

ORIGINAL ARTICLE

Open Access



# AI in the city: impact assessment of artificial intelligence uses in earth observation

Sanja Šćepanović<sup>1,2\*</sup> , Edyta Bogucka<sup>1,3</sup> and Daniele Quercia<sup>1,4</sup>

## Abstract

Earth Observation (EO) offers valuable insights into urban environments, and integrating artificial intelligence (AI) amplifies these benefits but also brings potential risks. AI practitioners often face challenges in envisioning diverse uses and conducting thorough impact assessments of their technology, particularly for less-studied uses. To address this, we developed UrbanGen, a framework validated through studies with urban EO practitioners and compliance experts. Practitioners and experts found UrbanGen valuable both for broad thinking and reflection by listing realistic AI uses for EO (91% accuracy) and identifying under-researched uses (57% accuracy), and for in-depth thinking and decision-making by providing risk (93% accuracy) and benefit (80% accuracy) assessments. UrbanGen highlighted less-studied and upcoming uses, such as analyzing foot traffic in retail areas, monitoring environmental law compliance, and detecting crowd sizes at election rallies. While most EO uses support sustainable cities, such novel uses pose higher risks, particularly in terms of surveillance, power imbalances, and decision-making detached from on-the-ground realities. Drawing from these insights, we propose an impact assessment checklist to help the EO community maximize benefits and reduce risks from AI deployments.

**Keywords** Urban AI, Earth observation, AI regulation, AI impact assessment, SDGs

## 1 Introduction

Earth Observation (EO) involves the collection, analysis, and presentation of data about Earth's physical and social systems through remote sensing from satellites, drones, or aircraft. This technology employs various sensing instruments, including optical, radar, lidar, and thermal sensors (Zhao et al., 2022). EO is particularly beneficial for urban environments, supporting sustainable urban planning (Musakwa & Van Niekerk, 2015; Hao et al., 2024; Xu et al., 2022; Šćepanović et al., 2021) and enhancing urban resilience (Cutter, 2021).

For instance, Guo et al. (2024) demonstrate how EO enables urban tree canopy mapping and comparison across 841 cities simultaneously in South America. EO aids in mitigating and responding to natural disasters (Schumann et al., 2018; Kondylatos et al., 2022) such as hurricanes (Li et al., 2016), wildfires (Crowley et al., 2023), and earthquakes (Elliott et al., 2016), as well as anthropogenic issues (Yang et al., 2013; Young & Onoda, 2017) such as air pollution (Li et al., 2005) and conflicts (Kaplan et al., 2022). Climate resilience efforts benefit from EO by understanding anthropogenic heat sources and measuring indicators such as land surface albedo, emissivity, temperature, and land cover and use (Šćepanović et al., 2021; Cartalis et al., 2015). Additionally, EO plays a crucial role in monitoring and reporting progress towards UN Sustainable Development Goals (SDGs) (United Nations: Department of Economic and Development, 2023). Chrysoulakis et al. (2014) provide a conceptual list of EO-based indicators for sustainable urban planning and management across various

\*Correspondence:

Sanja Šćepanović  
sanja.scepanovic@nokia-bell-labs.com

<sup>1</sup> AI Research Lab, Nokia Bell Labs, Cambridge, UK

<sup>2</sup> The Uehiro Oxford Institute, University of Oxford, Oxford, UK

<sup>3</sup> Department of Computer Science, University of Cambridge, Cambridge, UK

<sup>4</sup> Department of Control and Computer Engineering, Politecnico di Torino, Turin, Italy



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

sectors, including energy efficiency, air pollution, public health (Šćepanović et al., 2023), water, transportation, and hazard vulnerability.

Incorporating artificial intelligence (AI) technologies enhances the applicability and potential of EO, yet introduces and amplifies risks and ethical issues (Kochupillai et al., 2022; Tuia et al., 2023; Micheli et al., 2022). Some risks arise directly from AI, such as model transparency issues and the sustainability concerns of training large models. Others are inherent to EO but amplified by AI, including power imbalances in data acquisition, ownership, processing, detachment from on-the-ground societal complexities, data biases, privacy and surveillance issues, and communication challenges in EO analysis results (Kochupillai et al., 2022). AI practitioners must envision potential uses and produce AI impact assessment reports to mitigate these risks, overweighing both benefits and risks (Stahl et al., 2023). Increasing AI regulations also demand these impact assessments (Smuha, 2021; European Commission, 2023). According to the OECD, legal advisory costs to ensure compliance with various regulatory systems are estimated to be nearly 780 billion dollars annually worldwide (International Federation of Accountants, 2018). However, cataloguing AI uses and associated risks is both challenging and time-consuming (Moraes et al., 2021; Liang et al., 2024; Hassel & Özkiziltan, 2023; Prunkl et al., 2021). Traditional evaluations of benefits of EO technologies relied on experts panels by UN and EO specialists (Group on Earth Observation, Committee on Earth Observation Satellites, 2017; Zhang & Zhang, 2022). Evaluations of risks and harms are even rarer, usually focusing on technical aspects (Zhang & Zhang, 2022; Hassel & Özkiziltan, 2023), or specific use cases (Kochupillai et al., 2022). A part of the reason is, as recent research shows, that AI developers struggle to specify uses and impacts for model cards (Liang et al., 2024), data cards (Yang et al., 2024), and societal impacts sections (Prunkl et al., 2021; Ashurst et al., 2022) required by top AI conferences. These tasks are challenging because they require (1) envisioning possible uses of AI technology, including those that are under-studied (broad or divergent thinking) and (2) systematically assessing these uses (in-depth or convergent thinking) (Nanayakkara et al., 2021; Prunkl et al., 2021; Kochupillai et al., 2022; Hassel & Özkiziltan, 2023).

In summary, previous research has focused on the benefits of AI for urban EO while neglecting comprehensive risk assessment. This highlights the need for a systematic approach to help EO practitioners envision uses and evaluate both benefits and risks. To address this, we implemented a semi-automatic method to support AI for EO, making two main contributions:

1. We implemented and validated the UrbanGen (Urban Generative AI) framework, which comprises three LLM prompts and an automatic literature coverage check. This framework is designed to list realistic urban AI uses, highlight not well studied ones, assess the risks of each use based on the EU AI Act, and evaluate benefits using the SDG taxonomy. We validated our framework in 18 individual user studies with 9 AI EO practitioners and 9 AI compliance experts. The participants found that our framework enabled them to envision a broader set of uses than they could without it (Sec. 3).
2. We thematically analyzed the uses and found that some have significant potential for improving sustainable urban development, climate resilience, and citizen well-being. However, others pose significant risks, including privacy and surveillance issues, data biases, model over-reliance, automated decision-making, and power imbalances, particularly in disadvantaged communities. Notably, upcoming and novel uses tend to be riskier compared to existing ones. Expert studies revealed that both EO and compliance experts, despite using different strategies to envision EO uses and analyze their risks, found UrbanGen highly useful in their work (Sect. 4).

Lastly, we discuss our results, emphasizing the need for thorough impact assessments. Consequently, we propose a new Impact Assessment Checklist for EO in Urban AI (Sect. 5).

## 2 Background and related work

We first introduce the field of Responsible AI (Sect. 2.1), followed by a discussion on existing AI regulations and the challenges of impact assessment (Sect. 2.2), then review prior studies assessing the benefits and risks of AI in EO (Sect. 2.3). Lastly, we explore recent research on LLMs that has the potential to support practitioners in these assessments through a semi-automatic process (Sect. 2.4).

### 2.1 Responsible and trustworthy AI

As Artificial Intelligence (AI) becomes increasingly pervasive in industries and daily life, balancing its development with ethical considerations has become essential. The fields of responsible and trustworthy AI have emerged in both academia and industry to address the ethical challenges posed by AI technologies. Academic research has significantly contributed to this area, exploring issues such as algorithmic fairness, transparency, accountability, and the impact on various stakeholders (Deshpande & Sharp, 2022; Mikalef et al., 2022; Li et al., 2023; De Miguel Velazquez et al., 2024). Leading

industry players like Nokia and NVIDIA have recognized this necessity by establishing ethical principles for AI (Nokia Bell Labs, n.a.; NVIDIA Corporation, n.a.). These principles include: *fairness*-designing AI systems to maximize non-discrimination and accessibility; *reliability, safety, and security*-ensuring AI causes no direct harm and minimizes indirect harmful behavior; *privacy*-respecting individuals' control over their data and the decisions made using their data; *transparency*-making AI systems explainable and understandable to allow for human oversight; *sustainability*-striving for societal and environmental sustainability by empowering society and reducing energy consumption; and *accountability*-involving consultation and collaboration in AI development to ensure true accountability. In the Earth Observation community, there is a growing recognition of these ethical considerations, as evidenced by recent studies (Kochupillai et al., 2022; Šćepanović et al., 2023). Responsible AI aspects are also being discussed in urban science and urban informatics, highlighting the importance of ethical considerations in these fields (Khayyat et al., 2021; Zhang & Liu, 2022).

## 2.2 AI regulations and challenges in impact assessments

The increasing implementation of AI has led to calls for regulatory oversight (Tahaei et al., 2023a; Borenstein & Howard, 2021; Floridi et al., 2021). The US Office of Science and Technology Policy issued a non-binding AI Bill of Rights Blueprint, emphasizing principles such as safety, non-discrimination, data privacy, and AI transparency. Similarly, the European Commission's AI Act, a legally binding framework, strives to balance innovation with the protection of societal values and rights. This Act categorizes AI applications by risk levels, ranging from low (those that do not require conformity assessment) to high risk (those that do) to prohibited (harmful or manipulative AI applications) (European Commission, 2023). It also leaves space for updates acknowledging the need to adapt to further expected technological advancements (Hutson, 2023).

These regulations mandate impact assessments of AI applications. Typically, a report is generated to document the potential benefits and risks of AI. CredoAI's report template includes sections on formal system evaluations and compliance with laws, regulations, and standards (Sherman & Eisenberg, 2024). Similarly, the National Institute of Standards and Technology (NIST) report contains listings of potential biases, evaluations of the impact's magnitude and likelihood, technical specifications, third-party technologies, and legal and compliance issues (National Institute of Standards and Technology, 2023a). AI practitioners face significant challenges in performing impact assessments, particularly in mapping

intended and unintended AI uses and their risks. For example, a study of 32 K model cards (Mitchell et al., 2019) on HuggingFace reveals that while most cards detail *Training Information*, the *Intended Uses and Bias, Risks, and Limitations* sections have much lower completion rates (17–23%) (Liang et al., 2024). Similarly, Data Cards on the same platform show that the *Considerations for Using the Data* section is the least filled, comprising only 2.1% of the card's text length (Yang et al., 2024). Additionally, leading AI conferences like NeurIPS require sections on broader societal impacts (Nanayakkara et al., 2021), but researchers often struggle to complete these due to the complexity and high opportunity costs involved (Prunkl et al., 2021).

A recent study (Golpayegani et al., 2023) highlighted five key elements for conducting regulatory risk assessments under the EU AI Act: domain, purpose, capability, AI deployer, and AI subject. The *domain* pertains to the industry or sector, such as innovation and research. The *purpose* specifies the objective, like analyzing urban heat islands. The *capability* details the technology used, such as detecting temperature variations from infrared data. The *AI deployer* refers to the individual utilizing the system, such as urban planners. The *AI subject* is the person impacted by the system, such as city residents. We leverage these five components for semi-automatic use generation and risk assessment using LLMs.

## 2.3 Impact assessments of AI in EO

Previous efforts to assess AI for EO can be broadly categorized into two main groups: (1) benefit-oriented assessments, which primarily focus on highlighting potential benefits and opportunities; and (2) risk-oriented assessments, which also consider emerging ethical issues and potential risks.

**Benefit-oriented assessments.** AI is an essential component of EO applications, from vehicle detection (Gao et al., 2019) and land cover mapping (Šćepanović et al., 2021) to federated learning for distributed EO settings (Tuia et al., 2023) and physics-informed models integrating physical constraints (Reichstein et al., 2019). It also enhances visualization and storytelling (Sacha et al., 2017).

EO data's global availability and spatial-temporal consistency have led to mainly highlighting its benefits (Tuia et al., 2023). EO is traditionally recognized for its contribution to environmental sustainability (Jin et al., 2022; Wang et al., 2022), including climate (Yang et al., 2013), air quality (Chu et al., 2003), sea levels (Cazenave & Nerem, 2004), ocean surface (Sun et al., 2022), forests (Praks et al., 2012; Antropov et al., 2017), and biodiversity (Kuenzer et al., 2014) monitoring. These benefits also include improving crop resilience (Jung et al., 2021).

Another set of benefits relates to sustainable cities and society (Persello et al., 2022). Experts from the Group on Earth Observations and the Committee on Earth Observation Satellites emphasize EO's crucial role in the 2030 Agenda for Sustainable Development (United Nations: Department of Economic and Development, 2023). EO provides evidence and insights into environmental changes and resource management, aiding the understanding of SDG progress and challenges. They identified 16 SDGs and 71 targets benefiting from EO (Group on Earth Observation, Committee on Earth Observation Satellites, 2017).

**Risk-oriented assessments.** The challenges and risks of AI in EO are relatively less discussed, and often centered around technical issues. For instance, in the list of 12 challenges in AI for EO (Persello et al., 2022), only one addresses potential ethical risks, while the others focus on aspects like data preparation, model development, and training optimization. Similarly, four out of five challenges identified in Persello et al. (2022) concentrate on uncertainty quantification, data quality, model explainability, and transferability.

A broader risk assessment is performed by research focusing on specific use cases (Kochupillai et al., 2022; Tuia et al., 2023). This research devises ethical principles in AI for EO: privacy, honesty, fairness, integrity, responsibility, and sustainability (Kochupillai et al., 2022). *Privacy* concerns, for example, agricultural contexts where land surveillance often occurs without the knowledge or awareness of farmers (Kochupillai et al., 2022; Copa-Cogeca et al., 2018). The EU Code of Conduct on agricultural data sharing (Copa-Cogeca et al., 2018) emphasizes the requirement for farmers to know and participate in how their data is used. *Honesty and transparency* in communication of results is crucial for applications such as prioritizing rural electrification (Kochupillai et al., 2022). *Fairness* is crucial in labeling data to identify informal settlements, often termed slums (Micheli et al., 2022; Tuia et al., 2023), because labeling an area as a 'slum' can stigmatize the residents, as the term carries a negative connotation (Kochupillai & Taubenböck, 2023). The concept of *integrity* involves the model's accuracy and reliability, questioning if the system delivers on its promises. An example from Kochupillai et al. (2022) is the assumed higher phytoplankton concentration near the poles, later found to be a wrong result due to atmospheric correction software, highlighting the importance of accuracy in EO data (Friedlingstein et al., 2022). The principle of *responsibility* places a strong emphasis on human agency and oversight, a requirement articulated not only in the EU's guidelines for trustworthy AI (High-Level Expert Group on AI, 2019) but also in the EU AI Act (European Commission, 2023), particularly for high risk AI systems.

This principle calls for attention to complex societal issues on the ground and engagement with stakeholders. Neglecting these considerations could lead to providing recommendations based on EO data that are irrelevant and unlikely to be embraced by local populations (High-Level Expert Group on AI, 2019; Kochupillai et al., 2022; Tuia et al., 2023). The final principle, *sustainability*, encompasses both environmental and social dimensions. In addition to environmental sustainability benefits (Jin et al., 2022; Wang et al., 2022) and concerns due to the vast amount of data involved in training EO models (Šćepanović et al., 2023), EO should also address social sustainability, such as the living and working conditions of local and indigenous communities (Kochupillai et al., 2022). Moreover, the perceptions of these ethical principles vary among EO experts, adding to the complexity of the impact assessments (Šćepanović et al., 2023).

To sum up, risk-benefit analyses for AI in EO have been hitherto hindered by two primary limitations. First, these analyses often focused on technical issues, and those that went further only analysed a select few individual AI use cases, potentially overlooking the broader spectrum of applications. Moreover, the assessments are predominantly conducted manually by experts in EO and regulatory professionals, leading to a laborious and costly process. Overcoming these limitations requires the development of new approaches that automate or semi-automate part of the risk-benefit assessment process. LLMs promise to achieve this.

#### 2.4 LLMs and prompt engineering best practices

LLMs like OpenAI's ChatGPT (OpenAI, 2023b), trained on trillions of tokens using advanced deep learning methods (Liu et al., 2023; Ganaie et al., 2022), generate human-like text in response to prompts. They are ideal for answering questions and crafting content, leveraging the vast knowledge they were trained on (Wei et al., 2023). With a proper care and caution taken while using them, LLMs have proven useful in a variety of tasks, from text annotation (Gilardi et al., 2023; Chung et al., 2022), and co-creation (Chung et al., 2022; Tolmeijer et al., 2022), to brainstorming assistants (Lukowicz et al., 2023; Bouschery et al., 2024), to *interpreting regulatory texts* (Zheng et al., 2023; Cui et al., 2023) and *applications in urban science research* (Fu et al., 2024). To achieve the desired output from LLMs, a set of best practices for prompt engineering has emerged (Shieh, 2023). We detail the main ones next.

*System Role:* Defining a system role in the prompt initiates interactions with language models by specifying the assistant's persona. This priming technique is instrumental in customizing language model behavior (Azure, 2023).

**Specific and Actionable Instruction:** Effective instruction involves articulating tasks clearly and precisely, establishing unambiguous expectations. This entails selecting appropriate action words (e.g., describe, summarize, explain) and defining the desired output format (e.g., JSON, one sentence).

**Optional Prompt Elements:** To refine results, optional prompt elements like cues or nudges offer additional guidance to steer the model's output in a specific direction (Azure, 2023; Shieh, 2023). For example, this can be achieved by asking the model to provide answers in a particular domain or area.

**Order Matters:** The arrangement of prompt elements significantly influences outputs, a phenomenon known as recency bias. It is advisable to position the most crucial prompt elements at the beginning and end of the prompt (Liu et al., 2024).

**Learning Techniques:** Various learning techniques can enhance performance. Beyond zero-shot learning, where models rely solely on internal knowledge, few-shot learning, with task-specific example input-output pairs, improves niche task performance. Chain-of-thought (CoT) reasoning guides the model's thinking by requesting step-by-step responses, particularly useful for tasks involving counting and mathematics (Wei et al., 2022a, b).

**Model Parameters:** Tweaking specific model parameters, such as temperature provides enhanced control over model outputs. The temperature parameter balances between randomness and determinism in generated text, with higher values fostering diversity, and lower values yielding more focused, consistent outputs.

**Evaluation Techniques:** When working with LLMs, one must consider the risk of the model producing biased or inaccurate information, referred to as hallucinations (Dziri et al., 2022). Also, for complex tasks with specific quality or creativity expectations, evaluation techniques are recommended. These approaches ideally involve human evaluators to ensure content meets quality standards and is realistic (Eldan & Li, 2023; Dziri et al., 2022).

As highlighted, traditional risk-benefit assessments of AI for EO have been sporadic, fragmented, and resource-intensive. LLMs incorporating the above best practices show the potential to uncover benefits and risks in a semi-automated manner and support urban AI practitioners.

### 3 Gathering insights from study participants

#### 3.1 Implementing UrbanGen

Our framework, UrbanGen (Fig. 1), includes three integrated LLM prompts and an automatic literature coverage check. Prompt 1 lists diverse and realistic uses of EO technology; Prompt 2 assesses the risk of these uses with relevant text from the EU AI Act; Prompt 3 evaluates if a

given use supports UN SDG targets. The automatic literature coverage check analyses 3 million Semantic Scholar papers to identify under-researched uses. In this section, we provide a short description of each UrbanGen module, while Appendix 2 provides full details.

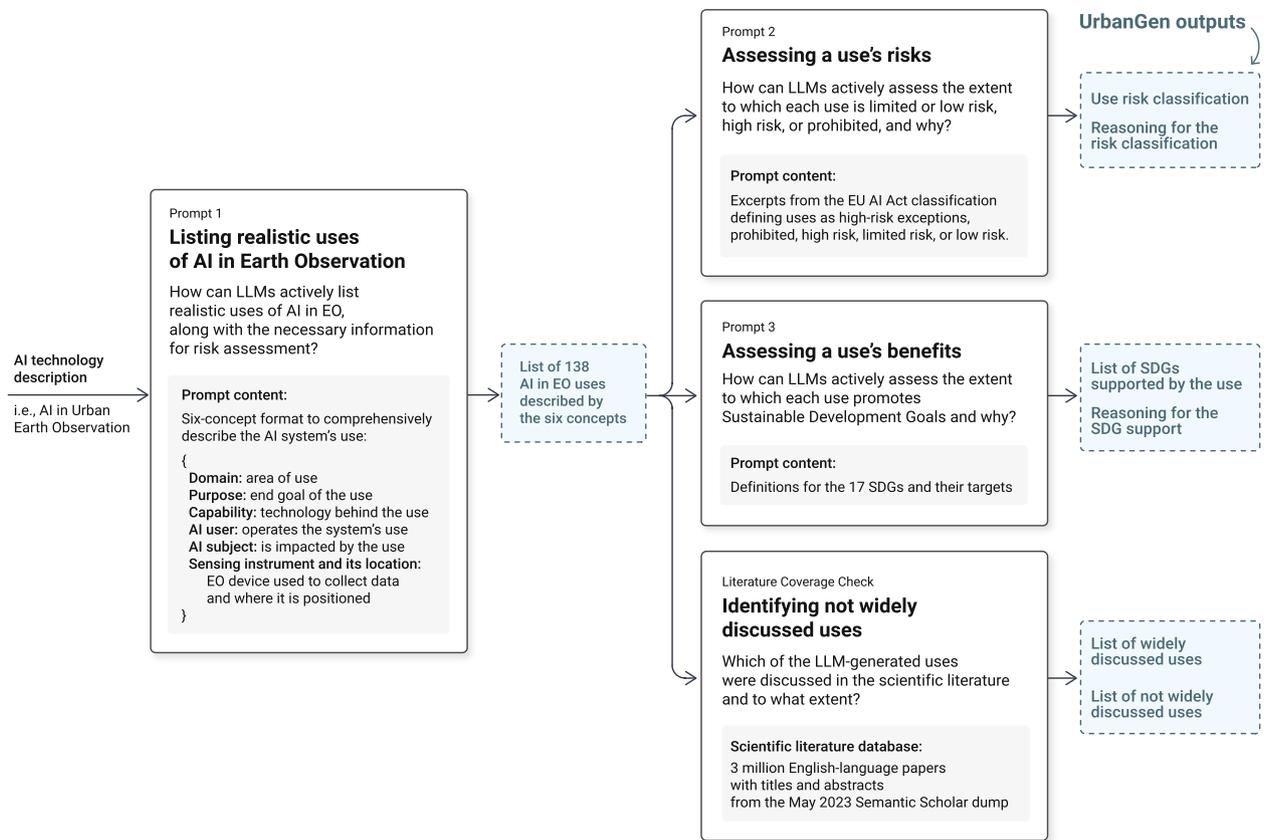
All our prompts were run using the GPT-4 model (OpenAI, 2023a), chosen for its leading performance across leaderboards for various tasks (LMSYS, 2024). Specifically, we used the recent GPT-4o variant, which offers similar performance while being more cost-effective for the large number of calls we required. To limit hallucinations, we set the temperature parameter to zero for all of our prompts.

**Listing realistic uses of AI in EO** We identified realistic EO technology uses through a three-step process. First, we followed guidelines from prior literature to ensure the risk level could be reliably assessed according to the EU AI Act (Golpayegani et al., 2023). This involved describing the uses along the five key concepts: domain, purpose, capability, AI deployer, and AI subject. Additionally, we introduced a sixth concept specific to EO technology: sensing instrument and location. Second, we covered uses from a wide range of relevant domains. We curated 46 domains, listed in Appendix 1: Table 2, starting with the EU AI Act and its Amendments (European Commission, 2023), and incorporating additional domains based on focus group discussions using the Social-Ecological Model (Golden et al., 2015). Third, we engineered an LLM prompt using Chain-of-Thought (CoT) reasoning to generate uses in a predefined format, minimizing unrealistic outputs (Wei et al., 2022b). The prompt is structured into five parts: *System role*, *Multi-part instructions*, *Six Concepts, Domains*, and *Examples* (Giray, 2023; OpenAI, 2023a; Brown et al., 2020).

The full prompt, including placeholders for domains and examples, is provided in the Appendix 2 as *Urban EO Uses Generation Prompt*.

**Assessing a use's risks** We collated relevant sections from the EU AI Act (e.g., Article 5 for assessing prohibited uses) (European Commission, 2023) and leveraged rules from the official AI Act Compliance Checker<sup>1</sup> for a more accurate AI use risk assessment. Prompt 2, like Prompt 1, is structured around five components: *System role*, *Multi-part instructions*, *Legal clauses*, *Placeholder*, and *Output structure*. The key difference lies in the multi-part instructions, which guide the LLM to "reason in steps" by evaluating whether the input AI use meets

<sup>1</sup> <https://artificialintelligenceact.eu/assessment/eu-ai-act-compliance-checker/>



**Fig. 1** Overview of our UrbanGen framework. It comprises four integrated components: three LLM prompts and an automatic literature coverage check. The prompts serve to (1) list realistic uses of AI for Earth Observation (Prompt 1), (2) assess the risks associated with these uses (Prompt 2), and (3) evaluate their potential benefits (Prompt 3). The automatic literature coverage check then identifies uses that are not widely discussed in the scientific literature. UrbanGen takes in input a given technology (in our case, AI in Urban Earth Observation) and generates 138 diverse uses of that technology, along with risk classifications and the Sustainable Development Goals (SDGs) supported by each. Additionally, it highlights the subset of generated uses that are underrepresented in the literature

specific criteria. For instance, to determine if the use is excluded from the EU AI Act, the question posed is *Does the input AI system fall within any of the following categories?*, along with a list of excluded categories from Article 2. Depending on the response, the model either moves on to the next step or labels the use as excluded. Subsequent steps assess whether the use is prohibited, high-risk, etc.

The full prompt text, including placeholders for relevant EU AI Act references, is provided in the Appendix 2 as *Risk Assessment Prompt*.

**Assessing a use's benefits** We aggregated the official UN descriptions for each of the 17 SDGs and their targets (United Nations: Department of Economic and Development, 2023). Prompt 3, like Prompts 1 and 2, is structured around *System role, Multi-part instructions, SDG Definitions, Placeholder, and Output structure*.

The full prompt text, with placeholders for SDG definitions, indicators, and targets, is provided in the Appendix 2 as *Benefits Assessment Prompt*.

**Identifying not widely discussed uses of AI in EO** To identify potentially under-researched LLM-generated uses, we analyzed 200 million papers from the May 2023 Semantic Scholar dump, filtering to 3 million English-language papers with titles and abstracts. We embedded the title + abstract fields of each paper and the LLM-generated use descriptions using the *all-mpnet-base-v2* model (Reimers & Gurevych, 2019). This model, fine-tuned on over 1 billion sentences, balances effectiveness and efficiency. For each paper, we identified the highest similarity use based on cosine similarity of their embeddings. We manually set a similarity threshold to ensure accurate matching, experimenting with various thresholds. The 99.9<sup>th</sup> percentile yielded 3,295 paper-use pairs, for which our

manual validation on those with the lowest similarity scores confirmed that the paper indeed discussed the use. Upon this, uses associated with zero papers were classified as potentially under-researched.

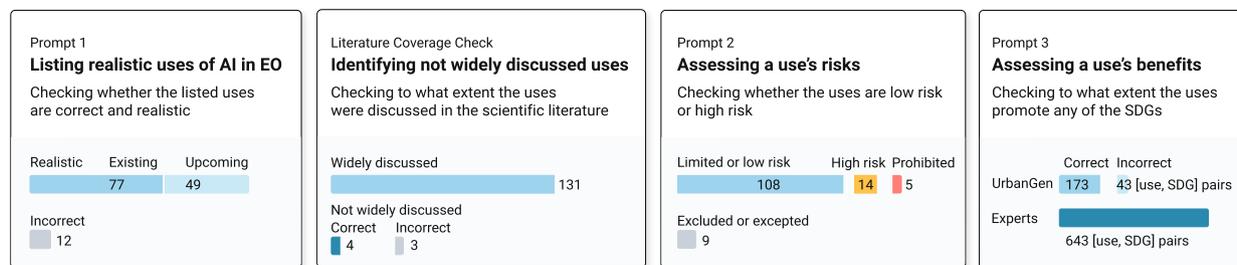
### 3.2 Running UrbanGen

UrbanGen identified 138 EO technology uses (Appendix 3, Table C). LLM categorized these uses as follows (Fig. 2): 9 (7%) excluded from EU AI Act classification, 2 (1%) as high-risk exceptions, 5 (4%) as prohibited, 14 (10%) as high risk, and 108 (78%) as limited or low risk. Examples are shown in Table 1. From analyzing 3 million Semantic Scholar abstracts, we found that 7 of the

138 LLM-generated uses were not discussed in any abstracts. These “under-researched” uses are presented in Table 3 in Appendix 4, suggesting they might be mentioned in non-scientific sources but lack detailed scientific research. LLM identified 216 benefits of the EO uses across 17 SDGs, with the most uses benefiting SDG 11: sustainable cities (41 uses) and the least benefiting SDG 5: gender equality (0 uses).

### 3.3 Gathering expert labels and insights

To gather validation labels and insights based on UrbanGen’s output, participants were recruited and given specific tasks aligned with the study protocol.



**Fig. 2** UrbanGen generated 138 uses, out of which 7 were found as not widely discussed by the literature, and it labelled their risk level and SDG benefits. Expert manual validation was conducted for each framework component. EO experts validated 126 (91%) realistic uses, including 49 upcoming ones, with 4 uses confirmed as under-researched by the authors. Compliance experts agreed with the risk classification in 93% of cases, identifying 14 high-risk uses and 5 prohibited under the EU AI Act. Each use supported at least one SDG, with LLM covering 27% of the SDGs identified by EO experts (or 44% when confirmed by at least two of three experts) and achieving an 80% accuracy

**Table 1** Example urban EO AI uses in each of the four risk categories

ID	Description	Risk Classification	Justification
58	Studying climate change effects by comparing historical and current satellite data	Excluded	Article 2(6): AI systems which are not placed on the market or put into service, exclusively used for research, testing, or development
82	Detecting military activity from satellite imagery	Excluded	Article 2(3): AI systems developed and used exclusively for military purposes
1	Identifying individuals in crowded areas using facial recognition by police	Prohibited	Article 5(1)(h): Real-time remote biometric identification for law enforcement
3	Categorizing population demographics by visual traits	Prohibited	Article 5(1)(g): Biometric categorisation systems
95	Detecting illegal fishing by identifying fishing vessels from radar data	High-Risk Exception	Annex I, Section B: AI system’s involvement in marine equipment
55	Navigating autonomous agricultural robots	High-Risk Exception	Annex I, Section B: AI systems used as safety components in agricultural and forestry vehicles
62	Identifying water pipeline leaks from radar data	High Risk	Annex III, Section 2: AI system’s role in the management and operation of critical infrastructure
64	Detecting suspicious behavior from aerial footage using optical drones	High Risk	Annex III, Section 6: used by or on behalf of law enforcement authorities in support of law enforcement
75	Analyzing environmental changes used by journalists	Limited or Low Risk	Article 6(3a): AI system performs a narrow procedural task by aiding journalists in data analysis, minimally impacting health, safety, or fundamental rights
107	Designing green spaces by urban planners	Limited or Low Risk	Article 6(3a): performing a narrow procedural task to assist urban planning

**Participants** We recruited participants in two cohorts: urban EO practitioners and compliance experts.

**Recruiting urban EO practitioners.** Using snowball sampling, we started with participants from our professional networks, who then referred others. We sought individuals developing urban AI systems with EO using machine learning, computer vision, and image recognition. We recruited 9 practitioners: five females and four males, aged 30 to 46 years ( $\mu = 37$ ,  $\sigma = 7.2$ ). Their countries of residence included Germany ( $n=4$ ), Italy ( $n=2$ ), the Netherlands, Luxembourg, and the United Kingdom ( $n=1$  each). They had significant AI expertise ( $\mu = 7$  years,  $\sigma = 3.6$  years), primarily in academia ( $n=7$ ) and industry ( $n=2$ ).

**Recruiting compliance experts.** We used snowball sampling, starting from an internal mailing list at a large tech company, seeking experts familiar with the EU AI Act and experienced in AI use case reviews. We recruited 9 practitioners: seven females and two males, aged 35 to 57 years ( $\mu = 42$ ,  $\sigma = 7.7$ ). They had AI assessment experience ( $\mu = 2$  years,  $\sigma = 2.6$  years), primarily in industry ( $n=6$ ) and academia ( $n=2$ ). Their countries of residence included Finland ( $n=4$ ), the Netherlands ( $n=2$ ), Belgium, Germany, and Morocco ( $n=1$  each).

**Study Protocol** We conducted 18 one-on-one sessions: with the 9 urban EO practitioners and with the 9 compliance experts, each ranging between 45 and 85 minutes. Sessions were held via video conferencing, and audio, computer screens, and transcripts were recorded with participants' permission, resulting in 18-hour-long transcripts.

Each session included two semi-structured interviews and two tasks. First, we interviewed participants about their current practices and challenges in envisioning AI technology uses and their associated risks in general. Second, we presented an interactive list with our generated AI uses in EO, along with the uses' LLM risk assessment, asking them to select one use that balances interest and adherence to their company's code of conduct (Fig. 6A in Appendix 5). Third, we provided 46 interactive assessment cards, for a subset of our uses. EO AI practitioners annotated uses for realisticness, transformational impact, perceived societal and environmental risk (Bronfman et al., 2012), and potential SDG benefits (Fig. 6B in Appendix 5), while compliance experts annotated uses for realisticness and accuracy of risk assessment (Fig. 6C in Appendix 5). The details about how realisticness, SDG benefits, and accuracy of risk assessment were solicited from participants are presented in

the section on validation in Appendix 4. Moreover, participants were engaged in evaluating each use's societal and environmental risks using a 7-point Likert scale, spanning from insignificance to severe concerns. They also rated the likelihood of these uses on a similar scale, from improbable to highly probable. Furthermore, they assessed the framework's practicality using a 5-point Likert scale, rating its utility from 'Not useful' to 'Extremely Useful'. Qualitative analysis of session recordings and transcripts investigated participants' insights on challenges related to envisioning risks in EO applications and identified key factors influencing the framework's efficacy in addressing these issues.

To conduct the tasks, we developed two visualizations (Fig. 6 in Appendix 5). The first visualization presented all AI uses in EO as tiles, color-coded by risk level according to the EU AI Act (Fig. 6A in Appendix 5). Users could explore uses by hovering over tiles to read explanations about the risk level and possible benefits. The second visualization was a web-based survey with five pages (Fig. 6B in Appendix 5). The first page outlined the study and tasks. The second page defined risky uses per the EU AI Act. The third and fourth pages presented assessment cards for 46 uses (23 per page) with checkboxes and input boxes for annotations. The final page included a demographics survey and a confirmation note.

Each of the 138 uses was annotated by six different AI practitioners: three EO developers and three AI compliance experts. To ensure originality, pasting from external sources and editing previous responses were disabled.

### 3.4 Validating UrbanGen

When validating our framework (detailed results in Appendix 4), we found that it enabled our participants to envision a broader set of uses than they could without it.

As shown in Fig. 2, out of the 138 uses identified by the framework, 91% were deemed realistic by the EO experts, with 56% already existing (e.g., *forecasting water levels from radar data*, use #98) and 35% considered upcoming (e.g., *identifying suspicious cargo from satellite imagery*, use #111). Conversely, 9% were seen as unrealistic, often in domains with fewer EO applications, due to impractical or exaggerated EO capabilities (e.g., *suggesting romantic outdoor spots*, use #7).

Appendix 4: Table 3 lists 7 uses identified as under-researched from Semantic Scholar abstracts. For three uses (*use #43*, *use #85*, and *use #117*), recent Google Scholar articles suggest they should be removed from this category, likely due to these articles not being included in our 3M Semantic Scholar dataset with abstracts. The

remaining four uses (57%) were confirmed as under-researched, lacking specific research papers (e.g., *analyzing foot traffic from drone imagery in retail areas*, use #47), highlighting the need for more studies in urban EO applications.

Compliance expert annotations of uses' risk levels aligned with the LLM's for 93% of the uses, with substantial agreement (Cohen's kappa 0.80). Disagreements included cases where the LLM labeled uses as "Limited or Low Risk", but experts had concerns, such as use #70 (*monitoring election rallies*) and use #65 (*tracking stolen vehicles*). Conversely, some uses marked "High Risk" by the LLM, like use #26 (*detecting hazardous conditions for workers using thermal drones*), were considered low risk by experts because they provide safety to the workers. Inter-annotator agreement among experts was moderate (Fleiss' Kappa 0.57), lower than with the LLM, suggesting a possible anchoring effect (Furnham & Boo, 2011; Nourani et al., 2021). Mitigating this would require independent expert assessments requiring them to search the EU AI Act for each use's assessment, which was impractical due to time constraints.

Lastly, UrbanGen identified 216 (use, SDG) pairs promoting an SDG, compared to 643 pairs by the EO experts (Fig. 2). The LLM covered 27% of expert-identified pairs but had an 80% alignment with expert evaluations. When considering SDGs labeled by at least two experts per use, the LLM identified 44% of the 262 pairs found by experts, with 54% precision. While experts could not evaluate each of specific SDG targets due to time constraints, the LLM did so and effectively aligned with and expanded upon the UN report identified targets (Appendix 4, Fig. 5).

#### 4 Main insights

*Urban EO practitioners and compliance experts use different strategies to envision EO uses* EO practitioners focus on literature reviews, research questions, data exploration, and expert exchanges, while compliance experts rely on sector experience, specialized roles, regulatory updates, and training sessions.

Urban EO practitioners start from literature reviews, conducting them to find related tasks and methodologies that can be adapted for new uses of EO technologies. This includes categorizing technologies into common task types such as predictive tasks, generative tasks, classification, and anomaly detection. Second, they start from specific research questions or hypotheses within the geospatial domain, leveraging existing geographic information and domain knowledge to guide their approaches. Third, many EO practitioners start from exploring how to apply new methods to existing data, leading to the discovery of

new uses. P7, involved in developing crop reporting datasets from European Union countries, gives two examples of such new uses: *"Just knowing what is growing in Europe is already an incredibly valuable asset. We didn't have this in mind when we originally started developing the dataset, but it allowed us to study agricultural diversity and its potential impact on biodiversity. First, we simply gridded Europe and checked the number of different crops growing in certain areas to calculate agricultural landscape's diversity. Second, we investigated combinations of certain crops as proxies for the presence of certain species, such as a beetle population. Third, we compared the diversity in urban areas with that of non-urban areas to understand urbanization effects"*. Finally, practitioners engage with experts in relevant sociotechnical fields, to brainstorm and evaluate the feasibility of new uses. To see how, P4 gave an example when they *"worked with an expert in LLMs to explore the possibility of LLMs to summarize the results of analyses of areas affected by natural hazards, such as the number of people affected. These LLMs would then be used to send SMS warnings to people in areas where access to the internet is typically unavailable during natural hazards. We were excited about the technology, as we always are. But then we stopped and asked: Can we trust LLMs to provide useful and accurate information? We are speaking about people's lives! We jointly decided that, for this task, as of today, we cannot trust them"*.

Compliance experts, on the other hand, often draw on their previous work experiences in specific sectors, such as agriculture, and development, to envision new uses. If they lack such specific work experience, they delegate the task of learning about new uses and brainstorming to specialized roles, such as solution managers within their organization. Compliance experts also rely on updates from regulatory bodies to ensure alignment with the latest regulations and to anticipate changes that may influence technology uses. Additionally, they participate in training sessions and workshops.

*Both expert groups share challenges related to data governance and the reliability of AI-driven EO uses* Urban EO practitioners face five significant challenges when identifying uses and risks of AI technologies. First, they are concerned with effective data governance and management, which includes ensuring privacy and preventing data misuse, particularly with sensitive or personal information ( $n=3$ ), and mitigating risks associated with handling large volumes of data ( $n=2$ ). Second, they often find that the actual implementation of AI systems is far more complex and labor-intensive than anticipated ( $n=2$ ). P1, who works in developing countries, explains:

“there is no data available, and existing models often do not work out-of-the-box. It takes time to set up a new pipeline, and it is often not much faster than manual labelling, contrary to what people in the research community might think”. Third, practitioners are aware of inherent biases in data, AI models, and development workflows, which can result in unfair or inaccurate outcomes ( $n=2$ ). As noted by P6, “the first step in developing a robust architecture is data collection. This process should rely on expert knowledge because, without it, we may select variables that are not relevant to the final target. This can introduce many pseudo-correlations”. Fourth, they are concerned with identifying and mitigating unintended consequences of their uses, that can occur due to the unpredictable nature of AI systems or their changing contexts ( $n=2$ ). To see how, see the reflection of P3: “If you are designing a project, of course you have a clear target in mind, but along the way, there are some unintended uses. For example, we develop a fire monitoring system that is publicly available. We process the data and provide it to the public. However, the monitoring system also detects war activities. We did not expect that, but it is clear that war produces a lot of heat reflections from the Earth’s surface, making mapping war activities an unintended product. We did not think about it at all, but we realized we could see that in the data”. Fifth, practitioners face the challenge of minimizing over-reliance on AI models in decision-making, which can lead to a false sense of security ( $n=2$ ).

Compliance experts face three challenges. First, they often encounter regulatory issues late in the product development process, leading to significant setbacks or even project cancellations despite substantial R&D investments. Second, they are concerned about the potential for sensitive information to be compromised. Third, they encounter challenges related to the reliability and trustworthiness of AI-generated results, emphasizing

the need for human review and not taking AI outputs at face value.

*UrbanGen is highly useful for broad and in-depth thinking about EO uses across both expert groups* UrbanGen is rated highly by both urban EO and compliance experts (Fig. 3), with slightly stronger support for envisioning EO uses compared to understanding risks.

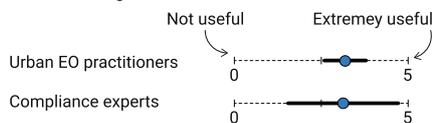
Urban EO practitioners imagined using the framework primarily at the beginning of the AI development process ( $n=9$ ) for four key tasks: landscape analysis ( $n=3$ ) to understand key stakeholders and risks, classify scientific fields, and prepare for funding proposals; project planning ( $n=3$ ) for early analysis and preparation; brainstorming ( $n=2$ ) to generate new ideas and explore unconsidered topics; and project approval ( $n=1$ ) to assist with ethics approval at the project’s start. Additionally, one participant mentioned using and updating the framework during later stages.

Compliance experts also imagined using the framework primarily at the design stage ( $n=6$ ), where it would be instrumental in project planning ( $n=3$ ) for conducting compliance assessments, identifying areas beneficial for the community, customers, and the company ( $n=2$ ), and ensuring that new inventions or proposals meet regulatory standards from the outset, even before a solution is fully developed ( $n=1$ ). In contrast to practitioners who focus on landscape analysis and brainstorming at the beginning, compliance experts also see the value of the framework in later stages ( $n=2$ ), applying it continuously depending on their role and the ongoing requirements of the project.

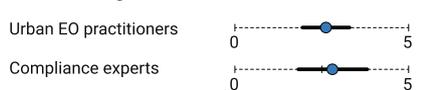
In summary, purposes such as landscape analysis, classifying scientific fields, brainstorming, and identifying areas beneficial for the company clearly demonstrate the usefulness of UrbanGen for broad thinking. Meanwhile, understanding key stakeholders, identifying risks, and

Usefulness of the framework per task and cohort

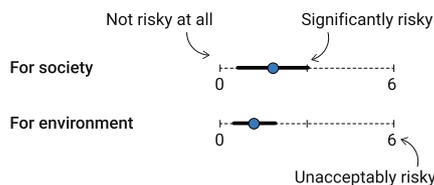
For envisioning uses



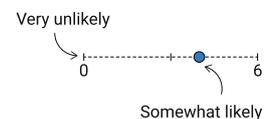
For envisioning risks



Perception of the riskiness of uses



Perception of whether the uses can fundamentally change business operations or daily life



**Fig. 3** UrbanGen is highly useful for considering EO uses broadly and in depth for both expert groups. For EO practitioners, it exposed them to risky societal uses and slightly risky environmental uses that are likely to affect business operations or daily life

ensuring new inventions meet regulatory standards highlight its usefulness for in-depth thinking as well. Specifically regarding risks, P7 stated, *'It is quite important to think about the impact that your project is going to have.'* On compliance and meeting regulatory standards, E6 remarked: *'So I guess one of the dreams for every lawyer is to have AI to do legal searches, even case law and legislation.'* Nevertheless, despite several participants noting that UrbanGen helped them think more in-depth about their specific uses, we reiterate that UrbanGen is designed solely to foster expert thinking. Final decisions and context-specific details should always be determined by humans.

*EO uses particularly benefit sustainable cities and communities* All 138 identified uses support at least one of the SDGs (Fig. 5 in Appendix 4). The five most supported SDGs by urban EO uses are SDG 11: sustainable cities and communities (41 uses), SDG 13: climate action (38 uses), SDG 9: industry, innovation, and infrastructure (22 uses), SDG 1: no poverty (19 uses), and SDG 2: zero hunger (17 uses). Examples include *mapping urban growth (use #106)*, *tracking carbon emissions (use #120)*, *detecting structural anomalies using radar (use #28)*, *identifying human trafficking networks*, and *estimating water sources from multispectral data (use #102)*.

*Upcoming EO uses present a higher risk compared to the existing ones* We identified 138 potential urban EO uses across various domains. EO experts classified the realisticness of these uses, revealing distinct differences between existing and upcoming uses. Among the *existing* uses, none were deemed prohibitively risky, and only 12% were classified as high risk. In contrast, among the *upcoming* uses, 12% were classified as prohibited, and 18% as high risk. This highlights the growing need for thorough impact assessments as urban AI and EO technologies continue to evolve, bringing novel and creative but risky urban applications.

The *prohibited* EO uses typically involve biometric identification, such as facial recognition from high-resolution imagery for surveillance, monitoring illegal activities, or population demographic analysis. According to the EU AI Act, *high risk* uses, requiring special conformity assessments, include applications in disaster management, workforce monitoring and optimization, critical infrastructure, and public safety and law enforcement. Most *low risk* uses are found in health and environmental monitoring, urban planning, commerce, economic activities, recreation, tourism, and creative applications in arts and entertainment.

We also analyzed the reasons behind the riskiness of the EO uses, drawing also from the discussions with our expert participants. These reasons are mainly linked to *privacy and surveillance issues* (e.g., capturing private objects while mapping slums with drone imagery), *data ownership and power imbalance* (e.g., most EO data is produced and owned by Western and developed countries, while the imaged areas are often in other countries), *data biases and model inaccuracies* (e.g., models trained and fine-tuned for certain cities often fail when applied to other cities), *vulnerable subjects and communities* (e.g., residents of flood-prone areas in their interactions with insurance firms), and *over-reliance on AI and decision-making detached from on-the-ground realities* (e.g., precision agriculture for urban farming using EO might suggest removing certain weed plants that actually support biodiversity and benefit farming practices).

Moreover, some experts argued that while certain uses can be risky, not acting upon them could be equally hazardous (e.g., disaster response or critical infrastructure monitoring).

*EO experts correctly identify societal and environmental risks despite their only moderate agreement* Urban EO experts rated the EO uses' risk between slightly and moderately risky, indicating slightly higher perceived risk for societal issues (Fig. 3). They also showed higher agreement on these risks (Krippendorff's Alpha of 0.481) compared to environmental risks (Krippendorff's Alpha of 0.014).

Uses perceived as extremely risky for society (scores above 5) were primarily in the domain of "Biometric identification and categorisation of natural persons," aligning with their classification as prohibited under the EU AI Act. The next four most risky uses (scores above 4) were related to military or counterterrorism and are excluded from the EU AI Act classification. Uses perceived as not risky at all for society (score of 0) were in "Arts and Entertainment" and "Education and vocational training," also classified as low risk under the EU AI Act.

The uses perceived as most risky for the environment (score of 3) involved monitoring remote work environments and migration patterns from satellite data. The least risky uses (score of 0.33) included raising awareness on climate issues and planning new transport routes.

*EO experts identify environmental monitoring EO uses as most transformational* Urban EO experts rated the potential transformational impact of uses with an average score of 4.0, indicating a somewhat likely impact (Fig. 3).

The agreement among experts was limited, with a Krippendorff's Alpha of 0.14.

The use with the highest perceived transformational impact (score of 5.67) was *estimating pollution levels from satellite data (use #114)*, followed by uses related to environmental monitoring, and natural resource management (scores of 5.33).

The uses with the lowest perceived transformational impact (score of 1.67) were in the domains of "Arts and Entertainment" and "Hobbies," such as *identifying dark areas for stargazing locations (use #126)*.

## 5 Discussion

Earth Observation (EO) data is a vital component of urban data infrastructure (Goodchild, 2022). Its significance in urban AI is expected to increase as EO spatial-temporal resolution improves (Selva & Krejci, 2012; Belward & Skøien, 2015), alongside ongoing urbanization (Kundu & Pandey, 2020; World Health Organization. Regional Office for Europe, 2017) and the digital transformation of cities (Shi et al., 2022; Guo et al., 2022).

We created a semi-automatic method, UrbanGen, which leverages LLMs to curate urban EO uses, alongside a preliminary benefit-risk assessment, and promotes responsible AI practices tailored to urban EO practitioners, as well as compliance experts in this domain. Both cohorts of experts rated UrbanGen as very useful for deliberation on the EO uses and their impact assessment. By analysing the 138 uses generated by UrbanGen, we identified that all the urban EO uses support some of the SDGs (Group on Earth Observation, Committee on Earth Observation Satellites, 2017); from promoting sustainable cities and communities (Chrysoulakis et al., 2014), to climate action (Cartalis et al., 2015) to zero hunger (Jung et al., 2021). However, at the same time, 17% of these uses were found either high risk or even prohibited as per the EU AI Act (European Commission, 2023), highlighting the need for responsible AI by design (Šćepanović et al., 2023; Tahaei et al., 2023b), i.e., from the earliest stages in the use development process (Kochupillai, 2021). For this reason, as the primary practical implication of our work, we created an *impact assessment checklist* (Fig. 4) (Stahl et al., 2023; Constantinides et al., 2024) for urban EO community, suggesting the questions to be asked even before starting the project, as well as during its development and application.

*Supporting urban EO community with impact assessment checklist* To maximize AI benefits for EO in urban environments and mitigate potential harms, we created an

impact assessment checklist (Fig. 4) for the community. This checklist is informed by the development of UrbanGen (Sect. 3) and insights from our user studies (Sect. 4). While not exhaustive, it serves as a starting point and complements existing ethical guidelines (Kochupillai et al., 2022) to promote transparency and accountability in the development and use of EO technology within this community.

The checklist comprises four sections, aligning with typical ways of reporting on AI system impacts, including intended use, risks, and mitigations (Sherman & Eisenberg, 2024; Microsoft, 2025; National Institute of Standards and Technology, 2023b).

The first section addresses the intended use (Fig. 4A), based on the EU AI Act's risk assessment requirements and a structured format for documenting system use, as employed in our prompts (Golpayegani et al., 2023). This format is comprehensive and accessible to both technical and non-technical stakeholders.

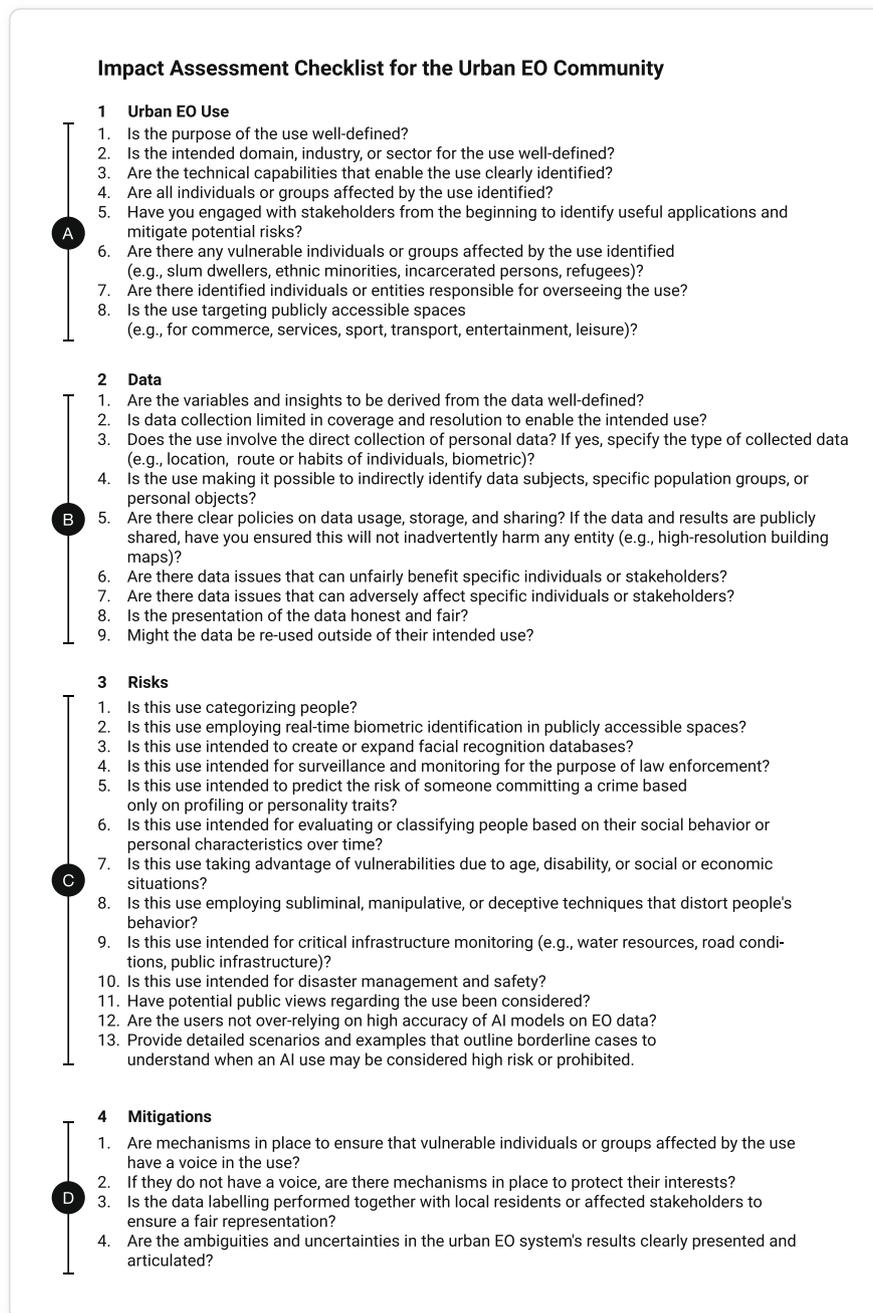
The second section focuses on data (Fig. 4B). Interviews with urban EO experts highlighted the risks of unintended privacy breaches associated with satellite and drone EO data (Vettorel, 2023; Purdy, 1999). As EO data resolution increases, it becomes suitable for facial recognition, raising concerns about prohibited uses, such as surveillance and expanding facial recognition databases.

The third section focuses on risks (Fig. 4C). Our analysis of prohibited and high-risk uses, as well as thematic analysis, revealed that urban EO uses often involve vulnerable groups (e.g., slum dwellers (Kochupillai & Taubenböck, 2023), farmers) or automatic policy-making (e.g., disaster response, land use distribution). Additionally, common EO application domains, such as disaster response and critical infrastructure monitoring, tend to be high-risk. This section formulates questions to identify and assess these risks.

Aiming to address and mitigate potential negative impacts on individuals and society, the fourth section addresses specific mitigations for the urban EO uses, relating to the types of outputs they produce, and the stakeholders affected (Fig. 4D).

Our work comes with additional two theoretical and two practical implications.

*Further implications* **Framework for integrating broad and in-depth AI evaluation.** The proposed framework, UrbanGen, demonstrates a novel method for integrating



**Fig. 4** Risk Assessment Checklist for the Urban EO Community. It helps to systematically consider an Urban Earth Observation: **A** use, **B** data, **C** risks, and **D** mitigations

both broad and in-depth thinking in the evaluation of AI applications (Bouschery et al., 2024; Lukowicz et al., 2023). The LLM method such as UrbanGen can support both the speed of mapping the landscape of an AI technology applications, such as urban EO (broad thinking), as well as uncovering subtleties and details, for which the experts would need to deliberate with other stakeholders

such as compliance experts (in-depth thinking). Using LLMs to generate diverse and realistic AI uses in EO, combined with automated literature searches, offers a new approach for comprehensive AI assessment. This dual approach can be theoretically extended to other domains, providing a robust model for ensuring that AI systems are both innovative and rigorously evaluated for

risk. However, human input for specific contexts and final decision-making are key in this framework.

**Dynamic and adaptive impact assessment tools.** The discussions on the usefulness of UrbanGen with both urban EO and compliance experts suggests that AI impact assessment tools, such as model cards (Mitchell et al., 2019) or our proposed checklist could benefit from being dynamic and adaptive, incorporating automated tools and expert validation to remain relevant and useful. For example, several among our experts suggested that as future tools, they would like to be able to see how other experts have rated some uses, or where they have employed similar AI capabilities. A further exploration and refinement of impact assessment tools (Stahl et al., 2023) by leveraging these ideas and similar approaches such as UrbanGen could enhance their applicability and utility in diverse AI contexts beyond urban EO.

**Improved risk mitigation and ethical urban EO deployment.** The UrbanGen’s ability to produce preliminary risk assessments for AI applications in EO, validated by experts, ensures that potential risks are systematically identified and evaluated (Kochupillai, 2021). This enhances the safety and ethical deployment of AI technologies in urban contexts. The provided impact assessment checklist serves as a practical tool for the Urban EO community, helping practitioners to systematically consider the ethical and societal implications of AI applications. By proactively addressing risks and benefits, the urban AI community can ensure that the deployment of AI in EO is both responsible by design (Lu et al., 2023) and beneficial (Tuia et al., 2023), minimizing potential harms and maximizing positive outcomes for urban populations.

**Informed decision-making for urban EO application development.** By listing realistic uses of AI for EO and identifying understudied areas through an automated search of scientific literature, this framework supports broad thinking and generates valuable insights. Urban AI practitioners can use this information to explore diverse and novel applications of AI in urban contexts. This helps in avoiding redundant solutions and encourages the development of innovative AI applications (Tuia et al., 2023) that address real-world urban challenges.

*Limitations* The use of LLMs presents five main challenges. First, generated uses and risks may be biased and limited to the GPT-4 training set (Luccioni et al., 2023). Enhancements could include fine-tuning or using specialized datasets (McGregor, 2021). Second, LLM hallucinations can lead to incorrect outputs (Mittelstadt et al., 2023). For example, UrbanGen identified 9% of

the uses which were deemed as unrealistic by the urban EO experts. Combining classifiers with manual checks could ensure accuracy (Mittelstadt et al., 2023). Third, LLMs may miss risky edge cases due to built-in guardrails (OpenAI, 2023a). For instance, the EO uses generated for certain of our domains could have been more risky if not for the guardrails. An example domain is *law enforcement*, where we can envision uses of drones for excessive surveillance of people, such as what happened during COVID (Couch et al., 2020). As another example, there are even riskier uses than those UrbanGen produced for the domain *military and defence*, including dynamic targeting with hypersonic missiles (Lambakis, 2024). Fourth, LLMs face the threat of being jailbroken (OpenAI, 2023a), increasing security risks (Wei et al., 2024). In our case, this translates to an adversarial possibility to envision undesired uses that are riskier than normally allowed by the guardrails. Fifth, while LLMs exhibit high accuracy in generating uses and their benefit-risk assessments, they currently lack the ability to generate creative or novel content or possess true understanding (Chakrabarty et al., 2024). On the one hand, this means UrbanGen does not propose novel uses that have not already been considered by others. On the other hand, this limitation is visible in the lower agreement rates in the SDG assessment, where human expertise was crucial for a thorough evaluation. Thus, AI should be viewed as an exploratory tool, with human experts responsible for making final decisions and evaluating the broader ethical and societal implications of each use. Finally, presenting LLM outputs may create a false sense of security (Pataranutaporn et al., 2023). Strategies like cognitive forcing functions and skill improvement can mitigate these issues (Buçinca et al., 2021).

## Appendix 1: Domains

**Table 2** We identified a list of 46 domains from the EU AI Act (European Commission, 2023) and an interactive session with our research team (N=8)

No.	Domain
1	Biometric identification and categorization of natural persons
2	Family
3	Romantic relationships and friendships
4	Health and Healthcare
5	Well-being
6	Human-Computer Interaction
7	Finance and Investment
8	Education and vocational training
9	Employment, workers management and access to self-employment

No.	Domain
10	Essential private services and public services and benefits
11	Recommender Systems and Personalization
12	Social Media
13	Sports and Recreation
14	Arts and Entertainment
15	Security and Cybersecurity
16	Marketing and Advertising
17	Agriculture and Farming
18	Entrepreneurship
19	Autonomous Robots and Robotics
20	Innovation and Research
21	Management and Operation of critical infrastructure
22	Law enforcement
23	Migration, Asylum and Border control management
24	Democracy
25	Media and Communication
26	Accessibility and Inclusion
27	Energy
28	Military and Defense
29	Administration of justice and democratic processes
30	Government Services and Administration
31	Diplomacy and Foreign Policy
32	Food Safety and Regulation
33	Crisis Management and Emergency Response
34	Humanitarian Aid
35	Transport and Logistics
36	Urban Planning
37	Counterterrorism
38	Environment and Sustainability
39	International Law Enforcement and Cooperation
40	Climate Change Mitigation and Adaptation
41	Gaming and interactive experiences*
42	Hobbies*
43	Smart home*
44	Social and Community Services*
45	Public and private transportation*
46	Interpersonal Communication*

Among these, 40 domains are from the EU AI Act, and 6 additional domains—not covered by the previous list of 40 domains—were identified during the session with our team (indicated by an asterisk \*)

coverage check module that automatically analyses 3 million Semantic Scholar papers to identify previously under-researched uses (Fig. 1).

### Listing realistic uses of AI in EO

```

Urban EO Uses Generation Prompt

SYSTEM ROLE: As a Senior Earth Observation Technology Specialist, you specialize in the latest developments in Earth Observation technology. Earth Observation involves gathering, analyzing, and presenting data about Earth's physical, chemical, and biological systems using remote sensing technologies. These observations can be made from satellites, drones, planes, or ground-based aerial platforms (ensuring instrument locations). The range of sensing instruments include optical, radar, thermal, multispectral, photographic, lidar, infrared and non-infrared cameras, sensors, and geosensors, among others. In this pivotal role, you are entrusted with reviewing and cataloguing the diverse applications and use cases of Earth Observation technology across multiple domains.

USER PROMPT: Create a comprehensive and self-explanatory list, in JSON format, detailing the various uses of Earth Observation technology. Each dictionary in the created list describes a particular use case or application of Earth Observation technology. Provide three uses for each of the {} domains listed below. The uses must contain specific details about how the technology is used, by using action verbs that clearly describe the actions, activities, or processes of the uses. The level of specificity should be consistent across all uses. For each of these uses, you must output the following 6 elements each in less than 7 words:
(1) Use: An element of a series of numbered uses, starting with 1. Each use should be listed consecutively.
(2) Domain: The domain that represents the area or sector the AI system is intended to be used in.
(3) Purpose: The purpose or objective that is intended to be accomplished by using an AI system.
(4) Capability: The capability of the AI system that enables the realization of its purpose and reflects the technological capability.
(5) Sensing instrument and location: The sensing technology that enables observation, i.e., gathering of information about the physical, chemical, and biological systems of the planet Earth (e.g., radar, optical), and where it is placed (e.g. satellite, drone).
(6) AI User: The entity or individual in charge of deploying and managing the AI system, including individuals, organizations, corporations, public authorities, and agencies responsible for its operation and management.
(7) AI subject: The individual directly affected by the use of the AI system, experiencing its effects and consequences. They interact with or are impacted by the AI system's processes, decisions, or outcomes.
Ensure that each concept is specific and easy to understand for non-experts. Avoid duplicate purposes or objectives and use clear and precise language to describe the uses' concepts.

Domains to be included are the following:
{"}

For some domains, e.g., Human-Computer Interaction, Well-being, and Social Media, it might be more challenging to identify Earth Observation uses and you might think of wearable sensors or cameras on computer screens, but do NOT. Instead, review your use and double check that the sensors you are outputting are indeed Earth Observation sensors. Try to be creative with Earth Observation uses even in these less obvious domains of its application. For example, for Well-being, you can think of suggesting pairs for health-promoting activities to the user, and for Human-Computer Interaction you can think of applications bringing satellite data closer to the public. Double-check that you are outputting realistic, i.e., plausible, meaningful, and useful uses. It is OK to output upcoming, i.e., uses that are not yet widespread but could be in research, development, or talked about in expert circles. However, try to make sure that it also makes sense to use satellite or aerial data for the use you are describing; and that it is not much more straightforward to use some other sensors, if it is much easier to achieve the same purpose with some other already existing sensors (e.g., phone cameras or wearable devices), then think of another use.

For the "Capability", write it by combining action verbs in gerund form (i.e., ending with "ing"), inferences and data, entity or metric.
(1) Action verbs clearly describe the actions, activities, or processes taken by the AI system, e.g., identify. Choose the most suitable action verb from the following list. If none can be assigned, propose a new verb, and mark it with an asterisk.
(A) Extracting (e.g., Mining, Gracing, Measuring, Assessing)
(B) Forecasting (e.g., Predicting, Guessing, Speculating)
(C) Comparing (e.g., Ranking, Ordering, Finding Best, Finding Cheapest, Recommending)
(D) Detecting (e.g., Monitoring, Sensing, Noticing, Classifying, Discriminating)
(E) Identifying (e.g., Recognizing, Discerning, Finding, Classifying, Perceiving)
(F) Discovering (e.g., Extracting, Noticing, Organizing, Clustering, Grouping, Connecting, Revealing)
(G) Generating (e.g., Making, Composing, Constructing, Creating, Authoring)
(H) Acting (e.g., Doing, Executing, Playing, Doing, Learning, Operating)
(2) Inference clearly describes the output or conclusion drawn by the AI system based on the data it processes, e.g., crop yield, floods, trend, anomaly, wildfires, pattern, and probability.
(3) Data, Entity or Metric clearly describes the source, type, or nature of the data used by the AI system, e.g., from an optical camera, from an infrared camera, user input, sensor readings, transaction records, biometric data, environmental data, social media posts, geographical information, medical records, and financial metrics.

For "Purpose", write it also in a gerund verb form (i.e., ending with "ing").
*** Ensure to output 3 uses per each of the {} domains above, i.e., {} uses in total. ***

Follow this example structure for reporting the identified uses:
{
  {{
    "Use": 1,
    "Domain": "Biometric identification and categorization of natural persons",
    "Purpose": "Identifying individuals in crowded areas",
    "Capability": "Facial recognition from aerial footage",
    "Sensing instrument and location": "Optical cameras on drones",
    "AI User": "Law enforcement agencies",
    "AI Subject": "Individuals in public spaces"
  }},
  {{
    "Use": 2,
    "Domain": "Health and Healthcare",
    "Purpose": "Monitoring disease spread",
    "Capability": "Derive environmental parameters that influence the emergence and spreading of diseases",
    "Sensing instrument and location": "Optical satellites",
    "AI User": "Public health agencies",
    "AI Subject": "People"
  }},
  ...
  {{
    "Use": 44,
    "Domain": "Education and vocational training",
    "Purpose": "Realistic geography education using virtual reality or augmented reality",
    "Capability": "Mapping and presenting urban and rural terrain",
    "Sensing instrument and location": "Multispectral and optical satellites",
    "AI User": "Educators",
    "AI Subject": "Students"
  }},
  ...
  ...
  ...
}
*** Ensure to output only the correctly formatted JSON and nothing else. ***
}
    
```

## Appendix 2: UsesGen methodology

Our framework, UrbanGen, consists of three LLM prompts: Prompt 1, designed to list diverse and realistic uses of EO technology; Prompt 2, provided with the relevant text from the EU AI Act to assess the risk associated with the output from Prompt 1; Prompt 3, given the text of the official UN definitions of SDGs and their targets to assess whether a given use supports each of the targets/respective goals; and an automatic literature

To list realistic uses of EO technology, which will allow for a further risk assessment, we followed a three-step process. First, we resorted to previous literature giving guidance as to how technology uses should be described at a minimum so that their risk level according to the EU AI Act can be reliably assessed (Golpayegani et al., 2023). Second, to generate a diverse set of uses, we curated a list of 46 relevant domains covering all the five levels of the Social-Ecological Model (Golden et al., 2015): individual,

interpersonal, institutional, community, and public policy. Third, we engineered an LLM prompt for generating a list of uses in the predefined format and across the curated set of domains, while employing the Chain-of-Thought (CoT) reasoning (Wei et al., 2022b) in the process to minimise generating unrealistic uses.

**Defining a format for describing an AI use.** To allow for the risk assessment as per EU AI Act, previous work (Golpayegani et al., 2023) showed it is necessary to break down a technology use into five risk concepts:

*Domain:* “The area or sector the AI system is intended to be used in” (e.g., urban planning).

*Purpose:* “The objective that is intended to be accomplished by using an AI system” (e.g., mapping urban growth).

*Capability:* “Technological capability that enables the realization of the purpose” (e.g., object and pattern recognition).

*AI deployer:* “The entity or individual in charge of deploying and managing the AI system, including individuals, organizations, corporations, public authorities, and agencies responsible for its operation and management” (e.g., city governments).

*AI subject:* “The individual directly affected by the use of the AI system, experiencing its effects and consequences. They interact with or are impacted by the AI system’s processes, decisions, or outcomes” (e.g., city residents).

In our initial experiments, we also found out that the prompt was sometimes outputting incorrect or inconsistent combinations of the EO sensing technology and its placement (e.g., placed GPS on drones to track moving objects instead of placing it directly on the object). For this reason, we introduced another concept, specific to EO technology uses, which would serve as *cue* to the LLM and guide it to output technically feasible uses:

*Sensing instrument and location:* “The sensing technology that enables observation, i.e., gathering of information about the physical, chemical, and biological systems of the planet Earth (e.g., radar, optical), and where it is placed (e.g., satellite, drone).”

Requiring Prompt 1 to output the uses along with their corresponding risk concepts information will allow the subsequent Prompt 2 (in Sect.7.2) to reliably evaluate the use’s risks.

**Curating a set of relevant and diverse domains.** We systematically curated the domains starting from the EU AI Act, in which 8 domains are explicitly listed in Annex III and referred to in Article 6(2) (European Commission, 2023) (e.g., “Education and vocational training”). We then reviewed the full text of the EU AI Act and its Amendments (European Commission, 2023) to add 32 domains briefly mentioned but not explicitly

listed. For instance, we added the “Social Media” domain based on Amendment 51, stating: “*The indiscriminate and untargeted scraping of biometric data from social media [...] adds to the feeling of mass surveillance [...]*” Additionally, we conducted a focus group using a think-aloud protocol ( $N=8$ ) to ensure all significant domains beyond the EU AI Act were captured. The group consisted of our research team (3 F, 5M, mean age: 31.8,  $SD$ : 6.74, range: 22-45). Using a Miro board, participants identified domains affecting their lives across five levels of the Social-Ecological Model (Golden et al., 2015). This resulted in 6 more domains, completing our final list of 46 diverse domains listed in Appendix 2, Table 2.

**Engineering an LLM prompt to generate the uses of AI for EO.** To generate a list of uses along with risk concepts, we divided Prompt 1 into five parts:

*System role:* This is the designated function, characteristics, or responsibilities assigned to the LLM. This has been shown to improve the quality of the output, as it allows to generate content from specific perspectives, accommodating diverse stakeholder viewpoints or specific role requirements (Giray, 2023; OpenAI, 2023a). We specified the role of a *Senior Earth Observation Technology Specialist* and described its responsibilities as *reviewing and cataloguing the diverse applications and use cases of Earth Observation technology across multiple domains* (which are expanded in the “Domains” input described below).

*Multi-part instructions:* These are generally commands that instruct the LLM to complete two or more tasks in a specific order. In our case, we formulated two tasks. The first task was to create a JSON list of three AI technology uses for each domain, described in a format that facilitates risk assessment. The second task was to ensure that the use descriptions were formatted correctly. We requested that they must contain exactly six risk concepts, each precise, easy to understand for non-experts, and expressed under 7 words.

*Six Concepts:* We instructed the LLM what counts as a use by providing it with the definitions of the five risk concepts as specified above (i.e., domain, purpose, capability, AI uses, and AI system), along with the concept specific to EO technologies (i.e., sensing instrument and location). These concepts guided the LLM to generate realistic and meaningful uses.

*Domains:* These are specific areas of knowledge or expertise that LLMs can access, guiding them to generate more specialised, sophisticated, and nuanced text. We equipped the LLM with a list of 46 domains introduced above.

*Examples:* These are a specific instance of few-shot learning (Brown et al., 2020) used to achieve the desired





*microsoft/mpnet-base* model. We selected this model due to its balance between effectiveness (over 14 million downloads last month) and being lightweight (only 109 million parameters), enabling us to embed a large number of articles efficiently.

For each paper  $p$ , we paired it with the highest similarity  $maxs_p$  use  $u$ , measured using cosine similarity of their embeddings  $v_p$  and  $v_u$ :

$$maxs_p = \max_u \text{sim}(v_p, v_u),$$

where  $p$  is the paper,  $u$  is a use, and  $maxs_p$  is the highest similarity of the paper embeddings with any of the uses' embeddings. We then manually determined the similarity threshold across  $maxs_p$  that would ensure the paper's abstract indeed discussed the associated use  $u$ . Specifically, we experimented with various thresholds on  $maxs_p$  and examined the paper-use pairs for each. If a certain threshold included pairs where the paper did not discuss the matched use, we increased the threshold for stricter matching. After testing thresholds at the 95<sup>th</sup>, 99<sup>th</sup>, 99.5<sup>th</sup>, and 99.9<sup>th</sup> percentiles of  $maxs_p$ , we found that the 99.9<sup>th</sup> percentile yielded 3,295 paper-use pairs, so that the paper indeed discussed the paired EO use, and so we chose this threshold. In this way, each EO use was associated with 0 or more papers that discussed it. If the use was associated with 0 papers, we classified it as potentially under-researched in the literature.

### Appendix 3: List of urban EO uses

Urban EO Uses, their risk classification (E: excluded, HRE: high-risk exception, P: prohibited, HR: high risk, and LLR: limited or low risk), and relevant EU AI Act text.

ID	Description	Risk	EU AI Act Text
1	Identifying individuals in crowded areas using facial recognition by police	P	Article 5(1)(h): Real-time remote biometric identification for law enforcement
2	Tracking movement patterns of individuals from drone images	P	Article 5(1)(g): Biometric categorisation
3	Categorizing population demographics by visual traits	P	Article 5(1)(g): Biometric categorisation
4	Monitoring family safety during disasters by detecting locations	HR	Annex III, Section 5(d)
5	Tracking family outdoor activities by monitoring movements in parks	LLR	N/A
6	Ensuring safe travel routes by identifying safe paths from satellite data	LLR	N/A

ID	Description	Risk	EU AI Act Text
7	Suggesting romantic outdoor spots by recommending scenic locations	LLR	N/A
8	Planning group outdoor activities by finding suitable locations	LLR	N/A
9	Monitoring group safety in events by tracking group movements	LLR	N/A
10	Monitoring disease spread by deriving environmental parameters	LLR	Article 6(3a)
11	Tracking air quality by measuring pollutants from satellite data	LLR	Article 6(3a)
12	Identifying heatwave risks by detecting temperature anomalies	LLR	Article 6(3a,d)
13	Recommending parks for activities by finding green spaces	LLR	Article 6(3a)
14	Monitoring urban noise levels by detecting noise pollution	LLR	Article 6(3a)
15	Tracking outdoor exercise patterns by monitoring activity hotspots	LLR	N/A
16	Visualizing environmental changes by presenting satellite data	LLR	N/A
17	Enhancing virtual reality experiences by integrating real-world data	LLR	Article 6(3a)
18	Improving map applications by updating geographical data	LLR	N/A
19	Assessing agricultural investments by estimating crop yields	LLR	N/A
20	Evaluating real estate by analyzing land use patterns	LLR	N/A
21	Predicting market trends by forecasting economic activities	LLR	N/A
22	Teaching environmental science by providing real-time data	LLR	Article 6(3a)
23	Enhancing geography lessons by mapping terrains	LLR	Article 6(3a)
24	Training in disaster management by simulating disaster scenarios	HR	Annex III, Section 3
25	Monitoring remote work environments by tracking outdoor activities	HR	Annex III, Section 4(b)
26	Ensuring worker safety by detecting hazardous conditions	HR	Annex III, Section 4(b)

ID	Description	Risk	EU AI Act Text	ID	Description	Risk	EU AI Act Text
27	Optimizing field operations by analyzing work patterns	HR	Provisions 57 and Annex III, Section 4(b)	47	Analyzing foot traffic in retail areas by detecting human movement	LLR	N/A
28	Monitoring infrastructure health by detecting structural anomalies	HR	Annex III, Section 2	48	Optimizing billboard placements by comparing visibility metrics	LLR	N/A
29	Improving public transportation by analyzing traffic patterns	LLR	N/A	49	Monitoring crop health by detecting plant stress	LLR	N/A
30	Managing water resources by monitoring water bodies	HR	Annex III, Section 2	50	Predicting harvest times by forecasting crop maturity	LLR	Article 6(3a)
31	Suggesting travel destinations by recommending scenic spots	LLR	Article 6(3a)	51	Managing irrigation systems by estimating soil moisture	LLR	N/A
32	Personalizing outdoor activities by finding suitable locations	LLR	N/A	52	Identifying new business opportunities by discovering market trends	LLR	Article 6(3a)
33	Recommending agricultural practices by analyzing soil conditions	LLR	N/A	53	Assessing real estate investments by estimating property values	LLR	N/A
34	Sharing environmental insights by visualizing satellite data	LLR	N/A	54	Optimizing supply chain routes by forecasting traffic patterns	LLR	N/A
35	Tracking environmental campaigns by monitoring campaign activities	LLR	N/A	55	Navigating autonomous agricultural robots by identifying field boundaries	HRE	Annex I, Section B (safety components in agricultural vehicles)
36	Promoting eco-friendly practices by identifying green initiatives	LLR	N/A	56	Mapping disaster zones for rescue robots by generating 3D terrain models	LLR	N/A
37	Planning outdoor sports events by finding suitable locations	LLR	N/A	57	Inspecting infrastructure with drones by detecting structural anomalies	LLR	N/A
38	Monitoring sports facilities by detecting facility conditions	LLR	Article 6(3a)	58	Studying climate change effects by comparing historical and current data	E	Article 2(6)
39	Tracking outdoor activities by monitoring activity hotspots	LLR	N/A	59	Exploring new energy sources by discovering geothermal sites	E	Article 2(6)
40	Creating virtual landscapes by generating realistic terrains	LLR	N/A	60	Analyzing urban heat islands by detecting temperature variations	LLR	N/A
41	Enhancing film production by providing real-world backdrops	LLR	N/A	61	Monitoring power grid health by detecting faults	HR	Annex III, Section 2
42	Creating art installations by using satellite imagery	LLR	N/A	62	Inspecting water pipelines by identifying leaks	HR	Annex III, Section 2
43	Monitoring border security by detecting unauthorized crossings	LLR	N/A	63	Assessing road conditions by estimating wear and tear	HR	Annex III, Section 2
44	Identifying illegal activities by recognizing suspicious patterns	P	Article 5(1)(dh)	64	Monitoring illegal activities by detecting suspicious behavior	HR	Annex III, Section 6
45	Preventing national cyber threats by monitoring infrastructure	E	Article 2(3)	65	Tracking stolen vehicles by identifying vehicle movements	LLR	N/A
46	Targeting ads based on weather patterns by forecasting consumer behavior	LLR	N/A	66	Surveilling large public events by monitoring crowd dynamics	HR	Annex III, Section 6
				67	Monitoring border crossings by detecting unauthorized entries	HR	Annex III, Section 7

ID	Description	Risk	EU AI Act Text
68	Tracking refugee movements by identifying migration patterns	LLR	N/A
69	Assessing refugee camp conditions by estimating population density	LLR	N/A
70	Monitoring election rallies by detecting crowd sizes	HR	Annex III, Section 8
71	Ensuring fair voting processes by monitoring polling stations	LLR	N/A
72	Analyzing political campaign reach by comparing rally attendance	LLR	N/A
73	Reporting natural disasters by detecting disaster impacts	LLR	Article 6(3a)
74	Creating visual content by generating maps	LLR	Article 6(3a)
75	Analyzing environmental changes	LLR	Article 6(3a)
76	Mapping accessible routes by identifying terrain features	LLR	N/A
77	Designing inclusive public spaces by estimating foot traffic	LLR	Article 6(3a)
78	Monitoring air quality by detecting pollution levels	LLR	N/A
79	Identifying renewable energy sites by discovering solar potential	LLR	N/A
80	Monitoring oil spills by detecting spills	LLR	N/A
81	Assessing wind farm efficiency by estimating wind speeds	LLR	N/A
82	Monitoring troop movements by detecting military activity	E	Article 2(3)
83	Identifying strategic locations by discovering terrain features	E	Article 2(3)
84	Assessing battlefield conditions by estimating damage	E	Article 2(3)
85	Monitoring environmental law compliance by detecting illegal activities	LLR	Article 6(3a)
86	Ensuring fair land use by comparing land use patterns	LLR	N/A
87	Tracking deforestation activities by detecting forest cover changes	LLR	N/A
88	Planning urban development by generating land use maps	LLR	Article 6(32a)

ID	Description	Risk	EU AI Act Text
89	Monitoring public infrastructure by detecting structural issues	HR	Annex III, Section 2
90	Managing natural resources by estimating resource availability	LLR	N/A
91	Monitoring border activities by detecting unauthorized crossings	HR	Annex III, Section 7
92	Assessing treaty compliance by monitoring military installations	E	Article 2(3)
93	Tracking deforestation agreements by detecting forest cover changes	LLR	N/A
94	Monitoring crop health by assessing vegetation indices	LLR	N/A
95	Detecting illegal fishing by identifying fishing vessels	HR	Annex III, Section 6
96	Ensuring food supply chain integrity by tracking crop transport routes	LLR	Article 6(3a)
97	Assessing disaster damage by estimating affected areas	LLR	N/A
98	Predicting flood risks by forecasting water levels	LLR	N/A
99	Coordinating relief efforts by mapping affected regions	LLR	Article 6(3a)
100	Identifying refugee camps by detecting temporary settlements	LLR	N/A
101	Monitoring food distribution by tracking supply routes	LLR	Article 6(3a)
102	Assessing water availability by estimating water sources	LLR	N/A
103	Optimizing shipping routes by forecasting weather patterns	LLR	N/A
104	Monitoring traffic congestion by detecting vehicle density	LLR	N/A
105	Tracking cargo movements by identifying transport vehicles	LLR	N/A
106	Mapping urban growth by detecting land use changes	LLR	N/A
107	Designing green spaces by identifying suitable areas	LLR	N/A
108	Monitoring infrastructure development by tracking construction progress	LLR	Article 6(3a)

ID	Description	Risk	EU AI Act Text
109	Identifying terrorist hideouts by detecting suspicious activities	E	Article 2(3)
110	Monitoring border security by tracking movements	HR	Annex III, Section 7
111	Detecting illegal arms shipments by identifying suspicious cargo	E	Article 2(3)
112	Monitoring deforestation by detecting forest cover changes	LLR	N/A
113	Tracking wildlife habitats by identifying habitat changes	LLR	N/A
114	Assessing air quality by estimating pollution levels	LLR	Article 6(3a)
115	Tracking drug trafficking routes by identifying suspicious movements of individuals	P	Article 5(1)(h)
116	Monitoring illegal mining by detecting mining activities	LLR	N/A
117	Identifying human trafficking networks through suspicious movements	HR	Annex III, Section 6
118	Monitoring glacier retreat by detecting ice cover changes	LLR	N/A
119	Assessing sea level rise by estimating coastal changes	LLR	Article 6(3a)
120	Tracking carbon emissions by detecting pollution sources	LLR	N/A
121	Creating realistic game environments by generating terrain models	LLR	N/A
122	Simulating real-world scenarios by creating virtual landscapes	LLR	N/A
123	Enhancing virtual reality experiences by mapping real-world locations	LLR	N/A
124	Planning hiking routes by identifying trails	LLR	N/A
125	Tracking bird migrations by monitoring bird habitats	LLR	N/A
126	Mapping stargazing locations by identifying dark sky areas	LLR	N/A
127	Optimizing solar panel placement by assessing sunlight exposure	LLR	N/A
128	Monitoring home energy usage by estimating energy consumption	LLR	N/A
129	Enhancing home security by detecting intrusions	LLR	N/A

ID	Description	Risk	EU AI Act Text
130	Planning community gardens by identifying suitable areas	LLR	N/A
131	Monitoring urban heat islands by detecting temperature variations	LLR	Article 6(3a)
132	Assessing public park usage by estimating visitor numbers	LLR	N/A
133	Monitoring public transit systems by tracking vehicle movements	LLR	N/A
134	Optimizing traffic flow by analyzing congestion patterns	LLR	N/A
135	Planning new transport routes by identifying optimal paths	LLR	Article 6(3a)
136	Sharing environmental changes by generating reports	LLR	Article 6(3a)
137	Raising awareness on climate issues by creating visual content	LLR	N/A
138	Promoting sustainable practices by sharing data insights	LLR	N/A

**Appendix 4: Validating UrbanGen**

To validate UrbanGen, we answered four questions:

Q1 Did our framework generate realistic uses?

Answered by having three Urban EO practitioners categorize each use as existing, upcoming, or unlikely, with the majority label being assigned when at least two practitioners agreed.

Q2 Did our framework accurately identified under-researched uses?

Answered by having two authors independently search and review the top 10 Google Scholar articles for each under-researched use.

Q3 Did our framework correctly classify the riskiness of the uses according to the EU AI Act?

Answered by providing compliance experts with descriptions, LLM classifications, and justifications for each use, followed by measuring the agreement between the LLM and the experts on both classification and justification. Each use was evaluated by three compliance experts, with the majority label taken when at least two experts agreed.

Q4 Did our framework accurately evaluate the benefits of the uses?

Answered by having three Urban EO experts select all SDGs that each use supports and comparing them to similar uses supporting SDGs found in the literature.

*Answer to Q1: UrbanGen generated realistic uses* Each use was evaluated by three urban EO experts, with the majority label taken when at least two experts agreed.

In this way, 126 uses (91%) were deemed realistic. Within this subset, 77 uses (56% of the total) were identified as already existing (e.g., *forecasting water levels from radar data in flood-prone communities, use #98* or *generating land use maps from satellite data, use #88*). The remaining 49 uses (35%) were categorized as upcoming (e.g., *optimizing supply chain routes by forecasting traffic, use #54* or *identifying suspicious cargo from satellite imagery, use #111*).

Conversely, 12 uses (9%) were considered unrealistic. These unrealistic labels were more common in domains with fewer EO applications (e.g., “Romantic relationships and friendship,” “Family,” or “Smart home”). Inaccuracies in these areas stemmed from the LLM suggesting impractical or unnecessary uses (e.g., *suggesting romantic outdoor spots, use #7* or *monitoring family movements in parks, use #5*) and exaggerating EO capabilities (e.g., *classifying individuals from satellite data by visual traits, use #3* or *detecting home intrusions from satellite imagery, use #129*). These domains were included to provide a comprehensive set of AI applications and might be significant for other AI technologies (such as wearables), where hallucinations are less likely.

Interestingly, two uses labeled as unrealistic by two of the three annotators are referenced in recent literature, namely *monitoring remote work environments using drones* (Asadzadeh et al., 2022) and *creating art installations using satellite imagery* (Parks, 2002; Jackson, 2018). This highlights the rapid development of technology, making it challenging even for experts to stay updated.

This rapid development partly explains the low inter-annotator agreement among the experts, with a Fleiss’ Kappa of 0.19. Similar results have been reported in other studies with AI experts, such as when assessing

AI use’s risk level (Wang et al., 2024). To understand the specific reasons for the low agreement in our study, we examined the types of uses on which experts agreed versus those where they did not. Agreement was high for well-known EO application domains, such as environmental monitoring, urban development, border security and migrant monitoring, and climate studies. Conversely, there was less agreement on the realism of uses in security and law enforcement, fitness and health, energy exploration, finance and market analysis, and arts and entertainment. Uses of EO data in these domains are less commonly discussed and promoted, possibly leading to a lack of familiarity among some experts. Additionally, uses in these domains were more likely to be hallucinated. Overall, this shows that the increased task complexity and epistemic uncertainty affected the annotations (Jiang et al., 2022; Salimzadeh et al., 2024).

*Answer to Q2: UrbanGen identified seven under-researched uses* Table 3 presents 7 uses identified as under-researched, based on our analysis of Semantic Scholar abstracts.

For 3 uses (*use #43, use #85, and use #117*), our Google Scholar search yielded recent articles discussing these uses warranting removing them from the under-researched category. A likely reason for our automatic coverage check not marking these uses as covered by research is that the papers in which they are discussed were not among those having the abstract in the Semantic Scholar dump that we analysed (i.e., 3M out of 200M papers in total).

However, for the remaining 4 (57%) uses, we did not identify any research papers discussing these specific uses, confirming them as under-researched. For instance, while articles discussing analyzing foot traffic from drone imagery were found, none pertained to retail areas (*use #47*). Similarly, research on monitoring crowds from optical remote sensing imagery did not cover election rallies (*use #70*). Given that specific application contexts determine particular risks and benefits of AI technology, the identified uses indeed represent under-researched areas in previous literature, underscoring the need for more studies in these specific urban EO contexts.

**Table 3** Urban EO uses identified by our automatic literature coverage check as under-researched: 4 out of 7 are confirmed as such by the author’s manual validation

Use ID	Use description	Article
43	monitoring border security by detecting unauthorized crossings using radar satellites	Bolakis et al. (2021)
46	targeting ads based on weather patterns by forecasting consumer behavior from weather data using optical satellites	-
47	analyzing foot traffic in retail areas by detecting human movement from aerial imagery using optical drones	-
70	monitoring election rallies and detecting crowd sizes from aerial views using optical drones	-
71	monitoring polling stations from aerial imagery using optical drones to ensure fair voting processes	-
85	monitoring environmental law compliance by detecting illegal activities from satellite data using optical satellites	Patias et al. (2020)
117	identifying human trafficking networks by tracking suspicious movements from satellite data for law enforcement	Bamigbade et al. (2024)

*Answer to Q3: UrbanGen predominantly generated correct risk assessments* Each use was evaluated by three compliance experts, with the majority label taken when at least two experts agreed.

The expert annotations aligned with the LLM’s for 93% of the uses, with the Cohen’s kappa of 0.80 across the five possible risk labels (substantial agreement). Example cases where they disagreed include certain uses labelled as “Limited or Low Risk” by the LLM because the “AI system is performing a narrow and procedural task as per Article 6(2a).” For example, *detecting crowd sizes and monitoring election rallies, use #70, or tracking stolen vehicles, use #65*. However, for *use #70*, two of the annotators, instead, selected the choice of “Insufficient information to assess,” elaborating further: (LE7) “*This is risky because it may deter people from going to polling stations in order to vote. Voting rates are already very low across Europe.*” and (LE4) “*Could be also used for political influencing? The risk would really depend on how it is used.*” For *use #65*, two of the annotators, instead, selected the option “High Risk,” elaborating further: (LE7) “*This would basically require tracking of all vehicles — it depends if it is identifying vehicle registration and, in turn, owners.*”, and (LE2) “*In*

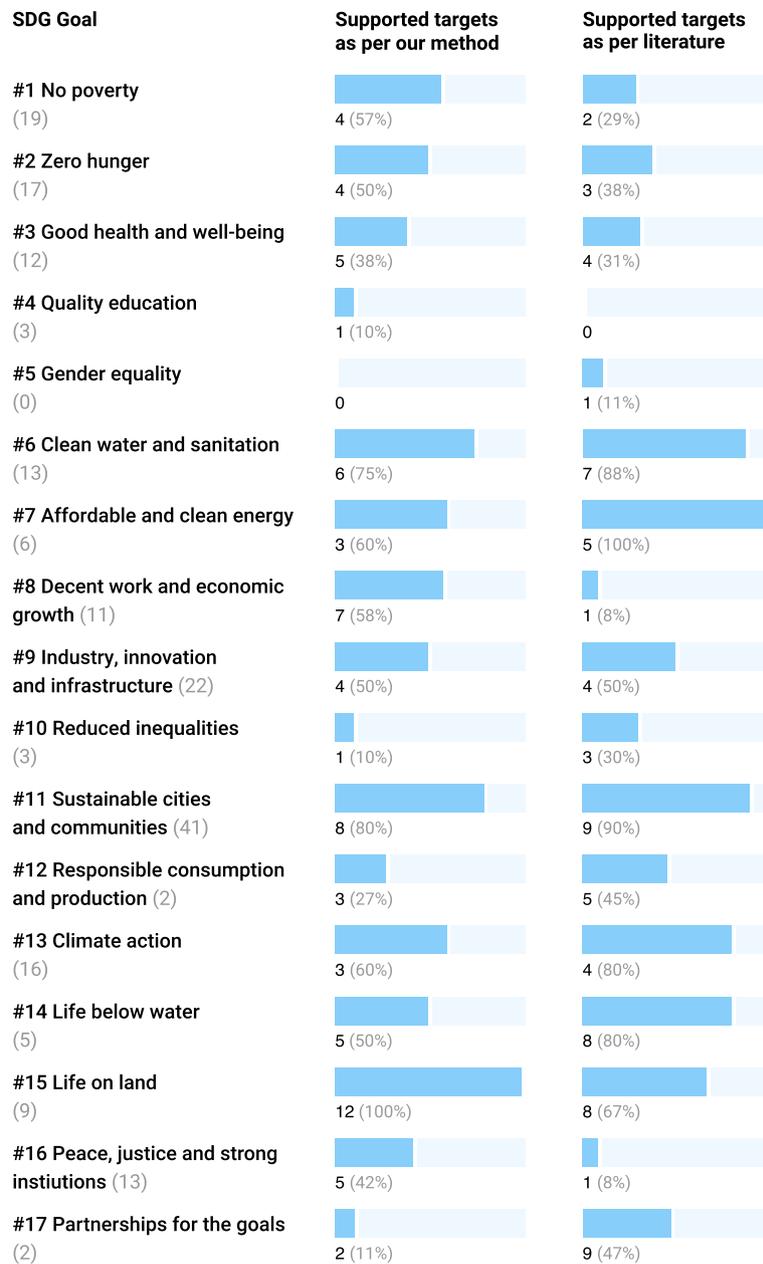
*my opinion, this could be used also to track individuals. If we would talk about the tracking of a container, ok, but a vehicle?*” However, there were also instances of uses labelled as “High Risk” by the LLM, for which the experts thought they are of low risk. For example, *detecting hazardous conditions for workers using thermal drones, use #26*, was labelled as high risk by the LLM as per the Annex III, 4(b): “AI intended to be used to make decisions affecting terms of the work related relationships, promotion and termination of work-related contractual relationships.” The experts, instead, marked it as low risk saying: (LE11) “*This information would benefit the health and safety without accessing any personal information,*” (LE1) “*This aims to record objective conditions in a space. Nothing about individuals. Seems low-risk.*”

The inter-annotator agreement among the annotators across the five possible risk labels was moderate, as indicated by a Fleiss’ Kappa of 0.52. This level of agreement, which is lower than the alignment between the majority label and the LLM, suggests a potential over-reliance on the LLM labels, likely influenced by the well-documented anchoring effect (Furnham & Boo, 2011; Nourani et al., 2021). It is plausible that this effect could have been mitigated if the experts were provided only with the use information, without being exposed to the LLM labels, and were asked to generate their own annotations. However, considering the necessity of referencing the EU AI Act’s text to justify the classification for each specific use context, this approach was deemed prohibitively time-consuming and impractical.

*Answer to Q4: UrbanGen identified fewer but accurate SDG-promoting uses compared to Urban EO experts* UrbanGen identified 216 (use, SDG) pairs where the use promotes an SDG, significantly fewer than the 643 pairs identified by experts when considering all uses listed by them. This means the LLM covered only 27% of the expert-identified pairs. However, the LLM demonstrated high precision, with 80% of its assessments aligning with expert evaluations.

When focusing only on SDGs labeled by at least two of the three experts for each use, the experts identified 262 (use, SDG) pairs, of which 44% were also identified by the LLM. In this scenario, the LLM’s precision was 54%.

Moreover, while it was prohibitive to ask experts to evaluate the support for specific targets within SDGs due to time constraints, the LLM could perform this task.



**Fig. 5** The number of targets EO supports according to our framework and to the UN report for each goal. The total number of supporting uses is shown in parentheses right below the goal's name

Appendix 5: Materials used during user studies

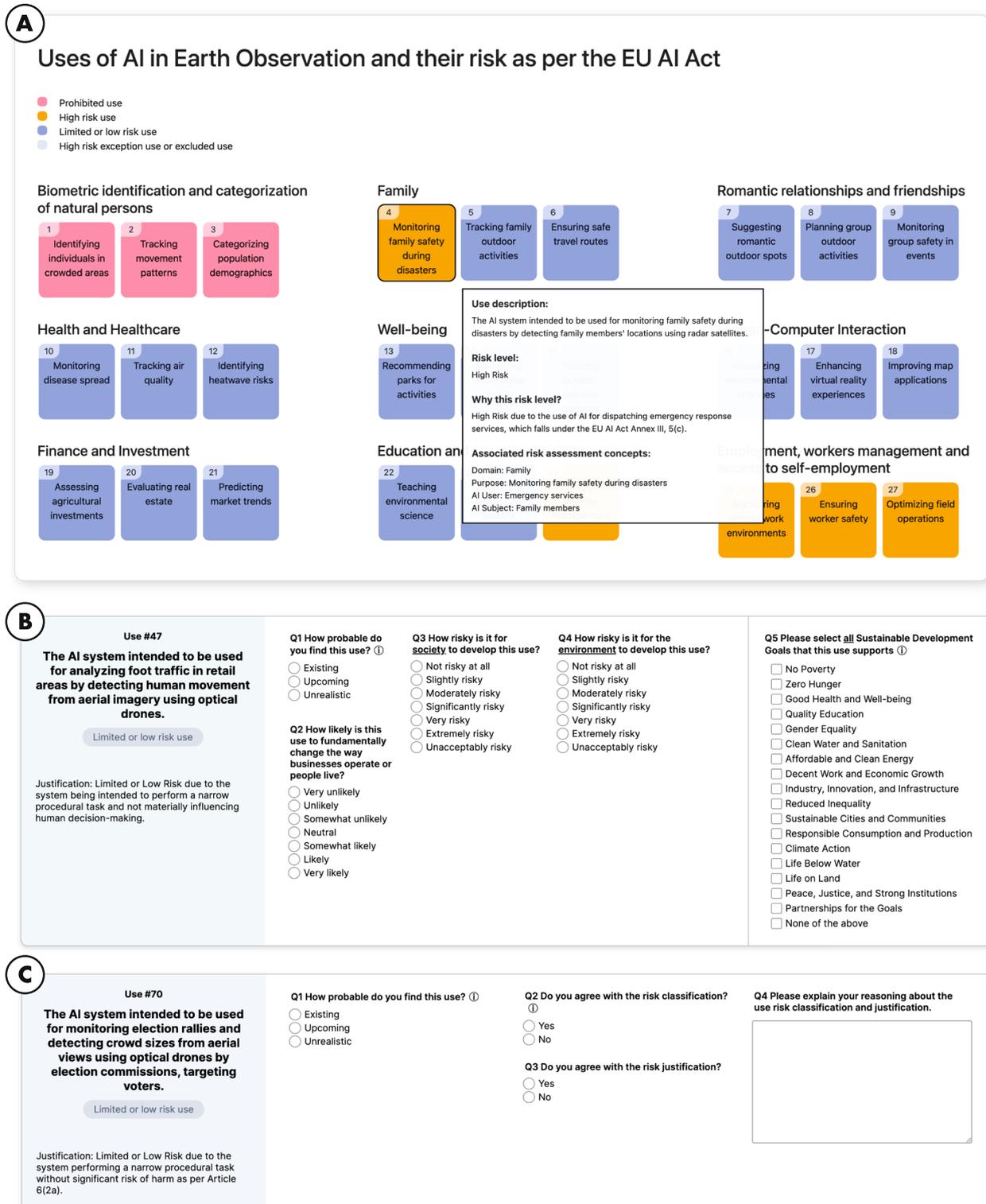


Fig. 6 Materials used during user studies with EO practitioners and compliance experts. We showed both participant groups an interactive list of 138 uses (A), followed by 46 interactive assessment cards for these uses (B, C). EO practitioners interacted with a more complex version of the card (B), while compliance experts used a simplified version of the cards (C), including the LLM-derived risk label and its justification

**Acknowledgements**

The authors sincerely thank all the study participants for evaluating our Urban EO uses and sharing their experiences and ideas.

**Authors' contributions**

SŠ conceptualized the problem, designed the experiments, conducted the user studies, prepared data for figures, and wrote the manuscript. EB conceptualized the problem, designed the experiments, prepared figures and materials for user studies, helped conduct the user studies, and contributed to writing and editing the manuscript. DQ conceptualized the problem, supervised the project, and contributed to writing and editing the manuscript.

**Funding**

The authors acknowledge that they received no funding in support for this research.

**Data availability**

All the uses produced by UrbanGen, and their risk-benefit assessments are shared in our public data repository <https://figshare.com/s/a36f17a385be85abbfb0>. We will also share our repository with the code <https://github.com/sanja7s/AI-in-the-City> for LLM prompting upon the publication.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

Received: 1 July 2024 Revised: 8 October 2024 Accepted: 24 April 2025  
Published online: 03 June 2025

**References**

Antropov, O., Rauste, Y., Häme, T., & Praks, J. (2017). Polarimetric alos palsar time series in mapping biomass of boreal forests. *Remote Sensing*, 9(10), 999

Asadzadeh, S., de Oliveira, W. J., & de Souza Filho, C. R. (2022). Uav-based remote sensing for the petroleum industry and environmental monitoring: State-of-the-art and perspectives. *Journal of Petroleum Science and Engineering*, 208, 109633

Ashurst, C., Hine, E., Sedille, P., & Carlier, A. (2022). AI ethics statements: analysis and lessons learnt from neurips broader impact statements. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency* (pp. 2047–2056). New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/3531146.3533780>

Azure, M. (2023). Introduction to prompt engineering. <https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/prompt-engineering#best-practices>. Accessed 11 May 2024

Bamigbade, O., Sheppard, J., & Scanlon, M. (2024). Computer vision for multimedia geolocation in human trafficking investigation: A systematic literature review. *arXiv preprint arXiv:2402.15448*

Belward, A. S., & Skøien, J. O. (2015). Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, 115–128

Bolaklis, C., Mantzana, V., Michalis, P., Vassiliou, A., Pflugfelder, R., Litzenberger, M., Hubner, M., Pastore, G., Oricchio, D., Desplas, M., Ansart, M., Santovito, M. R., Pica, G., Patino, L., Ferryman, J., spsampsps Kriechbaum-Zabini, A. (2021). Foldout: A through foliage surveillance system for border security. In *Technology Development for Security Practitioners*, (pp. 259–279). Springer

Borenstein, J., & Howard, A. (2021). Emerging challenges in ai and the need for ai ethics education. *AI and Ethics*, 1, 61–65

Bouschery, S. G., Blazevic, V., & Piller, F. T. (2024). Artificial intelligence-augmented brainstorming: How humans and ai beat humans alone. *SSRN Electronic Journal*. <https://ssrn.com/abstract=4724068>. Accessed 11 May 2024

Bronfman, N. C., Jiménez, R. B., Arévalo, P. C., & Cifuentes, L. A. (2012). Understanding social acceptance of electricity generation sources. *Energy Policy*, 46, 246–252. <https://www.sciencedirect.com/science/article/pii/S0301421512002625>

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodi, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901

Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1–21

Cartalis, C., Polydoros, A., Mavrouk, T.H., & Asimakopoulos, D. N. (2015). Earth observation in support of urban resilience and climate adaptability plans. *The Open Remote Sensing Journal*, 6(1):17–22. <https://benthamopen.com/ABSTRACT/TORMSJ-6-17>. Accessed 11 May 2024

Cazenave, A., & Nerem, R. S. (2004). Present-day sea level change: Observations and causes. *Reviews of Geophysics*, 42(3), 1–20. <https://doi.org/10.1029/2003rg000139>

Chakrabarty, T., Laban, P., Agarwal, D., Muresan, S., & Wu, C.-S. (2024). Art or artifice? Large language models and the false promise of creativity. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1–34). Honolulu: Association for Computing Machinery. <https://doi.org/10.1145/3613904.3642731>

Chrysoulakis, N., Feigenwinter, C., Triantakoustantis, D., Penyevskiy, I., Tal, A., Parlow, E., Fleishman, G., Düzgün, S., Esch, T., & Marconcini, M. (2014). A conceptual list of indicators for urban planning and management based on earth observation. *ISPRS International Journal of Geo-Information*, 3(3), 980–1002

Chu, D. A., Kaufman, Y. J., Zibordi, G., Chern, J., Mao, J., Li, C., & Holben, B. (2003). Global monitoring of air pollution over land from the earth observing system-terra moderate resolution imaging spectroradiometer (modis). *Journal of Geophysical Research: Atmospheres*, 108(D21), 1–18. <https://doi.org/10.1029/2002jd003179>

Chung, J. J. Y., Kim, W., Yoo, K. M., Lee, H., Adar, E., & Chang, M. (2022). Talebrush: Sketching stories with generative pretrained language models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery. <https://doi.org/10.1145/3491102.3501819>

Constantinides, M., Bogucka, E., Quercia, D., Kallio, S., & Tahaei, M. (2024). RAI Guidelines: Method for Generating Responsible AI Guidelines Grounded in Regulations and Usable by (Non-)Technical Roles. In *Proc. ACM Hum.-Comput. Interact.* vol. 8(388) (p. 28), New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/3686927>

Copa-Cogeca, CEMA, Fertilizers Europe., CEETAR, CEJA, ECPA, EFFAB, FEFAC, & ESA. (2018). EU Code of conduct on agricultural data sharing by contractual agreement. [https://cema-agri.org/images/publications/brochures/EU\\_Code\\_of\\_conduct\\_on\\_agricultural\\_data\\_sharing\\_by\\_contractual\\_agreement\\_2020\\_ENGLISH.pdf](https://cema-agri.org/images/publications/brochures/EU_Code_of_conduct_on_agricultural_data_sharing_by_contractual_agreement_2020_ENGLISH.pdf). Accessed 11 May 2024

Couch, D. L., Robinson, P., & Komisaroff, P. A. (2020). Covid-19-extending surveillance and the panopticon. *Journal of Bioethical Inquiry*, 17(4), 809–814

Crowley, M. A., Stockdale, C. A., Johnston, J. M., Wulder, M. A., Liu, T., McCarty, J. L., Rieb, J. T., Cardille, J. A., & White, J. C. (2023). Towards a whole-system framework for wildfire monitoring using earth observations. *Global Change Biology*, 29(6), 1423–1436

Cui, J., Li, Z., Yan, Y., Chen, B., & Yuan, L. (2023). Chatlaw: Open-source legal large language model with integrated external knowledge bases. *arXiv preprint arXiv:2306.16092*

Cutter, S. L. (2021). Urban risks and resilience. In W. Shi, M. F. Goodchild, M. Batty, M. P. Kwan, & A. Zhang (Eds.), *Urban Informatics*, Singapore: Springer 197–211. [https://doi.org/10.1007/978-981-15-8983-6\\_13](https://doi.org/10.1007/978-981-15-8983-6_13)

De Miguel Velazquez, J., Šćepanović, S., Gvirtz, A., & Quercia, D. (2024). Decoding Real-World Artificial Intelligence Incidents. *Computer*. 57(11):71–81. <https://doi.org/10.1109/MC.2024.3432492>

Deshpande, A., & Sharp, H. (2022). Responsible AI systems: who are the stakeholders? In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 227–236)

Dziri, N., Milton, S., Yu, M., Zaiane, O. R., & Reddy, S. (2022). On the origin of hallucinations in conversational models: Is it the datasets or the

models? In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 5271–5285)

Eldan, R., & Li, Y. (2023). Tinystories: How small can language models be and still speak coherent english? *arXiv preprint arXiv:2305.07759*. <https://arxiv.org/abs/2305.07759>

Elliott, J., Walters, R., & Wright, T. (2016). The role of space-based observation in understanding and responding to active tectonics and earthquakes. *Nature Communications*, 7(1), 13844

European Parliament. (2023). Amendments adopted by the european parliament on 14 june 2023 on the proposal for a regulation of the european parliament and of the council on laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. [https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html). Accessed 11 May 2024

Floridi, L., Cowsils, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., spsampsps Vayena, E. (2021). *An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations* (pp. 19–39). Springer International Publishing

Friedlingstein, P., Jones, M. W., O’Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., Le Quéré, C., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Bates, N. R., Becker, M., Bellouin, N., ... Zeng, J. (2022). Global carbon budget 2021. *Earth System Science Data*, 14(4), 1917–2005. <https://doi.org/10.5194/essd-14-1917-2022>

Fu, J., Han, H., Su, X., & Fan, C. (2024). Towards human-ai collaborative urban science research enabled by pre-trained large language models. *Urban Informatics*, 3(1), 8

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1), 35–42

Ganaie, M. A., Hu, M., Malik, A. K., Tanveer, M., & Suganthan, P. N. (2022). Ensemble deep learning: A review. *Engineering Applications of Artificial Intelligence*, 115, 105151. <https://www.sciencedirect.com/science/article/pii/S095219762200269X>

Gao, Z., Ji, H., Mei, T., Ramesh, B., & Liu, X. (2019). Eovnet: Earth-observation image-based vehicle detection network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(9), 3552–3561. <https://doi.org/10.1109/jstars.2019.2933501>

Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30). <https://doi.org/10.1073/pnas.2305016120>

Giray, L. (2023). Prompt engineering with chatgpt: A guide for academic writers. *Annals of Biomedical Engineering*, 51(12):2629–2633. <https://doi.org/10.1007/s10439-023-03272-4>. Accessed 11 May 2025

Golden, S. D., McLeroy, K. R., Green, L. W., Earp, J. A. L., & Lieberman, L. D. (2015). Upending the social ecological model to guide health promotion efforts toward policy and environmental change. *Health Education & Behavior*, 42(1\_suppl), 85–145. PMID: 25829123

Golpayegani, D., Pandit, H. J., & Lewis, D. (2023). To be high-risk, or not to be—semantic specifications and implications of the ai act’s high-risk ai applications and harmonised standards. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’23)*, pp. 905–915. ACM.

Goodchild, M. F. (2022). Elements of an infrastructure for big urban data. *Urban Informatics*, 1(1), 3

Group on Earth Observation, Committee on Earth Observation Satellites (2017). Earth Observations in support of the 2030 Agenda for Sustainable Development. Technical report

Guo, J., Hong, D., & Zhu, X. X. (2024). High-resolution satellite images reveal the prevalent positive indirect impact of urbanization on urban tree canopy coverage in South America. *Landscape and Urban Planning*, 247, 105076

Guo, C., Zhu, D., Ding, Y., Liu, H., & Zhao, Y. (2022). A systematic framework for the complex system engineering of city data governance. *Urban Informatics*, 1(1), 14

Hao, M., Chen, S., Lin, H., Zhang, H., & Zheng, N. (2024). A prior knowledge guided deep learning method for building extraction from high-resolution remote sensing images. *Urban Informatics*, 3(1), 6

Hassel, A., & Özkiziltan, D. (2023). Governing the work-related risks of ai: implications for the german government and trade unions. *Transfer: European Review of Labour and Research*, 29(1), 71–86

High-Level Expert Group on Artificial Intelligence. (2019). Ethics guidelines for trustworthy ai. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Accessed 11 May 2024

Hutson, M. (2023). Rules to keep AI in check: nations carve different paths for tech regulation. *Nature*. 620, 620-263. <https://www.nature.com/articles/d41586-023-02491-y>. Accessed 11 May 2024

International Federation of Accountants and Business at OECD. (2018). Regulatory divergence: costs, risks, impacts. [https://www.ifac.org/\\_flysystem/azure-private/publications/files/IFAC-OECD-Regulatory-Divergence.pdf](https://www.ifac.org/_flysystem/azure-private/publications/files/IFAC-OECD-Regulatory-Divergence.pdf). Accessed 11 May 2024

Jackson, D. (2018). *Data cities: how satellites are transforming architecture and design*. Lund Humphries

Jiang, J., Kahai, S., & Yang, M. (2022). Who needs explanation and when? juggling explainable ai and user epistemic uncertainty. *International Journal of Human-Computer Studies*, 165, 102839

Jin, Z., Wang, Y., Wang, Q., Ming, Y., Ma, T., & Qu, H. (2022). GNNlens: A visual analytics approach for prediction error diagnosis of graph neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 29(6), 3024–3038. <https://doi.org/10.1109/tvcg.2022.3148107>

Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A., & Landivar-Bowles, J. (2021). The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. *Current Opinion in Biotechnology*, 70, 15–22. <https://www.sciencedirect.com/science/article/pii/S0958166920301257>

Kaplan, G., Rashid, T., Gasparovic, M., Pietrelli, A., & Ferrara, V. (2022). Monitoring war-generated environmental security using remote sensing: A review. *Land Degradation & Development*, 33(10), 1513–1526

Khayyat, M. et al. (2021). Responsible ai in urban science. *Science and Justice*, 11(2), 123–130. <https://journal.unnes.ac.id/nju/sji/article/view/43818>

Kochupillai, M. (2021). Outline of a novel approach for identifying ethical issues in early stages of ai4eo research. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS* (pp. 1165–1168). IEEE

Kochupillai, M., Kahl, M., Schmitt, M., Taubenböck, H., & Zhu, X. X. (2022). Earth observation and artificial intelligence: Understanding emerging ethical issues and opportunities. *IEEE Geoscience and Remote Sensing Magazine*, 10(4), 90–124. <https://doi.org/10.1109/mgrs.2022.3208357>

Kochupillai, M., & Taubenböck, H. (2023). Conducting ethically mindful earth observation research: The case of slum mapping. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1937–1940). IEEE

Kondylatos, S., Prapas, I., Ronco, M., Papoutsis, I., Camps-Valls, G., Piles, M., Fernández-Torres, M.-Á., & Carvalhais, N. (2022). Wildfire danger prediction and understanding with deep learning. *Geophysical Research Letters*, 49(17). <https://doi.org/10.1029/2022gl099368>

Kuenzer, C., Ottinger, M., Wegmann, M., Guo, H., Wang, C., Zhang, J., Dech, S., & Wikelski, M. (2014). Earth observation satellite sensors for biodiversity monitoring: Potentials and bottlenecks. *International Journal of Remote Sensing*, 35(18), 6599–6647. <https://doi.org/10.1080/01431161.2014.964349>

Kundu, D., & Pandey, A. K. (2020). World urbanisation: trends and patterns. In D. Kundu, R. Sietchingip, & M. Kinyanjui (Eds.), *Developing national urban policies: Ways forward to green and smart cities* (pp. 13–49). Singapore: Springer Nature Singapore. [https://doi.org/10.1007/978-981-15-3738-7\\_2](https://doi.org/10.1007/978-981-15-3738-7_2)

Lambakis, S. (2024). Space sensors and missile defense. *Comparative Strategy*, 43(1), 1–57

Li, C., Mao, J., Lau, A. K., Yuan, Z., Wang, M., & Liu, X. (2005). Application of modis satellite products to the air pollution research in Beijing. *Science in China Series D(Earth Sciences)*, 48, 209–219

Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., Yi, J., & Zhou, B. (2023). Trustworthy AI: From principles to practices. *ACM Computing Surveys*, 55(9), 1–46

Li, X., Yu, L., Xu, Y., Yang, J., & Gong, P. (2016). Ten years after hurricane Katrina: Monitoring recovery in New Orleans and the surrounding areas using remote sensing. *Science Bulletin*, 61, 1460–1470

Liang, W., Rajani, N., Yang, X., Ozoani, E., Wu, E., Chen, Y., Smith, D. S., & Zou, J. (2024). What’s documented in ai? systematic analysis of 32k ai model cards. *arXiv preprint arXiv:2402.05160*

Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2024). Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12, 157–173

- Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., & Tang, J. (2023). Self-supervised learning: Generative or contrastive. *IEEE Transactions on Knowledge and Data Engineering*, 35(1), 857–876
- Large Model Systems Organization (LMSYS) (2024). LMSYS Chatbot Arena Leaderboard. <https://lmsys.org/blog/2023-06-22-leaderboard/>. Accessed 11 May 2024
- Lu, Q., Zhu, L., Xu, X., & Whittle, J. (2023). Responsible-ai-by-design: A pattern collection for designing responsible artificial intelligence systems. *IEEE Software*, 40(3), 63–71
- Luccioni, S., Akiki, C., Mitchell, M., & Jernite, Y. (2023). Stable bias: Evaluating societal representations in diffusion models. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, & S. Levine (Eds.), *Advances in Neural Information Processing Systems*, 36:56338–56351. Curran Associates, Inc. [https://proceedings.neurips.cc/paper\\_files/paper/2023/file/b01153e7112b347d8ed54f317840d8af-Paper-Datasets\\_and\\_Benchmarks.pdf](https://proceedings.neurips.cc/paper_files/paper/2023/file/b01153e7112b347d8ed54f317840d8af-Paper-Datasets_and_Benchmarks.pdf). Accessed 11 May 2024
- Lukowicz, P., Mayer, S., Koch, J., Shawe-Taylor, J., & Tiddi, I. (2023). Interacting with large language models: A case study on ai-aided brainstorming for guesstimation problems. In *HHAI 2023: Augmenting Human Intellect: Proceedings of the Second International Conference on Hybrid Human-Artificial Intelligence* (vol. 368, p. 153). IOS Press
- McGregor, S. (2021). Preventing repeated real world ai failures by cataloging incidents: The ai incident database. In *Proceedings of the AAAI Conference on Artificial Intelligence* (vol. 35, pp. 15458–15463). Association for Computing Machinery. <https://doi.org/10.1145/3287560.3287596>
- Micheli, M., Gevaert, C. M., Carman, M., Craglia, M., Daemen, E., Ibrahim, R. E., Kotsev, A., Mohamed-Ghouse, Z., Schade, S., Schneider, I., Shanley L. A., Tartaro, A., & Vespe, M. (2022). Ai ethics and data governance in the geospatial domain of digital earth. *Big Data & Society*, 9(2). <https://doi.org/10.1177/20539517221138767>
- Microsoft Azure. (2025). Prompt Engineering Techniques. <https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/prompt-engineering>. Accessed 11 May 2024
- Mikaléf, P., Conboy, K., Lundström, J. E., & Popović, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. *Eur. J. Inf. Syst.* 31(3):257–268. <https://doi.org/10.1080/0960085x.2022.2026621>
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency* (pp. 220–229). Atlanta: PUBLISHER: Association for Computing Machinery. <https://doi.org/10.1145/3287560.3287596>
- Mittelstadt, B., Wachter, S., & Russell, C. (2023). To protect science, we must use llms as zero-shot translators. *Nature Human Behaviour*, 7(11), 1830–1832
- Moraes, T. G., Almeida, E. C., & de Pereira, J. R. L. (2021). Smile, you are being identified! risks and measures for the use of facial recognition in (semi-) public spaces. *AI and Ethics*, 1(2), 159–172
- Musakwa, W., & Van Niekerk, A. (2015). Earth observation for sustainable urban planning in developing countries: Needs, trends, and future directions. *Journal of Planning Literature*, 30(2), 149–160
- Nanayakkara, P., Hullman, J., & Diakopoulos, N. (2021). Unpacking the expressed consequences of ai research in broader impact statements. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 795–806). New York: Association for Computing Machinery. <https://doi.org/10.1145/3461702.3462608>
- National Institute of Standards and Technology (2023a). <https://www.nist.gov/itl/ai-risk-management-framework>. Accessed 11 May 2024
- National Institute of Standards and Technology (2023b). The EqualAI Algorithmic Impact Assessment Tool. <https://www.equalai.org/aia/>. Accessed 11 May 2024
- Nokia Bell Labs. (n.a.) Responsible AI. <https://www.bell-labs.com/research-innovation/ai-software-systems/responsible-ai/>. Accessed 17 Sep 2023
- Nourani, M., Roy, C., Block, J. E., Honeycutt, D. R., Rahman, T., Ragan, E., & Gogate, V. (2021). Anchoring bias affects mental model formation and user reliance in explainable ai systems. In *26th International Conference on Intelligent User Interfaces* (pp. 340–350). College Station: Association for Computing Machinery. <https://doi.org/10.1145/3397481.3450639>
- NVIDIA Corporation. (n.a.) Trustworthy AI. <https://www.nvidia.com/en-us/ai-data-science/trustworthy-ai/>. Accessed 17 Sep 2023
- OpenAI. (2023a). Gpt-4 technical report. 2303.08774
- OpenAI. (2023b). Research on gpt-4 - latest updates. <https://openai.com/research/gpt-4>. Accessed 11 May 2024
- Parks, L. (2002). Satellite and cybervisualities: Analyzing "digital earth". *The Visual Culture Reader*, 2, 279–294
- Pataranutaporn, P., Liu, R., Finn, E., & Maes, P. (2023). Influencing human-ai interaction by priming beliefs about ai can increase perceived trustworthiness, empathy and effectiveness. *Nature Machine Intelligence*, 5(10), 1076–1086
- Patias, P., Mallinis, G., Tsioukas, V., Georgiadis, C., Kaimaris, D., Tassopoulou, M., Verde, N., Dohr, M., & Riffler, M. (2020). Earth observations as a tool for detecting and monitoring potential environmental violations and policy implementation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 1491–1496
- Persello, C., Wegner, J. D., Hänsch, R., Tuia, D., Ghamisi, P., Koeva, M., & Camps-Valls, G. (2022). Deep learning and earth observation to support the sustainable development goals: Current approaches, open challenges, and future opportunities. *IEEE Geoscience and Remote Sensing Magazine*, 10(2), 172–200
- Praks, J., Hallikainen, M., Antropov, O., & Molina, D. (2012). Boreal forest tree height estimation from interferometric tandem-x images. In *2012 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1262–1265). IEEE. <https://doi.org/10.1109/igars.2012.6351309>
- Prunkl, C. E., Ashurst, C., Anderljung, M., Webb, H., Leike, J., & Dafoe, A. (2021). Institutionalizing ethics in ai through broader impact requirements. *Nature Machine Intelligence*, 3(2), 104–110
- Purdy, R. (1999). Legal and privacy implications of 'spy in the sky' satellites. *Mountbatten Journal of Legal Studies*, 3(1), 63–79
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). Deep learning and process understanding for data-driven earth system science. *Nature*, 566(7743), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
- Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. In K. Inui, J. Jiang, V. Ng, & X Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3982–3992). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1410>
- Sacha, D., Sedlmair, M., Zhang, L., Lee, J. A., Peltonen, J., Weiskopf, D., North, S. C., & Keim, D. A. (2017). *What you see is what you can change: Human-centered machine learning by interactive visualization* (vol. 268, pp. 164–175). Elsevier BV. <https://doi.org/10.1016/j.neucom.2017.01.105>
- Salimzadeh, S., He, G., & Gadiraju, U. (2024). Dealing with uncertainty: Understanding the impact of prognostic versus diagnostic tasks on trust and reliance in human-ai decision-making. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. Honolulu, HI: Association for Computing Machinery. <https://doi.org/10.1145/3613904.3641905>
- Šćepanović, S., Antropov, O., Laurila, P., Rauste, Y., Ignatenko, V., & Praks, J. (2021). Wide-area land cover mapping with sentinel-1 imagery using deep learning semantic segmentation models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 10357–10374. <https://doi.org/10.1109/jstars.2021.3116094>
- Šćepanović, S., Bogucka, E. P., Quercia, D., & Nattero, C. (2023). Responsible ai for earth observation: Attitudes among experts. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1934–1936). IEEE. <https://doi.org/10.1109/igars52108.2023.10282983>
- Sćepanovic, S., Obadic, I., Joglekar, S., Giustarini, L., Nattero, C., Quercia, D., & Zhu, X. (2023). Medsat: A public health dataset for england featuring medical prescriptions and satellite imagery. *Advances in Neural Information Processing Systems*, 36:77810–77851. [https://proceedings.neurips.cc/paper\\_files/paper/2023/file/f4fd676c3b21f20f8c391d929188386-Paper-Datasets\\_and\\_Benchmarks.pdf](https://proceedings.neurips.cc/paper_files/paper/2023/file/f4fd676c3b21f20f8c391d929188386-Paper-Datasets_and_Benchmarks.pdf). Accessed 11 May 2024
- Schumann, G. J., Brakenridge, G. R., Kettner, A. J., Kashif, R., & Niebuhr, E. (2018). Assisting flood disaster response with earth observation data and products: A critical assessment. *Remote Sensing*, 10(8), 1230. <https://doi.org/10.3390/rs10081230>
- Selva, D., & Krejci, D. (2012). A survey and assessment of the capabilities of cubesats for earth observation. *Acta Astronautica*, 74, 50–68
- Sherman, E., & Eisenberg, I. W. (2024). AI Risk Profiles: A Standards Proposal for Pre-Deployment AI Risk Disclosures. In *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI*. [AAAI] Press. 23047–23052. <https://doi.org/10.1609/AAAI.V38I21.30348>

- Sherman, E., & Eisenberg, I. (2024). Ai risk profiles: A standards proposal for pre-deployment ai risk disclosures. In *Proceedings of the AAAI Conference on Artificial Intelligence* (vol. 38, pp. 23047–23052)
- Shi, W., Batty, M., Goodchild, M., & Li, Q. (2022). The digital transformation of cities. *Urban Informatics*, 1(1), 1
- Shieh, J. (2023). Best practices for prompt engineering with openai api. <https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-openai-api>. Accessed 11 May 2024
- Smuha, N. A. (2021). From a 'race to ai' to a 'race to ai regulation': Regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57–84
- Stahl, B. C., Antoniou, J., Bhalla, N., Brooks, L., Jansen, P., Lindqvist, B., Kirichenko, A., Marchal, S., Rodrigues, R., Santiago, N., Warso, Z., & Wright, D. (2023). A systematic review of artificial intelligence impact assessments. *Artificial Intelligence Review*, 56(11), 12799–12831
- Sun, Z., Sandoval, L., Crystal-Omelas, R., Mousavi, S. M., Wang, J., Lin, C., Cristea, N., Tong, D., Carande, W. H., Ma, X., Rao, Y., Bednar, J. A., Tan, A., Wang, J., Purushotham, S., Gill, T. E., Chastang, J., Howard, D., Holt, B., ... John, A. (2022). A review of earth artificial intelligence. *Computers and Geosciences*, 159, 105034. <https://www.sciencedirect.com/science/article/pii/S0098300422000036>
- Tahaee, M., Constantinides, M., Quercia, D., Kennedy, S., Muller, M., Stumpf, S., Liao, Q. V., Baeza-Yates, R., Aroyo, L., Holbrook, J., Luger, E., Madaio, M., Blumenfeld, I. G., De-Arteaga, M., Vitak, J., & Olteanu, A. (2023a). Human-centered responsible artificial intelligence: Current & future trends. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–4)
- Tahaee, M., Constantinides, M., Quercia, D., & Muller, M. (2023b). A systematic literature review of human-centered, ethical, and responsible ai. *arXiv preprint arXiv:2302.05284*
- Tolmeijer, S., Christen, M., Kandul, S., Kneer, M., & Bernstein, A. (2022). Capable but amoral? comparing ai and human expert collaboration in ethical decision making. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22). Association for Computing Machinery. <https://doi.org/10.1145/3491102.3517732>
- Tuia, D., Schindler, K., Demir, B., Camps-Valls, G., Zhu, X. X., Kochupillai, M., Džeroski, S., van Rijn, J. N., Hoos, H. H., Del Frate, F., Dătu, M., Volker Markl, Bertrand Le Saux, Rochelle Schneider, Gustau Camps-Valls (2023). Artificial intelligence to advance earth observation: a perspective. *arXiv preprint arXiv:2305.08413*
- United Nations: Department of Economic and Development, S. A. S. (2023). The 17 goals. <https://sdgs.un.org/goals>
- Vettorel, A. (2023). Earth observation, satellite navigation and privacy: The international, european and italian legal framework. In *Rights of Individuals in an Earth Observation and Satellite Navigation Environment* (pp. 81–113). Brill Nijhoff
- Šćepanović, S., Joglekar, S., Law, S., & Quercia, D. (2021). Jane jacobson in the sky: Predicting urban vitality with open satellite data. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1–25
- Wang, G.-G., Cheng, H., Zhang, Y., & Yu, H. (2022). Enso analysis and prediction using deep learning: A review. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2022.11.078>
- Wang, Z. J., Kulkarni, C., Wilcox, L., Terry, M., & Madaio, M. (2024). Far-sight: Fostering responsible ai awareness during ai application prototyping. *arXiv preprint arXiv:2402.15350*
- Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., Du, N., Dai, A. M., & Le, Q. V. (2022a). Finetuned language models are zero-shot learners. *The Tenth International Conference on Learning Representations, (ICLR) 2022, Virtual Event*, April 25–29, 2022. OpenReview.net. <https://openreview.net/forum?id=gEZrGCozdqR>
- Wei, A., Haghtalab, N., & Steinhardt, J. (2024). Jailbroken: How does LLM safety training fail? *Advances in Neural Information Processing Systems*, 36
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022b). Chain-of-thought prompting elicits reasoning in large language models. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., & Oh, A., editors, *Advances in Neural Information Processing Systems* (vol. 35, pp. 24824–24837). Curran Associates, Inc. <https://proceedings.neurips.cc/paperfiles/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf>. Accessed 11 May 2024.
- Wei, C., Wang, Y.-C., Wang, B., & Kuo, C.-C. J. (2023). An overview on language models: Recent developments and outlook. *CoRR: ArXiv, abs/2303.05759*. <https://doi.org/10.48550/ARXIV.2303.05759>. Accessed 11 May 2024
- World Health Organization. Regional Office for Europe. (2017). *Urban green spaces: a brief for action*. World Health Organization. Regional Office for Europe
- Xu, F., Heremans, S., & Somers, B. (2022). Urban land cover mapping with sentinel-2: A spectro-spatio-temporal analysis. *Urban Informatics*, 1(1), 8
- Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., & Dickinson, R. (2013). The role of satellite remote sensing in climate change studies. *Nature Climate Change*, 3(10), 875–883. <https://doi.org/10.1038/nclimate1908>
- Yang, X., Liang, W., & Zou, J. (2024). Navigating dataset documentations in AI: A large-scale analysis of dataset cards on huggingface. In *The Twelfth International Conference on Learning Representations (ICLR)*. OpenReview.net. <https://openreview.net/forum?id=xC8xh2RSs2>
- Young, O. R., spamspsps Onoda, M. (2017). *Satellite Earth Observations in Environmental Problem-Solving* (pp. 3–27). Springer Singapore. [https://doi.org/10.1007/978-981-10-3713-9\\_1](https://doi.org/10.1007/978-981-10-3713-9_1)
- Zhang, Y., & Liu, M. (2022). Responsible ai in urban informatics. *arXiv preprint arXiv:2208.04727*. <https://arxiv.org/pdf/2208.04727>. Accessed 11 May 2024
- Zhang, L., & Zhang, L. (2022). Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities. *IEEE Geoscience and Remote Sensing Magazine*, 10(2), 270–294
- Zhao, Q., Yu, L., Du, Z., Peng, D., Hao, P., Zhang, Y., & Gong, P. (2022). An overview of the applications of earth observation satellite data: Impacts and future trends. *Remote Sensing*, 14(8), 1863
- Zheng, Z., Chen, K.-Y., Cao, X.-Y., Lu, X.-Z., & Lin, J.-R. (2023). Llm-funcmapper: Function identification for interpreting complex clauses in building codes via llm. *CoRR arXiv preprint arXiv, abs/2308.08728*. <https://doi.org/10.48550/arXiv.2308.08728>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.