CHATGPT IN SCHOOL GEOGRAPHY: SCOPING REVIEW (2023–2025) AND DESIGN OF A SUPPORT/RISK MATRIX BY TASKS IN SECONDARY EDUCATION

CHATGPT NA GEOGRAFIA ESCOLAR: REVISÃO DE ESCOPO (2023–2025) E DESENHO DE UMA MATRIZ DE APOIO/RISCO POR TAREFAS NO ENSINO SECUNDÁRIO

CHATGPT EN LA GEOGRAFÍA ESCOLAR: SCOPING REVIEW (2023–2025) Y DISEÑO DE UNA MATRIZ DE APOYO/RIESGO POR TAREAS EN SECUNDARIA

https://doi.org/10.56238/sevened2026.011-004

Carlos Guallart[1]

ABSTRACT

This article analyzes the possibilities and limitations of ChatGPT as a support tool for teaching Geography in Secondary Education, from a "teaching task system" approach. Based on a recent review of generative artificial intelligence (GAI) in education and an applied categorization of curricular and professional tasks specific to Geography teaching, a unified matrix is constructed that relates macroprocesses (planning, resource design, instruction, and assessment) to the expected level of support provided by ChatGPT and the risk of cognitive substitution. The study is based on a documentary scoping analysis (2023–2025; n = 54 documents) and an applied task categorization, without empirical classroom validation at this stage. The results show high performance in linguistic structuring tasks (drafting materials, text adaptation, rubrics, and item banks with verification), medium performance in tutoring and methodological support under supervision, and significant limitations in tasks requiring factual accuracy, professional judgment in summative assessment, and advanced spatial reasoning without GIS mediation. A workflow and four demonstrative learning situations are proposed, aimed at preserving authorship, documenting processes, and strengthening geographic thinking through verification, source comparison, and reasoned argumentation.

Keywords: Generative Artificial Intelligence. ChatGPT. Geography Education. Secondary Education. Process-Oriented Assessment. Academic Integrity. Spatial Thinking. Teaching Task Matrix. Documentary Analysis. Scoping Review.

RESUMO

Este artigo analisa as possibilidades e os limites do ChatGPT como ferramenta de apoio ao ensino de Geografia na Educação Secundária, a partir de uma abordagem de "sistema de tarefas docentes". Com base em uma revisão recente sobre inteligência artificial generativa

[1] Master's Degree in Territorial and Environmental Planning. Colegio Santa María del Pilar (marianistas) de Zaragoza. E-mail: cguallart@gmail.com Orcid: 0000-0001-5232-1082

(IAG) na educação e em uma categorização aplicada de tarefas curriculares e profissionais próprias do ensino de Geografia, constrói-se uma matriz unificada que relaciona macroprocessos (planejamento, design de recursos, instrução e avaliação) com o grau de apoio esperado do ChatGPT e com o risco de substituição cognitiva. O estudo fundamenta-se em uma análise documental de escopo (2023–2025; n = 54 documentos) e em uma categorização aplicada de tarefas, sem validação empírica em sala de aula nesta fase. Os resultados indicam alto desempenho em tarefas de estruturação linguística (rascunhos de materiais, adaptação de textos, rubricas e bancos de itens com verificação), desempenho médio em tutoria e apoio metodológico sob supervisão, e limitações relevantes em tarefas que exigem precisão factual, julgamento profissional na avaliação somativa e raciocínio espacial avançado sem mediação de TIG. Propõe-se um fluxo de trabalho e quatro situações de aprendizagem demonstrativas, orientadas à preservação da autoria, à documentação do processo e ao fortalecimento do pensamento geográfico por meio de verificação, contraste de fontes e defesa argumentada.

Palavras-chave: Inteligência Artificial Generativa. ChatGPT. Didática da Geografia. Educação Secundária Obrigatória. Avaliação Processual. Integridade Acadêmica. Pensamento Espacial. Matriz de Tarefas Docentes. Análise Documental. Revisão de Escopo.

RESUMEN

Este artículo analiza las posibilidades y límites de ChatGPT como herramienta de apoyo en la docencia de Geografía en Educación Secundaria desde un enfoque de “sistema de tareas docentes”. A partir de una revisión reciente sobre inteligencia artificial generativa (IAG) en educación y de una categorización aplicada de tareas curriculares y profesionales propias de la enseñanza de la Geografía, se construye una matriz unificada que relaciona macroprocesos (planificación, diseño de recursos, instrucción y evaluación) con el grado de apoyo esperable de ChatGPT y con el riesgo de sustitución cognitiva. El estudio se basa en un análisis documental de alcance (2023–2025; n = 54 documentos) y en una categorización aplicada de tareas, sin validación empírica en aula en esta fase. Los resultados muestran un rendimiento alto en tareas de estructuración lingüística (borradores de materiales, adaptación de textos, rúbricas y bancos de ítems con verificación), un rendimiento medio en tutoría y apoyo metodológico bajo supervisión, y limitaciones relevantes en tareas que exigen precisión factual, juicio profesional en evaluación sumativa y razonamiento espacial avanzado sin mediación de TIG. Se propone un flujo de trabajo y cuatro situaciones de aprendizaje demostrativas orientadas a preservar la autoría, documentar proceso y reforzar pensamiento geográfico mediante verificación, contraste de fuentes y defensa argumentada.

Palabras clave: Inteligencia Artificial Generativa. Chatgpt. Didáctica de la Geografía. Educación Secundaria Obligatoria. Evaluación Procesual. Integridad Académica. Pensamiento Espacial. Matriz de Tareas Docentes. Análisis Documental. Scoping Review.

1 INTRODUCTION

The incorporation of ChatGPT in school geography poses a problem of design and validity of tasks, not only of technological adoption. The tool can support textual production and revision, but also displace decisions, inferences and contrast evidence, deteriorating the correspondence between the cognitive effort of the students and the "deliverable" that is evaluated. This article adopts a teaching task system approach and develops a scoping review (2023–2025) aimed at the applied design of a task-based support/risk operational matrix, organized by macro-processes (planning, resource design/adaptation, instruction, and evaluation/feedback). The contribution consists of offering criteria to decide where support is didactically defensible and where the risk of cognitive substitution increases, incorporating safeguards based on verification, traceability and argued defense.

In school geography, this technological disruption intersects with a discipline with its own epistemological and didactic features. The teaching of Geography is aimed at the development of spatial thinking, understood as the ability to understand, explain and project space through structuring concepts such as location, scale and interrelation. In a complementary way, spatial literacy involves thinking about space in an informed and critical way to address territorial problems. In addition, as a school subject, Geography requires explaining reality and territorial problems, handling spatial sources, and developing reasoning that connects local and global scales (Adamini, Cafferata, & Velo, 2024). In this sense, it offers a favourable framework for developing critical thinking linked to the analysis of the territory and the interrelations between society and the environment.

The central problem lies in the possible rupture between the learning process and the deliverable, given that GAI can generate results that do not have a solid relationship with the cognitive effort of the students. In subjects that require an understanding of complex spatial and social phenomena, an uncritical use of AI can lead to "metacognitive laziness" and the loss of essential cognitive skills when technology replaces, rather than supports, students' efforts (Vallejo, 2024; Xiaoyu et al., 2025). At the same time, AGI poses a dual challenge: it can facilitate the personalization and simulation of scenarios, but it also entails risks of substitution of reasoning and inquiry (García & Pujol, 2025). Although the specific literature on AGIs in Geography is still incipient, there is evidence in the field of social studies that indicates that chatbots can improve academic performance and knowledge retention if they are properly integrated into didactic design (Sutabri et al., 2025; Yetişensoy & Karaduman, 2024).

The purpose of this article is to analyze the possibilities and limits of ChatGPT as a support tool within the system of teaching tasks in the teaching of Geography in Secondary

Education, proposing an integration matrix that preserves academic integrity and reinforces the role of teachers as pedagogical designers. The analysis also delimits in which aspects tools such as ChatGPT are insufficient with respect to the geospatial skills required and frames its use as an instrument that, under appropriate mediation, can promote a dialectical reflection on the socio-territorial reality without replacing the teaching function.

It should be emphasized that this work is documentary in nature and of applied design: it does not present an intervention in the classroom or data on the performance of students or teachers. Therefore, the matrix is offered as a reasoned guide to support planning, instruction, and evaluation decisions, and not as empirical evidence of educational impact. Its usefulness and validity should be tested in subsequent implementation studies.

With this objective, the work poses three research questions: (1) what types of teaching and curricular tasks of Geography in Secondary Education are most likely to receive useful support from ChatGPT and to what degree? (2) In which tasks does the risk of cognitive substitution increase and decrease the validity of the final product as evidence of learning? and (3) what didactic design decisions (process evidence, verification, traceability and contrast of sources) allow the tool to be integrated responsibly? The main contribution is a unified support/risk matrix and an applicable workflow that translates general AI integration frameworks to specific conditions of School Geography.

Based on this synthesis, an operational matrix is designed that classifies tasks by degree of support and risk of substitution, and that is accompanied by conditions of use and process evidence to guide its application in departments and centers.

2 CONCEPTUAL FRAMEWORK

2.1 TEACHING TASKS: OPERATIONAL DEFINITION AND LEVELS

Teaching practice can be described as a system of tasks that is articulated in recurrent macro-processes. In an operational definition, these macro-processes include lesson planning, generation and adaptation of teaching materials, direct classroom instruction, and assessment with feedback (Faculty AI, 2024). In the context of AI, this perspective shifts the focus from the logic of content transmission – associated with "banking education" – to pedagogical leadership of a strategic nature, where teachers design, decide and guide the learning experience (Vallejo, 2024).

To analyze where AI can intervene, it is useful to distinguish levels of intervention within the task system. At the macro-process level, focused on planning, teachers act as learning architects: they select content, sequence basic knowledge and structure the curriculum, with the possibility of algorithmic support to organise and propose design alternatives (Vallejo,

2024). At the task level, understood as didactic execution, there are specific actions (explanations, activities, guided practices) where AI can operate as an assistant or tutor, supporting the creation of resources, the generation of questionnaires, the preparation of case studies, and the preparation of review activities (Adiguzel et al., 2023; Cronje et al., 2023). At the product level, associated with evaluation and final evidence, IAG stresses traditional approaches based solely on deliverables, since it facilitates the generation of products without necessary correspondence with the learning process; therefore, it is proposed to prioritize evidence distributed during the process and designed to document the cognitive effort of students (García & Pujol, 2025).

This operationalization is aligned with frameworks that structure teaching practice in broad dimensions. From the EurekAI competency model, adapted from the DigCompEdu framework, teaching tasks range from professional development and pedagogical reflection to classroom practice and student training, understanding each task as a means to produce learning evidence and reduce the gap between effort and result that HAI can cause (García & Pujol, 2025). In parallel, specific use cases where AI intervenes directly have been described, such as the "lesson plan generator", the "review and feedback activity generator" and the "essay marker". Collectively, the level of the task can range from automating routine administrative processes to supporting the design of complex learning experiences (Faculty AI, 2024).

2.2 CHATGPT AS A LINGUISTIC AND COGNITIVE SUPPORT

In the case of language models, this translates into the production of coherent text and the response to queries in natural language, with variability in precision and stability depending on the context of use and the design of the prompt. ChatGPT, based on the *Generative Pre-trained Transformer* (GPT) architecture, operates through probabilistic inference, predicting the next word in a sequence; it does not "understand" the world, but simulates human conversations with high fidelity (Office of Science and Technology of the Congress of Deputies, 2024; Organisation for Economic Co-operation and Development [OECD], 2023; Osamor et al., 2023). In addition, in the current ecosystem it coexists with other tools – such as Copilot (Microsoft) and Gemini (Google) – described as multimodal and capable of processing text, images and code.

From a didactic perspective, its value is concentrated in linguistic and cognitive support functions: it can facilitate conceptual clarification, offer immediate feedback, support the generation of ideas, and sustain continuous interactions, which opens up options for personalizing itineraries and maintaining support outside school hours (Osamor et al., 2023;

Vallejo, 2024). They may also play specific conversational roles—for example, as a "Socratic opponent" or as a "co-designer"—that favor the testing of arguments and the revision of drafts, provided that a demanding verification framework is adopted (Area, 2025; Urmeneta & Romero, 2024).

Its limits, however, are structural and condition its educational use. These systems are described as "black boxes" with explainability problems, and can generate "hallucinations", that is, inaccurate or false information with a plausible appearance (Area, 2025; Bobula, 2024; Ali Elkot et al., 2025). In addition, they incorporate biases associated with the training data and lack human ethical judgment, which forces us to assess the adequacy of their responses in terms of equity, relevance, and pedagogical responsibility (Lorenz, Perset, & Berryhill, 2023). In addition, they do not have real-world understanding and do not guarantee reliable performance in complex mathematical logic without additional tools, which limits their use when formal validations or high precision are required (Cornell University, 2023). For these reasons, a use with continuous human supervision ("*human-in-the-loop*") is recommended, aimed at validating accuracy, didactic quality, and adequacy to the context (Bernal et al., 2025; U.S. Department of Education, 2023). Along these lines, a "golden rule" is formulated for formative use: ask the AI about what is already known, so that students contrast, validate and criticize the answer from previous knowledge (García & Pujol, 2025).

2.3 PARTICULARITIES OF SCHOOL GEOGRAPHY

School Geography presents specific requirements that condition the integration of the IAG. In the twenty-first century, its teaching requires moving from the classroom to the territory, incorporating field trips and Geographic Information Technologies (GITs) to promote learning linked to sustainability and environmental awareness (Correa, López & Díaz, 2024). In terms of competencies, the discipline requires developing skills such as spatial perception, spatial reasoning and spatial citizenship, which involve understanding spatial relationships, interpreting patterns and taking an informed position on territorial problems. This approach includes cartographic literacy and the use of TIG to analyse and interpret spatial phenomena, placing working with data, representations and scales at the core of geographical learning (Guallart, 2025b).

These particularities introduce specific challenges for mainly language-based tools. Unlike purely linguistic tasks, geographical thinking requires manipulating and interpreting spatial data and topological relationships, so the intersection with textual language models presents limitations when factual accuracy, complex spatial reasoning, or direct generation of cartographic representations are required (Tao & Xu, 2023). The evidence available in

nearby areas suggests, however, that chatbots can add value if they are designed with clear objectives: in the teaching of Geography and Social Sciences, experiences such as "*Sosyalcibot*" have been associated with improvements in motivation and performance by providing explanations (including visuals) and immediate feedback (Yetişensoy & Karaduman, 2024). At the same time, variable performance has been observed in tasks that require high precision or demanding spatial reasoning, with results conditioned by the difficulty of the task (Salman et al., 2025).

In this framework, the IAG can support school geography without replacing spatial experience or working with sources, but acting as a mediator in tasks such as explaining data analysis methodologies, generating draft maps in textual formats (e.g. ASCII code) or simulating future territorial scenarios, always with verification and contrast with reliable data and representations (Guallart, 2025a). Likewise, it can contribute to the search for information and the comparison of perspectives on conflicting and socially relevant territorial problems, while teachers maintain the responsibility of defining the territorial cut, selecting sources, demanding precision and promoting reasoned reflection (Adamini, Cafferata & Velo, 2024).

Given that part of the debate on the educational use of IAG is articulated in institutional guides and technical documents that are quickly updated, the work incorporates grey literature (reports, *policy papers* and guides) from educational bodies along with academic studies. In the analysis, guidelines are used to describe operational practices and recommendations, while claims about learning effects and risks are preferably supported by published reviews and studies.

3 METHODOLOGY

3.1 CORPUS AND UNIT OF ANALYSIS

The study adopts a design of documentary analysis and applied categorization aimed at producing a didactic support instrument (matrix). The corpus is composed of two sets: (a) recent academic literature on HAIs in education and their didactic integration (2023–2025), selected by thematic relevance (HAG/LLM, assessment, academic integrity, task design and/or Social Sciences) and by type of contribution (evidence-based reviews, case studies, frameworks and guides); and (b) a didactic-curricular corpus made up of representative tasks of the teaching of Geography in Secondary Education and associated common knowledge/content (e.g. territory and landscape, demographic dynamics, economic activities and globalisation, urbanisation and spatial planning, socio-environmental challenges and

geopolitics), in order to translate evidence and general frameworks into concrete planning decisions, instruction and assessment.

A bilingual search strategy based on three blocks of terms (IAG/LLM; geographic education; secondary stage) and iterative adjustment was applied through pilot tests. The body of the article describes the inclusion/exclusion criteria and the composition of the corpus, while the full equation and variants are presented in Annex I to facilitate replication. The differential methodological core is the coding by tasks: inventory of typical tasks, definition of criteria to estimate degree of support and risk of substitution, classification A/M/B/N and contrast with an integration framework (EIADD) to ensure coherence between level of intervention and didactic design.

Selection was completed by reference review (*snowballing*) and screening for thematic relevance and applicability to school contexts. In the didactic-curricular corpus, the tasks were defined based on their representativeness in usual units of Compulsory Secondary Education (ESO) and their correspondence with basic knowledge, prioritizing those that generate assessable products and, therefore, are more sensitive to cognitive substitution.

Table 1

Characteristics of the documentary corpus and search/selection strategy (n = 54)

Dimension	**Specification**
Study design	Scoping Desk Analysis (2023–2025) and Applied Categorization Aimed at Producing a Task-Based Operational Support/Risk Matrix.
Search window	September 1, 2025 – December 1, 2025.
Sources consulted	Scopus, Web of Science, ERIC and Google Scholar.
Search languages	Spanish and English (bilingual equations).
Search structure	Combination of three blocks of terms: (a) IAG/LLM; (b) Geography/geographic education; (c) educational stage (secondary). Iterative adjustment through pilot testing, incorporating additional terms where necessary (e.g., "spatial thinking", "geographical thinking", "GIS").
Inclusion criteria	(i) Publication 2023–2025; (ii) focus on educational uses of IAG/LLM or assessment/integrity; (iii) applicability to the school context or teacher training.
Exclusion Criteria	Informative entries without identifiable authorship or without bibliographic traceability; Documents

Dimension	Specification
	unrelated to geographic education or not relevant to secondary school.
Selection process	Screening by title/abstract and full reading; final selection of n = 54 documents. Complement by reference review (*snowballing*) and screening for relevance and applicability.
Role of the corpus in the study	Guidelines/policies are used for operational practices and recommendations; Claims about risks and effects are preferably supported by published reviews and studies.
Stated methodological limitation	Systematized scope search (*scoping review*), aimed at mapping trends and uses in the period analyzed. No formal evaluation of methodological quality or meta-analysis was performed, so the results should be interpreted as a mapping of the field and as a basis for the design of the matrix, not as evidence of effects.

3.2 CODING AND CATEGORISATION PROCEDURE

The procedure was carried out in four steps. First, an inventory of teaching tasks and learning tasks typical of the teaching of Geography in Secondary Education was developed, organized by macroprocess (planning, design/adaptation of resources, instruction and evaluation/feedback). Second, categories of analysis were defined to estimate the "degree of support" and the "risk of substitution" based on explicit criteria (expected precision, need for verification, dependence on the context, requirement of professional judgment and potential for delegation of reasoning). Third, each task was classified on the A/M/B/N operational scale and contrasted with integration frameworks (EIADD) to ensure coherence between intervention level and didactic design. Fourth, the classifications were reviewed looking for inconsistencies and adjusting definitions until a unified matrix with operational observations was obtained for use in centers and departments.

To make the level of AI intervention comparable between tasks, the AI Integration and Didactic Design Scale (EIADD) is adopted, which allows the intervention to be classified from non-AI scenarios to full delegation scenarios, including intermediate levels of assistance and collaboration (Caldeiro & Odetti, 2024). In addition, a taxonomy of educational uses of AI is used that distinguishes purposes such as creating resources, personalizing, correcting (in a formative key) and making materials and activities accessible, which facilitates the analysis of the didactic function pursued in each task (Ezcurra, Corica & Micco, 2025). In addition,

categories aimed at controlling didactic risk are incorporated: practices to avoid (cognitive substitution), practices to be reviewed (for example, tasks vulnerable to plagiarism or the opaque use of AI) and practices to be promoted (when they favour critical thinking and creativity) (García & Pujol, 2025). In the design and analysis of interactions, *structured prompts* are used using models such as R-I-T-A (Role, Information, Task, Adaptability) to standardize the degree of specification and facilitate comparison between results (Guallart, 2025a).

To increase consistency, the categorization was reviewed in two rounds: a first coding and a subsequent review of discrepancies (analogous tasks with different levels). When doubts arose, operational definitions were reformulated and borderline cases were re-classified until internal coherence between 'degree of support' and 'risk of substitution' was achieved. A change log was retained to ensure traceability.

3.3 CRITERIA FOR "DEGREE OF SUPPORT" AND SCALE

The "degree of support" is defined as the practical and pedagogical utility of the output of the IAG for a particular task, considering both the quality of the outcome and the associated risks and the need for supervision. To estimate it, convergent criteria are applied: accuracy and reliability of the response, time savings for teachers, need for human supervision, real autonomy of the tool and risk of undesired effects (e.g., systematic errors or overconfidence in plausible answers) (Faculty AI, 2024). In addition, two pedagogical adequacy filters from the Effective Integration Matrix are incorporated: whether the task favours effective learning or is limited to automation, and whether the presence of IAG supports the process or replaces the students' reasoning; in this framework, the "golden rule" of use is prioritized: asking AI about what is already known, so that students validate and expand on previous knowledge and do not delegate to unknown areas (García & Pujol, 2025).

Based on these criteria, an operational scale of support is established that integrates equivalent formulations of the different blocks and maintains a gradation compatible with other scales of use (such as AIAS) (Furze, 2024). In practical terms, four levels are distinguished. High support (A) corresponds to tasks in which the AI produces useful drafts or proposals with minimal editing and low uncertainty (e.g., brainstorming or basic structures), always with teacher control. Medium support (M) is assigned when the AI provides structures or alternatives that require significant contextual adaptation through prompt iteration and teacher review (e.g., planning or adapting materials to the group). Low support (B) is reserved for tasks where AI shows frequent errors or variability that forces thorough verification (e.g., specific statistical data or high-accuracy claims) (Cornell University, 2023). Finally, null

support (N) is attributed to tasks that are not suitable for delegation in AI due to their material, ethical, or relational nature, such as physical fieldwork or dimensions that require deep and situated empathy, as well as final evaluative decisions without supervision (Ali Elkot et al., 2025; Faculty AI, 2024; Videla, 2025d).

3.4 ETHICAL AND PRIVACY CONSIDERATIONS

The research is aligned with principles of data protection and minimization of risk in the use of chatbots with minors, avoiding introducing students' personal information in the interactions analyzed and in any example of *prompts* (Ezcurra, Corica & Micco, 2025; U.S. Department of Education, 2023). It is warned that the use of free tools or personal accounts may involve reuse of data entered for training or other purposes, which is incompatible with the protection of sensitive student data and privacy requirements in school contexts (Holdsworth, 2023).

From a pedagogical governance point of view, the "*human-in-the-loop*" principle is adopted: AI does not make final decisions on assessment or learning trajectory without teacher supervision, and all relevant outputs are validated to mitigate biases and hallucinations (Amiri et al., 2025; U.S. Department of Education, 2023; Fonteneau & Anvar, 2025). Likewise, an AI literacy perspective is incorporated that includes understanding of bias and intellectual property, in order to sustain academic integrity and avoid opaque or uncritical use in school assignments (Giannini, 2024; U.S. Department of Education, 2023; Wilby & Esson, 2023).

This study has limitations derived from its documentary nature and applied design: the proposed matrix is not validated here by classroom intervention, nor is its impact on learning or academic integrity measured with student or teacher data. Consequently, the support/risk classifications should be interpreted as a reasoned guide for didactic design, susceptible to adjustment according to context, the school's evaluative culture, access to TIG and students' level of AI literacy.

Since the performance of the models changes with updates, the support/risk assessments are formulated as dependent on the state of the tool in the period analyzed. In real applications, it is recommended to revalidate critical tasks (e.g., factual accuracy, evaluation) when version or institutional configuration changes occur.

4 RESULTS

4.1 TAXONOMY OF TASKS IN GEOGRAPHY IN SECONDARY EDUCATION

The analysis allows to delimit a set of representative tasks of the teaching of Geography in Secondary Education, focused on the understanding of territorial and socioeconomic processes, the analysis of socio-environmental problems and the interpretation of phenomena at different scales. These tasks require students to observe, analyse, interpret, assess and infer from territorial and social information, with different degrees of cognitive complexity, incorporating cartographic literacy and, when possible, the use of TIG to support spatial reasoning.

To specify the variability of cognitive demands, the classification is triangulated with an adaptation of Bloom's taxonomy and the DOK (*Depth of Knowledge*) model to the context of AI. In tasks of remembering and understanding, definitions of physical elements and demographic data are located; in applying and analyzing, comparisons of population pyramids or analysis of the causes of geopolitical conflicts; and in evaluating and creating, designs of territorial planning proposals or argued debates on sustainability. Overall, the resulting taxonomy describes a repertoire that combines factual tasks, analytical tasks with the need to contrast sources, and territorial judgment tasks with an ethical and citizen component (Area, 2025; Correa, López & Díaz, 2024).

4.2 CHATGPT SUPPORT MATRIX: PATTERNS BY MACROPROCESS

The results show a consistent pattern: ChatGPT adds value as a generator of initial ideas, creator of rubrics and producer of administrative or communicative texts, with a generally high performance in verbal structuring tasks. In planning, support focuses on the sequencing of knowledge, the proposal of lesson structures, the diversification of resources and the ideation of projects; however, practical usefulness depends on teachers providing specific context and writing *detailed prompts, since direct usability decreases when working with generic instructions and without information from the specific group or curriculum (Faculty AI, 2024; U.S. Department of Education, 2023).*

In the production of resources and materials, support tends to be high: ChatGPT facilitates the adaptation of texts at different reading levels, the simplification of language, the generation of review questions, and the creation of case studies and scenarios (including fictional ones) that serve as the basis for teaching design (Prud'homme-Généreux, 2024; Wilby & Esson, 2023). In direct instruction, the support appears as medium or medium-low, because it can explain concepts and sustain continuous personalized tutoring, but the risk of hallucination requires systematic verification and use aimed at contrasting, not replacing,

study and comprehension (Osamor et al., 2023; Salman et al., 2025; Langreo, 2025). In assessment, the difference between formative and summative assessment is decisive: for formative feedback on drafts (grammatical, structural improvements and initial suggestions), the support ranges from medium to high; on the other hand, for summative grading of complex trials or products that require contextual nuance and professional judgment, support is at low levels and at risk, due to inconsistencies and possible biases (Apata & Oladapo, 2025; Salman et al., 2025; Yetişensoy & Karaduman, 2024).

In specifically geographical tasks, performance is high in conceptual clarification and in simulation of territorial debates, while it shows limitations when the task requires precise visual mapping or advanced spatial treatment without complementary tools. This result reinforces the idea that the value of ChatGPT is concentrated in the linguistic-cognitive, and that its integration into Geography needs to be supported by TIG when the evidence of learning requires cartographic representation or spatial analysis with georeferenced data (De la Torre, 2024; Tao & Xu, 2023).

4.3 NON-RECOMMENDED TASKS AND HIGH-RETURN TASKS WITH TEACHER CONTROL

In the set of tasks analysed, two problematic areas stand out. The first is the vulnerability to cognitive substitution in descriptive essays and generic essays for home, which AI can complete from start to finish without leaving a trace of the learning process; in these cases, the final product ceases to be a valid indicator of understanding (García & Pujol, 2025). The second is the automated summative assessment of complex work (including essays or products associated with fieldwork), where lack of precision, inconsistency, and potential bias advise against its use to assign final grades, which must be maintained under professional teaching judgment (Cornell University, 2023; Fonteneau & Anvar, 2025; Salman et al., 2025).

In contrast, high-return tasks are identified when teacher control directs use towards verification, contrast, and argumentation. These include the generation of practice questionnaires and question banks (with revision to avoid errors), linguistic adaptation to attend to diversity, and brainstorming for field trips, which free up teachers' time without displacing the cognitive core of the subject (Cornell University, 2023; Faculty AI, 2024). Tasks designed for students to critically analyse AI responses, identify biases or errors and contrast them with reliable sources, as well as the use of AI to generate opposing arguments in debates on geographical controversies, so that students evaluate evidence and refine their position (Guallart, 2025a; Prud'homme-Généreux, 2024). Finally, two frequent uses are

considered not recommended due to their specific risk: requesting "real" academic bibliography without verification, due to the possibility of inventing sources, and requesting the direct creation of visual maps without intermediation (code or GIS tools), due to technical limitations and the risk of producing incorrect representations (Tao & Xu, 2023; Wilby & Esson, 2023).

A synthetic version of the matrix (representative tasks by dimension) is presented; the extended version with prompts and traceability evidence is offered as Annex II (excerpt).

Table 2

Unified ChatGPT Support and Risk Matrix in Geography Tasks in Secondary Education

Dimension	**Key task (representative)**	**ChatGPT's role**	**G**	**R**	**Condition of use/safeguard (and evidence)**
Planning	Knowledge sequencing and session design	Structurer (Alternatives and Connections)	A–M	B	It requires curricular and group context; Final teaching decision. (P+B)
Planning	Itineraries / field trips (logistics and *checklist*)	Ideation + checklist	M	M	Verify coordinates/times/accessibility; it does not replace territorial recognition or security. (F+V)
Design/adaptation of resources	Adaptation of texts and instructions (easy reading, levels)	Editor/Adapter	To	M	Mandatory review for bias/errors; Pin to source text. (P+V)
Design/adaptation of resources	Summary of reports/news for critical reading	Structured Summarizer + Critical Questions	A–M	M	Require short appointments/location and check "VERIFY" on sensitive data. (F+V)
Instruction	Territorial debate and dilemmas (multi-perspective)	Socratic mediator + counterarguments	To	B	Forcing evidence and comparison of sources; Avoid caricatures of actors. (F+D)
Geographical task	Data/map/graph analysis (patterns and hypotheses)	Methodological assistant/verifier	M	M	Triangulation with data and sources; record checks and confidence level. (F+V)

Dimension	Key task (representative)	ChatGPT's role	G	R	Condition of use/safeguard (and evidence)
Geographical task	Spatial planning/urban solutions *(trade-offs*)	Alternative Generator + Impact Matrix	M	M	Demand assumptions, data and justification; avoid "nice solutions" without viability. (F+B)
TIG/SIG	Support for ArcGIS Online or others (steps, queries, debugging)	Procedural Assistant	M	B	It does not replace GIS; test in a real environment and apply cartographic criteria. (F+V)
Evaluation/feedback	Rubric for formative feedback (without final grade)	Justification by criteria + review questions	M	M	Not assigning final grade; use as support and contrast with evidence from the student. (B+D)
Administration and communication	Family Communications/Instructions/Minutes	Draft Writer	To	M	Privacy and tone control; not to enter personal data; Final review. (P+V)
Not recommended	Generic descriptive assay for home (no process)	Generate full text	A–B	To	Redesign: outline + drafts + oral defense/application. (B+D)
Not recommended	"Real" bibliography without verification	Generate referrals	B	To	Replace with search and verification strategy (DOI/ISBN). (F+V)
Not recommended	Direct visual mapping as a substitute for GIS	"Substitute" for GIS	N	To	Ask for flow in GIS (layers/symbology/projection) and quality criteria, not "image". (F+V)

The authors.

Note. G = estimated degree of support (A = high; M = medium; B = low); R = risk of substitution (A = high; M = medium; B = low). In "evidence" abbreviations are used: P = prompt; B = drafts/versions; F = sources/data; V = verification/contrast; D = oral defense/argumentation.

5 DIDACTIC IMPLICATIONS AND IMPLEMENTATION EXAMPLES

5.1 IMPLEMENTATION EXAMPLES: LEARNING SITUATIONS (SDA)

The responsible integration of ChatGPT in Geography can be specified, by way of **example**, in a set of learning situations (SdA) that can be adapted to Secondary Education that illustrate how to incorporate the tool without replacing the geographical reasoning of the students. In all cases, it is proposed that the final product requires decision-making and

justification with evidence, and that learning evidence includes traceability of the process (history of prompts, intermediate drafts, and critical reflection on the use of AI), in order to preserve authorship and academic integrity (Area, 2025; García & Pujol, 2025; Loos, 2025).

A first SdA is based on a field trip on the island of Tenerife (or an analogous local context) aimed at the analysis of sustainability and economic activities. The final product combines a digital poster and a field report focused on waste management and the energy transition, with evidence distributed in an initial questionnaire, field notebook and oral presentation. This option prioritizes direct observation and the link between classroom and territory as a didactic condition of school geography (Correa, López, & Díaz, 2024).

The second SdA, called "The Demographic Consultant", is located in Human Geography and addresses demographic and migratory challenges through a fictitious country with marked aging. The final product takes the form of a report of public policy recommendations, where students must justify decisions and anticipate medium-term effects. Evidence includes the history of interaction with AI, successive drafts, and a reasoned reflection on the help received and discarded (García & Pujol, 2025; Loos, 2025).

The third SdA, "Demographic dynamics and territorial development: depopulation in Spain", shifts the focus from the essay to a strategic proposal for territorial revitalization supported by data. The didactic itinerary incorporates conceptual clarification, analysis with digital maps and official sources (for example, INE) and a propositional closing in which the students defend measures and assess their limitations. In this SdA, AI supports methodological understanding and argument construction, but does not replace visual map reading or working with sources (Guallart, 2025a).

The fourth SdA focuses on "Socio-economic inequalities and migration" and combines a mock debate with a subsequent report that looks at causes and consequences. It is linked to learning objectives aligned with SDG 10 (Reduced inequalities) and requires teachers to select a clear territorial cut to avoid generalizations and to support spatial analysis with evidence. The evidence incorporates the registration of *prompts*, the validation of demographic data, and critical reflection on the reliability of the information generated by AI (Adamini, Cafferata, & Velo, 2024; Area, 2025).

5.2 IMPLEMENTATION GUIDELINES: PHASED WORKFLOW (TEACHERS AND CHATGPT)

Based on the support and risk matrix, a phased workflow is proposed that clearly separates pedagogical decisions (non-delegable), linguistic-cognitive supports (delegable with control) and moments of direct spatial experience (not mediatized by AI). In the design

phase, the teachers define objectives, select the territorial area, plan the sequence of activities and decide on the final product and process evidence; at this level, ChatGPT can support the generation of didactic scenarios (e.g., fictitious countries or conflicts), the diversification of resources, and the initial development of rubrics, always under teacher review and contextualization (Adamini, Cafferata & Velo, 2024; Wilby & Esson, 2023).

During the territorial work phase (exit or direct observation work), **it is** recommended that AI does not intervene directly, in order to favor spatial perception and in situ evidence collection; this decision is linked to the need to sustain spatial experience as the basis of geographical learning (Correa, López & Díaz, 2024).

In the synthesis and production phase, students can use ChatGPT to organize memory, improve writing or prepare a presentation, as long as they declare the use and do not delegate the central argument or interpretative decisions. In parallel, teachers can use AI to generate a poster or report rubric and to prepare formative assessment instruments, maintaining professional judgement in the final assessment (Correa, López & Díaz, 2024). To increase the pedagogical return, it is suggested to incorporate interactions with AI configured as "Socratic opponent", "skeptical demographer" or "territorial planner", so that students defend proposals, anticipate objections and refine their arguments with an empirical basis, instead of accepting answers as an authority (García & Pujol, 2025; Prud'homme-Généreux, 2024; Urmeneta & Romero, 2024).

5.3 OPERATIONAL RESOURCES: *PROMPTS*, TEMPLATES AND QUALITY CONTROLS

Prompts formulated as reusable templates *are proposed* , accompanied by quality controls aimed at precision, curricular alignment and traceability of the process.

- ✓ In preparation teaching tasks, an example of a prompt to generate an initial diagnostic instrument is: "*Act like a Geography teacher. Generate 5 multiple-choice questions to assess understanding of island waste management, adapted to a student aged [x] years. Include justification of the correct answer and mark with [VERIFY] any factual data*."
- ✓ In the design of case studies for debate, it is proposed: "*Acts as an expert in geography didactics. It creates a case study of a water conflict between two fictional regions, including conflicting climate and economic data to encourage discussion*."
- ✓ For the complete design of an SdA with evaluation criteria, a structured *prompt* (R-I-T-A model) can be used: "*You are a geography teacher... Design a learning situation about depopulation... including a rubric to assess critical analysis*".

- ✓ In role-playing games, ChatGPT can generate profiles of actors (e.g., migrant, local politician, business, NGO) as a starting point for the debate: "*It acts as an expert in geography didactics. Generate 4 short and contrasting profiles for a role-playing game about the installation of a wind farm in a rural municipality with depopulation: newly arrived migrant, local politician, businessman and environmental NGO. For each profile, include objectives, 3 arguments and 1 possible concession. Mark with [VERIFY] any factual data or cipher*"

In the students' work, it is recommended to prioritize *prompts* that prevent direct response and force verification, revision or argumentation.

- ✓ For controlled tutoring (study mode), it is proposed: "*Help me understand what the energy transition is, ask me questions to see if I understand it, don't give me the direct answer*".
- ✓ To activate critical thinking, *adversarial* prompts are used: "*Act like a skeptical demographer. He criticizes my proposal to increase the birth rate through subsidies, pointing out three possible long-term economic failures*", or "*He acts as an expert in demography who questions my proposal for repopulation based on rural tourism. It points out three economic weaknesses of my plan*."
- ✓ In content sensitive to bias, an explicit control of the framing is incorporated: "*It acts like a geographer expert in demography. It explains the causes of current migrations in the Mediterranean for secondary school students, avoiding Eurocentric biases. Includes 2 nuances/counterexamples and 3 verification questions*."

Quality controls are recommended as assessable requirements. At the factual and curricular level, teachers verify that the questions and statements do not contain errors and conform to the curriculum, given the possibility of plausible but incorrect answers (Cornell University, 2023; Faculty AI, 2024). In terms of authorship, students are required to submit the chat history and a reflection on what information was useful and what was discarded as inaccurate; in addition, it is requested to explain why it accepted or rejected suggestions from the AI, since the inability to justify decisions works as an indicator of dependency and lack of authorship. On the methodological level, triangulation is required: no data provided by AI is considered valid without contrast with an official or equivalent source (e.g., INE, World Bank) (Area, 2025; Loos, 2025). Finally, in terms of transparency, the students declare the use of AI in the final product when it has been used to structure or revise the writing, clearly differentiating linguistic support for the construction of the argument.

6 RISKS, LIMITS AND SAFEGUARDS

The risks associated with AGI (plausible errors, biases and privacy) only acquire didactic meaning when they are translated into **task decisions** and **enforceable evidence**. Therefore, in this article, the risks are not presented as a general inventory, but as criteria that explain the classification of the matrix: tasks that are "not recommended" because of high delegability (e.g., generic essay without process or bibliography without verification) and tasks that are "supportable" when verifiable phases are required (drafts, contrast with sources, and defense). Safeguards are formulated as assessable process requirements (sources, verification, interaction history, and justification of decisions), so that support is not a substitute for geographical reasoning.

A second risk is algorithmic bias. The models can reproduce biases present in their training data and project them onto territories and populations, which introduces interpretive distortions and equity issues (Gallent-Torres et al., 2024; Holdsworth, 2023; Lorenz, Perset & Berryhill, 2023). In Geography, this bias can take specific forms, such as cultural stereotypes about the "Global South" or an overrepresentation of perspectives from the "Global North", with the consequent impoverishment of approaches and the marginalization of underrepresented realities (Furze, 2025; Giannini, 2024; Lorenz, Perset & Berryhill, 2023; Wilby & Esson, 2023). For this reason, bias should not be treated as an external technical problem, but as content to be worked on explicitly in the classroom, linking it to multiperspective and critical analysis of sources.

The third limit concerns privacy and data protection. Students' personal data or work with sensitive information should not be entered on IAG's public platforms, especially because they are minors and due to the lack of control over the subsequent processing of what is entered (Gallent-Torres et al., 2024; Giannini, 2024; U.S. Department of Education, 2023). In line with this principle, it is proposed to reduce risk through tools that do not require personal data and, where available at the institution, through accounts managed by the school or the administration (Ezcurra, Corica & Micco, 2025).

In terms of academic integrity and assessment, an additional problem appears: the automatic detection of AI-generated text is unreliable, described as a "mission impossible" when the aim is to infer authorship only from the final product (Loos, 2025). Consequently, the most solid safeguard shifts the emphasis from the final text to the evaluation of the process: intermediate evidence, recording of decisions, contrast of sources and oral defense of arguments and data, so that students demonstrate understanding and control of reasoning (García & Pujol, 2025). Along the same lines, it is recommended to incorporate transparency

scales for students to declare the level of AI used, facilitating an adjusted and explicit evaluation of the type of aid received (Caldeiro & Odetti, 2024).

The operational safeguards are articulated in three complementary measures. First, institutionalize verification: teach to "re-question", demand contrast with reliable sources and, when possible, with primary sources, incorporating contrast as an assessable part of the task (Gallent-Torres et al., 2024; Office of Science and Technology of the Congress of Deputies, 2024). Second, ensure traceability: students document *prompts*, relevant responses, acceptance or rejection decisions, and the verification process, so that the final product is anchored in evidence of authorship and learning (Area, 2025; García & Pujol, 2025). Third, maintaining teacher supervision when AI intervenes in activities related to evaluation or feedback, preventing content generated without control from conditioning grades or educational decisions (Fonteneau & Anvar, 2025).

7 DISCUSSION

In the light of the proposed matrix, the central input is not 'if' to use IAG, but 'where' and 'under what conditions' its use preserves geographical learning. The discussion suggests that the differential value of ChatGPT is concentrated in linguistic-cognitive and design supports, while tasks that demand factual accuracy, situated evaluative judgment, and advanced spatial reasoning require teacher mediation and, frequently, TIG/GIS tools. Consequently, responsible integration depends on shifting evaluation towards process evidence (verification, traceability and defense), rather than towards easily delegated end products.

This change of role has direct implications for the departments of Geography and for the educational centers. The literature converges on the need for coherence and standardization of criteria: it is not feasible to sustain contradictory approaches within the same school, where one part of the teaching staff totally prohibits AI while another demands or encourages it without a common framework (Vallejo, 2024). Hence, school policies are required that define acceptable uses, conditions for citation and declaration of the use of AI, as well as minimum verification and data protection procedures, so that students receive clear and comparable expectations between subjects and groups (Gallent-Torres et al., 2024; Guallart, 2025c). This standardization is also linked to the reduction of digital and pedagogical gaps: without shared frameworks, the integration of AI can amplify inequalities according to the level of teaching competence, the evaluation culture of the school or access to resources. In this sense, it is proposed to rely on reference frameworks such as EurekAI or DigCompEdu

to align practices, train teachers and provide consistency to integration (García & Pujol, 2025).

On the didactic level, the results reinforce the idea that evaluation must move from the final product to the process. If AI can generate complete essays, the learning indicator ceases to be the text delivered and becomes the quality of the inquiry process: formulation of questions, selection and contrast of sources, decision-making, argumentation and oral defense. This approach opens a specific line in Geography: the question as evidence of learning, understood not as a formal feature, but as an ability to delimit a territorial problem, guide the search for data and construct a situated explanation. At the same time, the standardization of effective *prompts* is configured as an emerging teaching competence, not as an accessory resource, since it conditions the quality of interactions and, by extension, the quality of learning processes mediated by AI (Area, 2025; García & Pujol, 2025). This requirement is connected to AI literacy for both teachers and students: understanding the tool not as an oracle, but as a reasoning engine that requires expert supervision and validation strategies, with special attention to the risk of metacognitive laziness (Vallejo, 2024; Xiaoyu et al., 2025).

Finally, Geography introduces a disciplinary counterpoint that is particularly relevant in the face of the virtuality of AI. Due to its focus on territory and physical reality, the subject reaffirms the value of direct contact with the environment (e.g., fieldwork, in situ observation, reading the landscape) as a dimension that cannot be replaced by generative systems (Correa, López & Díaz, 2024). This specificity suggests a practical orientation for integration: to assess by application to local contexts and through evidence that combines spatial experience, contrasted data and argumentation, including situations in which AI lacks situated knowledge and, therefore, cannot replace the geographical criterion constructed by students (Area, 2025).

In terms of transferability, the matrix is especially useful as a departmental working tool to agree on minimum criteria (verification, traceability and process evidence), but its application must be adapted to specific regulatory and curricular contexts and technological availability (e.g., access to TIGs and institutional accounts). Future research should validate the instrument through classroom implementation studies (quasi-experimental designs or action-research), comparing tasks and process evidence with and without the mediation of AGI, and incorporating quality analyses of spatial reasoning and source contrast practices.

8 CONCLUSIONS

Geography teaching in Secondary Education, understood as a system of tasks, can benefit significantly from generative artificial intelligence (AGI) when integrated with a critical approach and teacher supervision. In this framework, tools such as ChatGPT are useful, above all, to support the personalization of learning, improve efficiency in the preparation of resources and facilitate scaffolding for writing and structuring activities, freeing up time for teachers to concentrate their intervention on pedagogical interaction and student accompaniment (Adiguzel et al., 2023; López-Vasco et al., 2025; Sharma & Rajarajan Eds, 2025). However, these benefits depend on the type of task, the degree of AI literacy and the existence of verification protocols; Without these conditions, the tool can increase the burden (revisions, correction of errors) and amplify inequalities between groups with different access or accompaniment.

The results also clearly delineate the limits of these tools. ChatGPT can fail as a substitute for factual accuracy, ethical judgment, and pedagogical criteria, which discourages its autonomous use in summative evaluation and in tasks that require deep territorial analysis without human guidance (Giannini, 2024; Salman et al., 2025; Wirzal et al., 2024). In Geography, moreover, it should not replace direct observation of the territory or spatial reasoning, which constitute the formative core of spatial competence; its role is justified when it acts as a scaffolding to manage complexity, contrast arguments, or organize evidence, not when it replaces spatial experience and work with sources (Guallart, 2025b).

In this way, the future of geographical education with AI does not lie in the automation of learning, but in the design of tasks that AI cannot solve on its own without critical student intervention. This orientation makes it possible to avoid a "techno-banking education", in which technology reinforces logics of transmission or substitution of cognitive effort, and favours, instead, a hybrid model where AI amplifies geographical thinking without replacing it (Adamini, Cafferata & Velo, 2024; Ubal et al., 2023; Vallejo et al., 2024). In terms of curriculum and evaluation, this implies prioritizing process evidence, tasks of application to local contexts, and argumentation and defense activities that force data to be contrasted, decisions justified, and territorial interpretations sustained.

The viability of responsible integration depends on the pedagogical leadership of teachers and on school agreements that regulate verification, traceability and privacy, but also on recognizing the limits of this work: the matrix is derived from documentary analysis and applied categorization, without empirical validation in the classroom, so its levels of support/risk should be considered indicative and reviewable. As future lines, it is proposed to evaluate the matrix through real implementations in Secondary Education classrooms that

measure (a) quality of spatial thinking and the use of sources, (b) changes in processes assessment and oral defense practices, and (c) effects on academic integrity and equity, especially in centers with different technological availability.

REFERENCES

Adamini, P., Cafferata, E., & Velo, S. (2024). "Inteligencia Artificial" tejiendo puentes con la didáctica de la geografía. En Actas del Congreso. ISFDN 54 “Olga Cossetini”.

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. Contemporary Educational Technology, 15(3), ep429. https://doi.org/10.30935/cedtech/13152

Ali Elkot, M., Alhalangy, A., AbdAlgane, M., & Ali, R. (2025). Pedagogical influence of AI-chatbots on learning outcomes: A systematic review. International Journal of Educational Methodology, 11(4), 527–540. https://doi.org/10.12973/ijem.11.4.527

Amiri, S. M. H., Islam, M. M., & Akter, N. (2025). Educating the future: AI’s role in shaping next-generation pedagogies. International Journal of Science and Research Archive, 15(1), 621–636. https://doi.org/10.30574/ijsra.2025.15.1.1036

Apata, S. B., & Oladapo, S. O. (2025). The transition from traditional teaching methods in universities to the integration of digital tools [Manuscrito no publicado]. Adekunle Ajasin University.

Area, M. (2025). Luces y sombras de la IA en la educación superior: Didáctica para el pensamiento crítico. Universidad de La Laguna. http://riull.ull.es/xmlui/handle/915/40470

Bernal, M. D. C., Bunhayag, G. A., Loyola, D. S. D., Tisado, J. C., & Genelza, G. G. (2025). Addressing the elephant in the room: The impact of using artificial intelligence in education. International Journal of Human Research and Social Science Studies, 2(4), 178–200. https://doi.org/10.55677/ijhrsss/08-2025-Vol02I4

Bobula, M. (2024). Generative Artificial Intelligence (AI) in higher education: A comprehensive review of challenges, opportunities, and implications. Journal of Learning Development in Higher Education, (30). https://doi.org/10.47408/jldhe.vi30

Caldeiro, G., & Odetti, V. (2024). La transparencia como factor crítico en los usos educativos de la IA. Virtualidad, Educación y Ciencia, 28(15), 46–53.

Cornell University. (2023). Generative Artificial Intelligence for education and pedagogy. Cornell University Center for Teaching Innovation. https://teaching.cornell.edu/generative-artificial-intelligence/cu-committee-report-generative-artificial-intelligence-education

Correa González, J., López Díez, A., & Salvador Díaz Pacheco, J. (2024). Enseñar geografía en el siglo XXI: Del aula al territorio [Teaching geography in the 21st century: From classroom to territory]. Espacio, Tiempo y Forma. Serie VI, Geografía, 17, 123–143. https://doi.org/10.5944/etfvi.17.2024.39779

Cronje, J., Swartz, B., & Valentine, L. (2023). Assessing the potential of chatbots as tutoring tools in engineering education: A South African case study [Manuscrito presentado para publicación]. Cape Peninsula University of Technology.

Cruz Argudo, F., García Varea, I., Ruiz Martínez, P. M., & Sánchez Campos, A. (2024). La inteligencia artificial generativa en la docencia universitaria: Oportunidades, desafíos y recomendaciones. CRUE Universidades Españolas. https://raeia.org/books/la-inteligencia-artificial-generativa-en-la-docencia-universitaria-oportunidades-desafios-y-recomendaciones/

De la Torre. (2024, 15 de julio). Creación de un generador de prompts: Empoderando la comunicación con IA. Delatorre.ai. https://delatorre.ai/creacion-de-un-generador-de-prompts-empoderando-la-comunicacion-con-ia/

Department for Education. (2025, 12 de agosto). Generative artificial intelligence (AI) in education [Policy paper]. GOV.UK. https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education

Ezcurra, V., Corica, C. & Micco, F. (2025). IA en la escuela: Guía para un uso crítico (Versión 2.0). Ministerio de Educación de la Ciudad Autónoma de Buenos Aires. https://www.studocu.com/es-ar/document/isfd-alberto-g-cavero/psicologia-educacional/ia-en-la-escuela-guia-critica-para-su-uso-educativo-version-20/136237855

Faculty AI. (2024). Use cases for generative AI in education: User research report. Department for Education (UK). https://assets.publishing.service.gov.uk/media/66cdb078f04c14b05511b322/Use_cases_for_generative_AI_in_education_user_research_report.pdf

Fonteneau, F., & Anvar, E. (2025, 2 de mayo). Exploring the safe and responsible use of generative AI across the data lifecycle: Early insights from the OECD's adaptive approach. OECD. https://unece.org/sites/default/files/2025-05/GenAI2025_S2_OECD_FONTENEAU_D.pdf

Furze, L. (2024, 28 de agosto). Updating the AI Assessment Scale. https://leonfurze.com/2024/08/28/updating-the-ai-assessment-scale/

Furze, L. (2025, 31 de julio). OpenAI has come for education. https://leonfurze.com/2025/07/31/openai-has-come-for-education/

Furze, L. (s. f.). Teaching AI ethics: Practical strategies for discussing AI ethics in K–12 and tertiary education [Ebook]. https://socialsci.libretexts.org/Bookshelves/Education_and_Professional_Development/Teaching_AI_Ethics%3A_Practical_Strategies_for_Discussing_AI_Ethics_in_K-12_and_Tertiary_Education_(Furze)

Gallent-Torres, C., Arenas Romero, B., Vallespir Adillón, M., & Foltýnek, T. (2024). Inteligencia Artificial en educación: Entre riesgos y potencialidades. Práxis Educativa, 19, e23760, 1–29. https://doi.org/10.5212/PraxEduc.v.19.23760.083

García Martínez, A. N., & Pujol, F. (2025). EurekAI: Modelo de integración de la IA en la docencia universitaria (Versión 1.1). Universidad de Navarra. https://doi.org/10.15581/028.00016

Giannini, S. (2024). Guía para el uso de IA generativa en educación e investigación. UNESCO. https://www.unesco.org/es/articles/guia-para-el-uso-de-ia-generativa-en-educacion-e-investigacion

Guallart Moreno, C. (2025a). Educar para pensar el territorio: pensamiento crítico y chatbots en la enseñanza de la geografía. Editorial Académica Española.

Guallart Moreno, C. (2025b). Competencia espacial, pensamiento geográfico y ciudadanía crítica: Un enfoque integrado en educación secundaria. Espacio, Tiempo y Forma. Serie VI, Geografía, 18, 151–180. https://doi.org/10.5944/etfvi.18.2025.45337

Guallart Moreno, C. (2025c). Necesidad de un protocolo institucional para la integración de la inteligencia artificial generativa en los centros educativos de primaria y secundaria. Revista DCS, 22(82), 1–17. https://doi.org/10.54899/dcs.v22i82.3396

Holdsworth, J. (2023, 22 de diciembre). ¿Qué es el sesgo de la IA? IBM. https://www.ibm.com/es-es/think/topics/ai-bias

Langreo, L. (2025, 29 de julio). What teachers should know about ChatGPT's new study mode feature. Education Week. https://www.edweek.org/technology/what-teachers-should-know-about-chatgpts-new-study-mode-feature/2025/07

Loos, E. (2025, January 31). Responsible chatbot use in higher education: Implications for exam boards: A narrative literature review (USO Project AI in Higher Education, Work package 4: AI literacy as a future skill) [Report]. Utrecht University School of Governance. https://www.uu.nl/staff/EJPLoos

López-Vasco, F. E., Angulo-Álvarez, M. R., & Sosa-Zúñiga, D. I. (2025). Formación docente en IA generativa: Impacto ético y retos en educación superior [Teacher training in Generative AI: Ethical impact and challenges in Higher Education]. Alteridad, Revista de Educación, 20(2), 166–177. https://doi.org/10.17163/alt.v20n2.2025.01

Lorenz, P., Perset, K., & Berryhill, J. (2023). Initial policy considerations for generative artificial intelligence (OECD Artificial Intelligence Papers, No. 1). OECD Publishing. https://doi.org/10.1787/fae2d1e6-en

Oficina de Ciencia y Tecnología del Congreso de los Diputados. (2024). Informe C: Inteligencia artificial y educación. https://oficinac.es/es/informes-c/inteligencia-artificial-y-educacion

Organisation for Economic Co-operation and Development. (2023, 15 de julio). Putting AI to the test: How does the performance of GPT and 15-year-old students in PISA compare? (OECD Education Spotlights, No. 6). https://doi.org/10.1787/2c297e0b-en

Osamor, A., Ifelebuegu, A., & Cherukut, P. (2023). Chatbots and AI in Education (AIEd) tools: The good, the bad, and the ugly. Journal of Applied Learning & Teaching, 6(2). https://doi.org/10.37074/jalt.2023.6.2.29

Pablo-Martí, F., Mir, C., & Santos, J. L. (2025). Introducción a la Inteligencia Artificial Generativa: Aplicaciones y Evaluación Crítica en Ciencias Sociales y Humanidades [Curso ISCE 2025-2026]. Departamento de Economía. Universidad Alcalá de Henares. https://www.researchgate.net/publication/396561612_Introduccion_a_la_Inteligencia_Artificial_Generativa_Aplicaciones_y_Evaluacion_Critica_en_Ciencias_Sociales_y_Humanidades

Prud'homme-Généreux, A. (2024, 3 de abril). Writing case studies using generative AI: Intimate debate case study. Faculty Focus. https://www.facultyfocus.com/articles/teaching-with-technology-articles/writing-case-studies-using-generative-ai-intimate-debate-case-study/

Salman, I. M., Ameer, O. Z., Khanfar, M. A., & Hsieh, Y.-H. (2025). Artificial intelligence in healthcare education: evaluating the accuracy of ChatGPT, Copilot, and Google Gemini in cardiovascular pharmacology. Frontiers in Medicine, 12, 1495378. https://doi.org/10.3389/fmed.2025.1495378

Sharma, M., & Rajarajan, A. K. (Eds.). (2025). Science Utsav 2025: Proceedings of online one-day teachers' conference on artificial intelligence in school education (Vashi, India, 8 de febrero de 2025). Navi Mumbai Science Foundation & Pillai College of Engineering. http://www.navimumbaisciencefoundation.org/common/pdf/SUPROC/NMSF_SUTC2025_Proceedings.pdf

Sutabri, T., Suroyo, H., & Kurniawan. (2025). Navigating the Generative AI Revolution in Education: A Systematic Review of Applications, Ethical Considerations, and Future Directions. Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI, 14(2), 374–384.

Tao, R., & Xu, J. (2023). Mapping with ChatGPT. ISPRS International Journal of Geo-Information, 12(7), 284. https://doi.org/10.3390/ijgi12070284

U.S. Department of Education, Office of Educational Technology. (2023, May). Artificial intelligence and the future of teaching and learning: Insights and recommendations. https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf

Ubal Camacho, M., Tambasco, P., & Martínez, S. (2023). El impacto de la Inteligencia Artificial en la educación. Riesgos y potencialidades de la IA en el aula. RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa, 15, 41–57. https://doi.org/10.6018/riite.584501

Urmeneta, A., & Romero, M. (Eds.). (2024). Creative Applications of Artificial Intelligence in Education. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-55272-4

Vallejo Echeverria, B. A., Basantes Arellano, L. C., Ardila Portocarrero, N. Y., Estupiñán Palomino, J. E., & Basantes Ortega, M. M. (2024). El papel de la inteligencia artificial en la formación del pensamiento crítico en las nuevas generaciones. Arandu UTIC. Revista Científica Internacional, 11(2). https://doi.org/10.69639/arandu.v11i2.391

Vallejo, A. E. (2024). La transformación del rol docente en la era de la Inteligencia Artificial: hacia un liderazgo pedagógico estratégico. Trayectorias Universitarias, 10(19), e165. https://doi.org/10.24215/24690090e165

Videla, F. (2025a). Diseña un asistente personal de inteligencia artificial: Caja de herramientas EdTech [Documento]. LinkedIn. https://www.linkedin.com/in/franco-videla-/

Videla, F. (2025b). Proyectos educativos con inteligencia artificial: Crear, pensar y diseñar experiencias para construir pensamiento crítico en contextos digitales [Guía práctica]. LinkedIn. https://www.linkedin.com/in/franco-videla-

Videla, F. (2025c). Innovación educativa con metodologías activas en tiempos de IA: Metodologías activas + inteligencia artificial [Documento]. LinkedIn. https://www.linkedin.com/in/franco-videla-/

Videla, F. (2025d). El desarrollo del pensamiento crítico en tiempos de IA: Pensamiento crítico + educación + inteligencia artificial [Documento]. LinkedIn. https://www.linkedin.com/in/franco-videla-/

Wilby, R. L., & Esson, J. (2023). AI literacy in geographic education and research: Capabilities, caveats, and criticality. The Geographical Journal, 190, e12548. https://doi.org/10.1111/geoj.12548

Wirzal, M. D. H., Md Nordin, N. A. H., Abd Halim, N. S., & Bustam, M. A. (2024). Generative AI in Science Education: A Learning Revolution or a Threat to Academic Integrity? A Bibliometric Analysis. Jurnal Penelitian dan Pengkajian Ilmu Pendidikan: e-Saintika, 8(3), 319–351. https://doi.org/10.36312/e-saintika.v8i3.2127

Xiaoyu, W., Zainuddin, Z., & Leng, C. H. (2025). Generative artificial intelligence in pedagogical practices: a systematic review of empirical studies (2022–2024). Cogent Education, 12(1), 2485499. https://doi.org/10.1080/2331186X.2025.2485499

Yetişensoy, O., & Karaduman, Hıdır. (2024). The effect of AI-powered chatbots in social studies education. Education and Information Technologies. Advance online publication. https://doi.org/10.1007/s10639-024-12485-6

APPENDIX

The table organizes common teacher tasks and describes the role of ChatGPT as a support. The level of support (G) reflects potential usefulness (High, Medium, Low, Null) with teacher supervision; the risk of substitution (R) estimates the probability (high, medium, low) that the use will shift towards the substitution of the intellectual work of the students or of the teacher's judgment.

Table 1

Matrix of teaching tasks and ChatGPT support (Geography 3rd year of Secondary)

Dimension	Task (examples)	ChatGPT's contribution/role	G	R	Operational Observations	Example of prompt	Minimum evidence to be archived (traceability)
Planning	Sequencing of basic knowledge, temporal structure, design of units and sessions.	Generator of structure, alternatives and connections between contents.	A-M	B	It requires curricular context, group profile and precise instructions; The utility falls with generic instructions.	Propose a sequence of 8 sessions for 3rd year of Secondary on "[topic]", with objectives, activities, resources and evidence; It adapts to [curriculum] and to a group with [characteristics].	Prompt and version; final sequence annotated with teaching decisions; List of sources/data used.
	Design of itineraries and field trips (objectives, stops, times, logistics, safety).	Ideation, organization and operational checklist assistant.	M	M	You can propose stops, order times and generate checklists; validates coordinates, times and accessibility if data is provided. It is not a substitute for territorial recognition or risk management on the ground.	Design a 4-hour outing in [place] to work on [knowledge]; it includes an itinerary with stops, times, approximate coordinates and contingency plan; points out risks and accessibility needs.	Final itinerary; map/track and cartographic sources; safety checklist; Record of changes after the previous visit.
Resource Design	Adaptation of texts and instructions (easy reading, simplification, support levels).	Rewriter and linguistic adaptor; graduated version generator.	To	M	Mandatory review to avoid factual errors, biases or inappropriate simplifications; risk of "hallucinations" if rewritten without anchoring to sources.	Rewrite this text for students aged 14-15 with easy reading: keep the key concepts and add a glossary of 8 terms. Don't make up data; mark with [DATA] what needs to be verified.	Original text and versions; changelist; verification of marked data; Adaptation criteria.
	Case studies, fictional scenarios and role-plays on	Generator of scenarios, scripts, roles, simulated sources and	A-M	B	Useful for contextualizing and activating the debate; it requires	Create a role-play on "[conflict/theme]" with 5 actors, conflicting	Script/roles; actual sources used to anchor the case; criteria and rubric;

Dimension	Task (examples)	ChatGPT's contribution/role	G	R	Operational Observations	Example of prompt	Minimum evidence to be archived (traceability)
	territorial conflicts, resources and geopolitics.	discussion questions.			adjustment to the curriculum, to evidence and to real sources when the product requires factuality.	objectives, a brief dossier per actor (max. 150 words), 6 guiding questions and a final task with evaluation criteria.	evidence of the debate (minutes, audio, notes).
	Immersive and interactive simulations (climate change, migrations, erosion, urban planning) with variables and consequences.	Rule, variable, event, and feedback designer; generator of materials for dynamization.	A-M	M	It provides speed to prototype dynamics; it needs validation of the causal model, the ranges of variables and territorial coherence. It is advisable to integrate real data (e.g., climatic or demographic series) to avoid plausible but false results.	Design a classroom simulation on "[topic]" with 4 controllable variables, 3 random events, and explicit rules; It includes a registration sheet for students and analysis questions with evidence requirements.	Rules and versions; dataset or data sources; student registration sheets; Final analysis with traceability of decisions.
	Review questions, item banks, practice quizzes, and distractors.	Generator of items, variants and explanatory feedback.	To	B	Suitable for practice and review; it requires error checking, ambiguities, and biases in distractors.	Generate 20 multiple-choice items on "[topic]" with 4 options, justification of the correct one and explanation of why the others are plausible distracters; indicates level (basic/medium/high).	Item bank with version; teacher review (errors and adjustments); Outcome statistics if applied in LMS.
Direct instruction	Explanation of concepts and processes (e.g., coastal erosion, tertiary sector) with examples, analogies, and counterexamples.	Adjustable explainer; Generator of analogies, graduated examples and verification questions.	M-A	To	It saves time and supports clarification, but students must contrast; risk of plausible but incorrect answers. It requires instructions that force us to cite evidence and identify uncertainties.	Explain "[concept]" in 120 words for 3rd year of Secondary, use 2 analogies and 1 counterexample; End with 3 check-through questions and a list of facts that need to be verified with sources.	Prompt and response; data verification; Contrast activity (source consulted + students' conclusions).
Mentoring and support	Personalized tutoring (e.g., study modes), clarification of doubts and generation of adapted exercises.	Conversational tutor; study guide; differentiation of tasks and scaffolding.	M	M	It provides customization and availability, but requires verification and design of tasks that document the process. It is advisable to limit the scope: explanation, guided practice	Act as a guardian. Ask me 5 questions first to diagnose my understanding of "[topic]". Then propose a 20-minute plan with graded exercises and self-correction, without giving the complete solution on the first try.	Diagnosis (questions/answers); plan and exercises; record of attempts; Brief metacognitive reflection.

Dimension	Task (examples)	ChatGPT's contribution/role	G	R	Operational Observations	Example of prompt	Minimum evidence to be archived (traceability)
					and metacognition, not "doing the task".		
	Response to motivational questions and connection with students' interests without trivializing the content.	Generator of contextualized connections (interests-people-places) and close examples.	M	M	Engagement can be improved by avoiding misleading analogies or stereotypes; It is advisable to require students to translate the connection into a geographical argument.	Answer "why learn [topic]?" Connect with 2 possible interests (e.g., sports, video games, music) and end with a short task: "apply the concept to a real case and cite a source."	Initial question; response; application task with cited source; self-assessment of students.
Formative assessment	Initial feedback on drafts (structure, clarity, grammar) and support on rubrics and criteria.	First Layer Reviewer; suggestor of improvements; Base Rubric Generator.	M-A	B	Useful before teacher feedback; It does not provide personal context or knowledge of the actual process. Review biases and criteria.	Evaluate this draft with a rubric (4 criteria, 4 levels). Point out 5 improvements in structure and 5 in argumentation, without rewriting the entire text. It returns questions for me to improve.	Draft and versions; rubric used; comments from the model and the teacher; Registration of student reviews.
	Qualification of complex tests or products with deep geographical reasoning.	Automated "dialer" (use restricted to support).	B	To	Inconsistency, biases and lack of nuance; Not suitable for final grade without teaching judgment. If used, it should be limited to detection of formal criteria and diagnostic support, with auditing and control samples.	Apply this rubric (included) and justify each criterion with textual evidence (short quotes). Point out uncertainties and do not assign a final grade; proposes 3 questions for an oral review.	Official Signature; model output; textual evidence cited; Teacher review and final decisions.
Administration and communication	Draft reports, communications to families, homework instructions, minutes and institutional messages.	Draft generator and adaptation of tone and clarity.	To	M	It speeds up writing and reduces the "blank page"; requires confidentiality control, regulatory compliance, and tone review. Avoid entering students' personal data and verifying statements.	Write a communication to families about "[topic]" in 180-220 words: objective, what students will do, evaluation criteria and how to support at home. Clear, non-paternalistic tone, without personal data.	Draft and final version; compliance checklist (privacy/style); Review and approval record.
Geographical task	Territorial debate and dilemmas (e.g.,	Mediator and Socratic adversary;	To	B	It provides value for multiperspective	Simulate a debate about "[dilemma]." It generates 4 actors	Discussion script; sources provided; record of evidence

Dimension	Task (examples)	ChatGPT's contribution/role	G	R	Operational Observations	Example of prompt	Minimum evidence to be archived (traceability)
	urbanization vs. agricultural park; renewables vs. landscape) with multiperspective.	generator of actors' positions and counterarguments.			if teachers demand evidence and contrast of sources. Useful for simulating the voices of actors (mayor's office, primary sector, NGOs, promoters, etc.) without caricaturing.	with different interests, 3 arguments and 3 expected evidences per actor, and cross-questions. It includes a guideline for evaluating the quality of evidence.	cited by students; Minutes of conclusions.
	Analysis of spatial data (method, interpretation, verification) and support for the reading of maps and graphs.	Methodological assistant, verifier and generator of interpretation questions.	M	M	It requires triangulation with sources and data; Variable performance according to difficulty. Useful for proposing hypotheses, checking coherence and detecting inconsistencies in calculations or readings.	With these data/table/map described: (1) explain patterns, (2) propose 3 hypotheses, (3) suggest 5 data quality checks and (4) write 3 tentative conclusions indicating their level of confidence.	Dataset and metadata; analysis steps; checks carried out; Conclusions with a level of confidence and sources.
	Summary and synthesis of reports, articles and news to prepare classes and design activities to contrast sources.	Structured Summarizer; key idea extractor; Critical reading question generator.	A-M	M	Useful for processing long texts, but can omit nuances or introduce errors. It is advisable to require: short textual quotations, location of paragraphs and external verification of figures and statements.	Summarize this text in 10 numbered ideas. For each idea, add: (a) short quote (max. 20 words), (b) location (section/paragraph), and (c) a critical question. Mark any sensitive figures or data with [VERIFY].	Source text; summary with locations; checklist; Critical reading activity for students.
	Design of urban and territorial planning solutions with future scenarios and impact assessment.	Scenario simulator; generator of alternatives and matrix of impacts.	M	M	Useful for ideation, but it requires criteria, data and evaluation argued by the students. Avoid "pretty" proposals without territorial viability or evaluation of *trade-offs*.	Propose 3 management alternatives for "[case]". For each one, it includes objectives, measures, affected actors and a matrix of impacts (environmental, social, economic, landscape) with explicit assumptions.	Alternatives and assumptions; data and sources; impact matrix; Reasoned justification of the students.
	TIG/GIS support (ArcGIS Online or others): step-by-step guides, consultations, expressions,	Procedural and debugging assistant; Translator of requirements into operational steps.	M	B	It does not replace the GIS tool or guarantee visual accuracy; It helps formulate queries, structure	In ArcGIS Online I want "[target]". Give me a sequence of steps with expected captures, possible errors and how to	Target and dataset; steps executed; screenshots or registration; cartographic quality checks.

Dimension	Task (examples)	ChatGPT's contribution/role	G	R	Operational Observations	Example of prompt	Minimum evidence to be archived (traceability)
	procedure review.				flows, and understand errors. It requires tests in the real environment and mapping criteria.	check them. If you propose expressions or queries, explain to me what they do and how to check results.	
Not recommended	Generic descriptive essays for home without evidence of trial or oral defense.	Full text generator.	A-B	To	It breaks the relationship between process and product; it is advisable to redesign tasks with process evidence (drafts, sources, decisions) and oral or applied defense.	(Do not use for final delivery) Instead: "Help me plan the essay: Generate an outline, guiding questions, and a list of sources, but don't write the text. Ask me for evidence and review my draft."	Scheme; drafts; sources; registration of revisions; oral defense or applied product.
	Request "real" academic literature without systematic external verification.	Reference generator.	B	To	Risk of invention of sources; requires external verification (catalogs, DOIs, databases). Integrate a protocol for the verification and registration of bibliographic evidence.	Give me possible search lines and keywords (in ES/EN) on "[topic]". Don't make up references. Suggests databases and how to check DOI/ISBNs.	Search strategy; database captures/exports; verified list with DOI/ISBN.
	Direct creation of visual mapping as a substitute for a GIS tool.	GIS tool replacement.	N	To	You need mediation with code or TIG (e.g., ArcGIS Online); limits representational accuracy.	If you need a final map: describe the cartographic objective and ask for a flow in GIS (layers, symbology, scale, projection) and quality criteria, not a "made-up" image.	GIS Project; data sources; cartographic decisions; Exported final map with metadata.