, 588. https://doi.org/10.3390/ijgi10090588

Bucher, B., Raimbault, J., Ndim, M., Raimond, A.-M., Perret, J., Dembski, S. & Jehling, M. (2025). A Model of Building Changes to Support Comparative Studies and Open Discussions on Densification. *ISPRS International Journal of Geo-Information*, 14(4), 155. https://doi.org/10.3390/ijgi14040155

Callcut, M., Cerceau Agliozzo, J., Varga, L. & McMillan, L. (2021). Digital twins in civil infrastructure systems. *Sustainability*, 13(20). https://doi.org/10.3390/su132011549

Chachoua, M. (1998). *Une théorie symbolique de l'entropie pour le traitement des informations incertaines*. Thèse de doctorat, Université d'Angers.

Chachoua, M. & Pacholczyk, D. (2002). *Gradual relevance and data imperfection handling*. Proceedings of the 10th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU).

Chachoua, M. & Pacholczyk, D. (2006). Reasoning under Uncertainty and Incompleteness: Towards a

Unified Framework. *In Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (pp. 53–64). Springer.

Devillers, R., Stein, A., Bédard, Y., Chrisman, N., Fisher, P. & Shi, W. (2010). Thirty Years of Research on Spatial Data Quality: Achievements, Failures, and Opportunities. *Transactions in GIS*, 14, 387-400. https://doi.org/10.1111/j.1467-9671.2010.01212.x

Dubois, D. & Prade, H. (1988). *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Springer.

Ellul, C., Stoter, J. & Bucher, B. (2022). Location-enabled digital twins – understanding the role of NMCAs in a European context. *3D GeoInfo* 2022, 17th 3D GeoInfo Conference, ISPRS, Oct 2022, Sydney, Australia, pp.53-60. https://doi.org/10.5194/isprs-annals-X-4-W2-2022-53-2022

Ellul, C., Stoter, J., Bucher, B., Olsson, P.-O., Billen, R. & DeLathouwer, B. (2024). Towards National Connected Digital Twins: A Geospatial Perspective. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10(4/W5), 147-154. https://doi.org/10.5194/isprs-annals-X-4-W5-2024-147-2024

Ehrhardt, D., Sommer, F., Ndim, M., Raimbault, J., Hartmann, T. & Jehling, M. (2025). Who drives urban densification? *Linking landownership and spatial dynamics*. https://doi.org/10.14512/rur.2977

Evers, D., Katuric, I.& van der Wouden, R. (2024). *Urbanization in Europe, Past Developments and Pathways to a Sustainable Future.* Palgrave Macmillan.

Floridi, L. (2011). *Semantic conceptions of information*. Stanford Encyclopedia of Philosophy.

Fuller, A., Fan, Z., Day, C. & Barlow, C. (2020). Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access*, 8, 108952–108971.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. & Schneider, L. (2002). Sweetening ontologies with DOLCE. *In Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web; Springer: Berlin/Heidelberg*, Germany; pp. 166–181.

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), 211–221.

Goodchild, M. F. & Li, L. (2021). Digital twins. *In Geographical Information Science* (Springer).

Grieves, M. & Vickers, J. (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. *In Transdisciplinary Perspectives on Complex Systems* (pp. 85–113). Springer.

Kokla, M. & Guilbert, E. (2020). A Review of Geospatial Semantic Information Modeling and Elicitation Approaches. *ISPRS International Journal of Geo-Information*, 9(3), 146. https://doi.org/10.3390/ijgi9030146

Kritzinger, W., Karner, M., Traar, G., Henjes, J. & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11), 1016-1022.

Kuhn, W. (2012). Core concepts of spatial information for transdisciplinary research. 2012, *Int. J. Geogr. Inf. Sci*., 26, 2267–2276.

Leszczynski, A. (2020). Platform urbanism: negotiating 'deep learning' spatial data infrastructures. *Urban Geography*, 41(9), 1268–1275.

Li, H., Yue, P., Wu, H., Teng, B., Zhao, Y. & Liu, C. (2025). A question-answering framework for geospatial data retrieval enhanced by a knowledge graph and large language models. *International Journal of Digital Earth*, 18(1). https://doi.org/10.1080/17538947.2025.2510566

Mehanna, S. (2023). *Data quality issues in mobile crowdsensing environments*. PhD thesis, Paris-Saclay University.

Monmonier, M. (1991). *How to lie with maps.* University of Chicago Press, Chicago.

Qanazi, S., Leclerc, E. & Bosredon, P. (2025). *Toward a Comprehensive Definition of Urban Digital Twins: Bridging Gaps for Smart City Development*. https://doi.org/10.1007/978-3-031-88653-9_64

Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 58(13), 1915–1933.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379–423, et 623–656.

Shannon, C. E. & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.

Smets, P. (1994). What is Dempster-Shafer's model? *In Advances in the Dempster-Shafer Theory of Evidence* (pp. 5–34). Wiley.

Smets, P. & Kennes, R. (1994). The Transferable Belief Model. *Artificial Intelligence*, 66(2), 191–234.

Tarski, A. (1936). The concept of truth in formalized languages. *In Logic, Semantics, Metamathematics*. Oxford University Press (translated edition, 1956).

Tarski, A. (1944). The Semantic Conception of Truth and the Foundations of Semantics. *Philosophy and Phenomenological Research*, 4(3), 341–376.

Tufte ER. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire; Graphics Press; 1997

United Nations (2017). Resolution Adopted by the General Assembly, 71/313. *Work of the Statistical Commission Pertaining to the 2030 Agenda for Sustainable. Development* A/RES/71/313. http://ggim.un.org/documents/A_RES_71_313.pdf

Wood, D. (2010). *Rethinking the Power of Maps*. Guilford Press.

Zabala, A., Masó, J., Bastin, L., Giuliani, G. & Pons, X. (2021). Geospatial User Feedback: How to Raise Users' Voices and Collectively Build Knowledge at the Same Time. *ISPRS Int. J. Geo-Inf.*, 10, 141. https://doi.org/10.3390/ijgi10030141

Zhou, K., Liu, T. & Zhou, L. (2015). Industry 4.0: Towards future industrial opportunities and challenges. *Procedia CIRP*, 63, 8–13.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8(3), 338–353.

Coordinates of authors:

Bénédicte BUCHER
Université Gustave Eiffel, LASTIG, IGN-ENSG
FRANCE
benedicte.bucher@ign.fr

Mohamed CHACHOUA
Université Gustave Eiffel, LASTIG, IGN-ENSG
FRANCE
chachoua@eivp-paris.fr