PROGETTI VI CALL 2025/2026

Titolo progetto:	Degrado urbano: analisi delle immagini di Google Street View mediante modelli di intelligenza artificiale
Descrizione	I progetto ha esplorato l'impiego di nuove fonti di dati non campionari, basate su immagini di street imagery di provenienza corporate, e di modelli generalizzati di intelligenza artificiale per lo studio del degrado urbano. L'attività ha previsto la simulazione di un processo di osservazione sociale sistematica (Systematic Social Observation – SSO), finalizzato alla rilevazione oggettiva delle caratteristiche fisiche dello spazio urbano attraverso immagini georeferenziate.
Obiettivi	Il progetto si è posto due obiettivi principali: - Produrre, attraverso nuove fonti di dati non campionari, informazioni sul degrado urbano, attualmente rilevate anche nell'ambito delle indagini survey dell'ex Sistema Multiscopo. - Utilizzare tali informazioni per valutare la qualità dei dati d'indagine, verificandone l'attendibilità e l'eventuale presenza di bias cognitivi nelle risposte dei rispondenti.
Metodologia	1. Fase amministrativa preliminare: - Attivazione dello spazio nell'ambito della convenzione tra l'ente corporate e la Pubblica Amministrazione. - Accesso ai dati elementari di street imagery. - Impostazione di un metodo di copertura dello spazio urbano. 2. Scelta dello studio di caso, valutazione dell'affidabilità dei modelli AI, costruzione degli indicatori e analisi della coerenza: Il flusso di dati è stato costruito per scaricare le immagini attraverso le coordinate geografiche delle stesse, corrispondenti a quelle dei quartieri di residenza degli intervistati dell'indagine campionaria, creazione su server di uno storage delle immagini che conservassero identificativi e metadati geografici forniti dall'app Google Tiles. I metadati geografici contengono l'orientamento e l'inclinazione dell'obiettivo della fotocamera che ha catturato l'immagine, la data della rilevazione e le coordinate geografiche del punto di rilevazione. Questo materiale viene poi assorbito dentro tre modelli linguistici generalizzati (due modelli di Gemini ed uno di Open AI). I tre modelli sono interrogati con un prompt che chiede loro la presenza assenza di 15 elementi che abbiamo ritenuto costitutivi del concetto di degrado urbano: • Veicoli abbandonati • Edifici abbandonati • Finestre rotte • Facciate degradate • Strade dissestate • Buche • Vegetazione spontanea • Graffiti • Presenza di rifiuti • Arredo urbano rovinato • Vetri rotti • Escrementi sulle strade • Popolazione senza fissa dimora • Cantieri stradali Oltre alla presenza/assenza di questi elementi è stato richiesto ai modelli un indice numerico da 0 a 100 che corrispondesse ad una valutazione generale del luogo osservato. La valutazione dell'efficienza del riconoscimento visuale dei modelli linguistici generalizzati, è stata valutata su un numero di alcune centinaia di immagini attraverso uno scrutinio parallelo dell'occhio umano. Questa operazione ha consentito di affinare il prompt verbale di richiesta al modello linguistico, scegliendo FLASH 2.5 come quello di produzione ufficiale. Nella fase di produzione la variabile dicotomica di presenza/assenza è stata trasformata in una variabile tricotomica presenza/assenza/presenza cospicua. Questi dati sono stati poi utilizzati per un'analisi delle Corrispondenza al fine di trasformare gli eigenvalues di ognuno degli ITEM in pesi su cui costruire un indice oggettivo, diverso dalla valutazione "soggettiva" dell'intelligenza artificiale. E' stato poi costruito un modello logistico multilivello per determinare la coerenza tra le risposte degli intervistati delle survey e la rilevazione oggettiva dell'IA. Sulla

	mancata coerenza pesano alcune caratteristiche sociodemografiche soggettive, come età, sesso e posizione professionale.
Risultati ottenuti	Riferendoci alla letteratura esistente abbiamo deciso di approssimare, in merito alla produzione di informazione sul degrado tratta dall'analisi di fonti visuali, lo spazio continuo. Volendo dare un'informazione al dettaglio di unità ecologiche subcomunali, per evitare che il bias della scelta casuale di un singolo punto urbano non rappresentativo determinasse il risultato. Si è quindi proceduto a calcolare su un grafo stradale la distanza mediana tra due segmenti, il cui risultato è stato approssimativamente quello di 80 metri. Sono stati eliminati i tratti di strada inferiori a 20 metri e i tratti corrispondenti agli assi a grande scorrimento come il GRA e le Autostrade (usando il campo max speed), lungo i quali non ci sono abitazioni ed è perciò inutile generare punti. E' stato impostato uno sfalsamento iniziale 30 (in modo che la misura della distanza ogni 80 metri non riparta da zero ad ogni incrocio, ma dopo 30 metri) e uno sfalsamento finale pari a 0. Naturalmente ogni volta che il tratto di strada cambia nome, ricomincia il conteggio degli 80/30/0 metri, perciò come output è stato generato un layer temporaneo di circa 86.000 punti, alcuni dei quali molto vicini. Volendo eliminare i punti entro un raggio di 20 metri (non a distanza lineare lungo il grafo), ho fatto un buffer a 20 metri. Nei poligoni in cui si intersecano i poligoni di buffer possono essere presenti più punti perciò volendo lasciarne uno solo all'interno di ciascun poligono (mantenendo la sua posizione geometrica originale sull'asse stradale, cioè senza spostarlo al centro del poligono), l'approccio seguito è stato quello di eseguire un'intersezione spaziale e poi rimuovere i punti multipli in base all'identificativo dello stesso poligono in cui ricadono (usando il join tra punti e poligono one to one). In tal modo è possibile utilizzare in ArcGIS il tool "Delete identical" con il campo comune ORIG_FID (in questo modo il tool capisce quali punti condividono lo stesso poligono e ne lascia uno solo). Alla fine del processo ho ottenuto un layer con 79.601 punti dagli 86.000 (circa) iniziali. Questi punti sono stati infine suddivisi in base ai poligoni dei 15 municipi su cui sono stati prodotti gli indici di degrado. Prospettive Future: - Individuazione di strategie di campionamento spaziale in grado di ottenere risultati comparabili utilizzando un numero limitato di punti. - Costruzione di un flusso di raccolta basato su dati non proprietari, per consentire la produzione autonoma dell'informazione senza ricorrere a enti corporate.
Membri del Team	Marina Bertollini, Isabella Corazziari, Claudia De Vitiis, Marco Dionisio Terribili, Donato Summa (referenti informatico) e Alberto Violante (referente)