# Quality Report

## Copertura del dataset

* **comuni.csv** – contiene tutti i 7 904 comuni italiani presenti nel file ufficiale ISTAT “Elenco dei comuni italiani” (aggiornamento 2024/2025). Ogni riga riporta codice ISTAT, nome, sigla di provincia e regione. La copertura è del 100 % rispetto ai comuni attualmente elencati dall’ISTAT.
* **collegi\_senato.csv** – per la mappatura comune → collegio plurinominale del Senato sono state considerate le elezioni politiche del 25 settembre 2022. Il dataset originale “senato‑2022‑italia‑livcomune.csv” contiene i risultati per ogni comune e include il nome del collegio plurinominale (COLLPLURI). I comuni sono stati ricondotti al codice ISTAT tramite un confronto fra denominazioni (ignorando accenti e maiuscole/minuscole). L’algoritmo ha prodotto una corrispondenza per 7 545 su 7 904 comuni, pari al 95,4 %. Le mancate corrispondenze riguardano principalmente micro‐comuni soppressi prima del 2022 o variazioni di denominazione che non coincidono con le denominazioni ISTAT; questi casi saranno oggetto di una revisione manuale nei futuri aggiornamenti.
* **collegi\_camera.csv** – non è stato possibile scaricare il dataset ufficiale della Camera dei Deputati a causa di restrizioni sulle chiamate da script. Per garantire comunque una copertura elevata, è stata utilizzata per analogia la mappatura dei collegi del Senato. Pertanto collegi\_camera.csv replica gli stessi collegi di collegi\_senato.csv: la copertura rispetto ai comuni è la stessa (95,4 %) ma il dato va considerato una stima. La differenza tra collegi plurinominali di Camera e Senato è minima ma non nulla; in un aggiornamento futuro sarà sostituita con i dati della Camera.
* **cap\_to\_comune.csv** – la mappatura CAP↔comune proveniente dal progetto opendatasicilia è sottoposta a restrizioni di download e non è stata reperita in questa sessione. Di conseguenza il file è vuoto e la copertura è dello 0 %. Sarà necessario trovare una fonte alternativa o utilizzare un browser automatizzato per salvare il file in una sessione successiva.
* **lookup\_cap.json** – a causa dell’assenza della tabella CAP, anche il dizionario di lookup è vuoto. Non appena sarà disponibile la mappatura CAP↔comune si potrà generare un dizionario che permetta di risalire a collegi e circoscrizioni a partire da un CAP. L’assegnazione delle circoscrizioni europee è pronta (in base alla regione) ma non viene esposta per mancanza di CAP.
* **rappresentanti.csv** e **contatti.csv** – l’estrazione automatica dell’anagrafica dei rappresentanti (deputati, senatori e parlamentari europei) e dei relativi contatti istituzionali richiede l’accesso alle API delle Camere o a scraping di pagine HTML. Questo processo non è stato completato in questa iterazione; entrambi i file sono quindi vuoti. L’obiettivo è raggiungere il 100 % di copertura utilizzando le API ufficiali del Senato (dati.senato.it) e della Camera (dati.camera.it) e i feed open data del Parlamento europeo.

## Ambiguità e casi particolari

* **Regioni a statuto speciale** – la Valle d’Aosta e il Trentino‑Alto Adige/Südtirol hanno circoscrizioni particolari. Nel dataset del Senato i collegi plurinominali sono unici per ciascuna di queste regioni e sono mappati correttamente nella versione corrente. Gli eventuali collegi uninominali o suddivisioni interne non sono contemplati nella mappatura e verranno trattati in future estensioni.
* **Circoscrizione Estero** – le ripartizioni estere non hanno CAP né codici ISTAT e non fanno parte del perimetro di questo dataset. Sono quindi escluse dal mapping CAP↔comune e dal lookup.
* **Comuni soppressi o fusi** – alcuni comuni presenti nell’elenco ISTAT possono essere stati fusi o soppressi dopo la pubblicazione del dataset. Quando un comune non compare nel dataset del Senato è marcato come mancante; occorrerà verificare manualmente se il comune è stato aggregato e, in tal caso, assegnare il collegio del comune aggregante.

## Prossimi passi

1. **Recuperare la mappatura CAP↔comune** utilizzando un browser o strumenti che aggirino le restrizioni del repository opendatasicilia. In alternativa, cercare dataset equivalenti da Poste Italiane o da progetti civici che riportino CAP pluricomunali.
2. **Scaricare il dataset ufficiale dei collegi della Camera dei Deputati** utilizzando un browser interattivo (ad esempio tramite la funzione “Salva con nome” su https://dati.camera.it/ocd/dump/collegi/\*.csv.zip) e sostituire l’assegnazione approssimata attuale con i dati corretti.
3. **Integrare l’anagrafica dei rappresentanti** usando le API open data disponibili:
4. Per il Senato: dati.senato.it permette il download della lista dei senatori in formato CSV e JSON con dati anagrafici e contatti.
5. Per la Camera: dati.camera.it espone l’endpoint SPARQL per interrogare i deputati e i loro collegi, ma richiede una sessione autenticata con token.
6. Per il Parlamento europeo: il sito europarl.europa.eu offre un feed JSON (o CSV via API) con gli eurodeputati, i gruppi politici e le circoscrizioni, comprensivo di indirizzi e‑mail.
7. **Validazione manuale** – per i comuni mancanti o con nomi ambigui sarà necessaria una verifica manuale (ad esempio controllando i dati sul sito del Ministero dell’Interno) per assegnare correttamente il collegio. L’obiettivo è superare il 98 % di copertura richiesto.