From open data to smart cities: How open source and open data enable urban transformation.

Table of Contents

1. Abstract 3
2. Introduction 3
3. Methodology 4
3.1. Data sources and openness 4
3.2. Workflow overview 4
3.3. Implementation 5
4. Applications for Smart-City Transformation 6
4.1. Operational characteristics 6
4.2. Decision-support outputs enabled by the pipeline 6
5. Results 7
5.1. Outcomes of the open-data connection 7
5.2. Process improvements (qualitative) 8
6. References 9

1. Abstract

Open geodata and an opensource stack (PostGIS/QGIS/QField) can be arranged in a reproducible pipeline that turns authoritative cadastral bases, such as ALKIS within the AAA model, into planning-ready information for urban decision-making. This here-presented workflow covers the intake and capture of licensing, lossless import to PostGIS, harmonization of the parcel and building layers, and controlled vocabularies for planning attributes, publishing read-only views for desktop, web, and field clients, and round-trip mobile capture with scripted merges and Quality Assurance. As such, this is in line with smart-city reference architectures centered around openness of interfaces and datasets, placing open data as the key ingredient of city ICT platforms [1].

Evidence from research into smart-city strategy also suggests that such data-centric, interoperable arrangements assist practical governance improvements and help connect digital initiatives with broader urban development objectives [2].

The result is a practical path from "open data in" to "planning decisions out," enabling applications such as renewal targeting, climate-adaptation siting, accessibility audits, and energy/emissions insights without dependence on proprietary formats or one-off manual processes.

2. Introduction

Open-source tools and open geodata increasingly form the basis for practical workflows in smart cities. Drawing on the integration of authoritative cadastral layers-e.g., ALKIS-with an open stack, QGIS, PostGIS, and field data collection tools like QField-allow municipalities to create end-to-end evidence-based processes that support mobile surveying, versioned database updates, and planning tasks such as renovation prioritization or CO_2 assessments. Recent reviews show how geospatial data-driven approaches translate this kind of pipeline into measurable urban sustainability gains, ranging from multisource spatial analytics to decision support for mobility, energy, and land use [3].

Policy and governance studies also find that open-data readiness in local administrations is a decisive enabler for such workflows; improving transparency, inter-departmental coordination, and the reuse of APIs/standardized datasets in day-to-day city management [4,5]. Looking ahead, these same open foundations; interoperable schemas, versioned databases, repeatable ETL are what allow teams to evolve toward urban "digital twins" where continuously updated views and indicators (e.g., building-level attributes collected in the field) are kept in sync with operational platforms and used to simulate and optimize interventions at district scale [6]. Taken together, the literature supports this pipeline; open geodata in, open tooling throughout, and planning decisions out, as a pragmatic path from open data to smarter, more sustainable cities.

3. Methodology

3.1. Data sources and openness

Open data are those published under permissive licenses, in machine readable formats that permit transparent reuse between tools and departments. This reduces costs, removes vendor lock-in, and permits reproducible workflows.

The main open source used here is ALKIS; other open layers, for example base maps or administrative boundaries, can be added where available to provide context.

ALKIS is the Amtliches Liegenschaftskatasterinformationssystem, Germany's authoritative cadaster provided by the state surveying/mapping authorities and coordinated by the AdV according to the AAA model AFIS–ALKIS–ATKIS.

Examples of ALKIS-provided content: geometries of parcels and buildings with permanent identifiers, addresses/house numbers and selected topographic features, typically referenced to ETRS89/UTM, licensed and with metadata describing the date of data extraction.

3.2. Workflow overview

The end-to-end process is executed as a repeatable pipeline:

I. Project registration with core identifiers
II. ALKIS download for specified area and import into PostGIS via norgis
III. Harmonization and derivations by Python/SQL
IV. Publication of read-only views to clients
V. Packaging to QField and field data acquisition on tablet
VI. Synchronization of edits and media back to database
VII. Automated merge & Quality Assurance

3.3. Implementation

3.3.1. Project initialization

A registry entry is created with project_id, gemeinde_id, gemarkung_id, source URLs, license text, CRS, and dataset vintage. A namespaced PostGIS schema and standardized folders for raw/staging/processed data are prepared.

3.3.2. Ingestion of ALKIS

ALKIS delivery files are checked concerning hash and file count and will be imported using norgis into tables in the Postgres database. It checks geometry validity like and for projection conformity, creates spatial indexes, and primary keys. Import logs are run ID, duration, feature counts, and geometric repair activities.

3.3.3. Harmonization and derivations

The core entities are normalized into the domain tables of flurstueck, gebaeude, adresse, plus generalized flaeche, linie, punkt. Compute derived attributes: parcel area, building footprint, centroids and address joins. Where the identifiers in ALKIS are stable, avoid surrogate keys; otherwise, add surrogate keys while retaining originals for traceability.

3.3.4. Controlled value dictionary

single dictionary table holds codes for Nutzung, Bestand, Missstände, Maßnahmen, Zielkoncept. Feature tables reference these codes via foreign keys, making sure there is consistent labeling in all clients -QGIS, WebGIS, FMS - and allowing non-breaking updates to labels.

3.3.5. Semantic layer (read-only views)

Consumption-ready views join geometry, attributes, and human-readable labels. Views are read-only and include lightweight denormalizations needed by map clients (e.g., floor counts, construction year, label strings). All downstream systems consume the same views, eliminating divergence between desktop GIS, WebGIS, and FMS.

3.3.6. Field package preparation and Field data acquisition protocol

A QGIS project is set up to refer only to the published views. The edit forms make use of dropdowns sourced from the dictionary table-to avoid free text-and constraints. There are media directories for photos; the photo path fields are pre-bound. Optional edit tables can be used to isolate field inputs from base tables where strict change control is required. The QField package (for offline) is exported with basemap, form logic, and project metadata.

Then, at the field, surveyors record, at minimum: deficiencies (Missstände), proposed measures (Maßnahmen), number of floors, construction year, and photos. Conditional logic is applied, and geometry edits are only allowed for specific layers (e.g., points for issues; polygons restricted).

3.3.7. Synchronization and merge

Returned QField data are ingested into tables with one row per edited feature and a link to media (UUID filenames). A merge script applies a conflict policy: attribute updates follow last-write-wins for non-critical fields; geometry changes and code changes are queued for manual review. Audit columns (editor information, modified time, and etc.) are populated via triggers. Successful merges refresh views so that all clients see the edits immediately.

4. Applications for Smart-City Transformation

4.1. Operational characteristics

Through a shared, read-only semantic layer, the same harmonized data are viewed in desktop GIS, WebGIS, and field tools. This reduces manual transfers and version drift. Light automation, such as scripted merges and Quality Assurance checks, is applied to keep datasets consistent without emphasizing system autonomy.

4.2. Decision-support outputs enabled by the pipeline

The integrated use of open cadastral data, harmonized attributes, and field observations enables a wide range of smart-city tasks, for example:

Urban renewal and building renovation: It is possible to prioritize either parcels or buildings using recorded deficiencies, construction periods, floor counts, and photo evidence. Candidate areas for either facade or roof improvements are surfaced for investment planning.

Climate adaptation and green infrastructure: By intersecting land-use classes with recorded issues (e.g., heat-exposed blocks, lack of shade) and available parcel geometries, opportunities may be identified for pocket parks and green spaces.

Energy and Emissions Insights: Building-age bands and basic attributes provide a measure of thermal performance, while external metrics allow districts to be indexed for potential CO_2 reductions under renovation scenarios.

Mobility and Accessibility: Field-captured barriers, crossings, and path conditions enhance base data to develop safer routes, improve school zones, or act in zones that are prioritized for pedestrians.

Quality and safety of the public sphere: lighting defects, damaged surfaces, and micro-hazards are georeferenced with photos, allowing for targeted maintenance and faster response cycles.

Social-infrastructure planning: Proximity analyses to schools, health care, parks, combined with parcel and building attributes, help in identifying service gaps and provide just distribution of resources.

5. Results

5.1. Outcomes of the open-data connection

Combining open cadastral data (ALKIS) with an open-source stack of PostGIS / QGIS / QField had the following results:

- Single source of truth: Harmonized, read-only views are published for desktop GIS, WebGIS, and field tools, to reduce version drift.
- Repeatable data flow: Ingestion, synchronization, and publishing are scripted to enable repeatable executions across data and projects, reducing the need for manual work.
- Field-ready data model: Controlled data dictionaries (e.g. Nutzung (Utilization/Land Use), Missstände (Deficiencies/Problems), Maßnahmen (Measures/Interventions), Zielkonzept (Goal Concept/Desirable Plan) are applied via foreign keys, which eliminate free text and ensure consistent labeling across sections. An example thematic map of deficiencies is shown in Fig. 1, illustrating how coded categories in the pipeline appear in planning views.

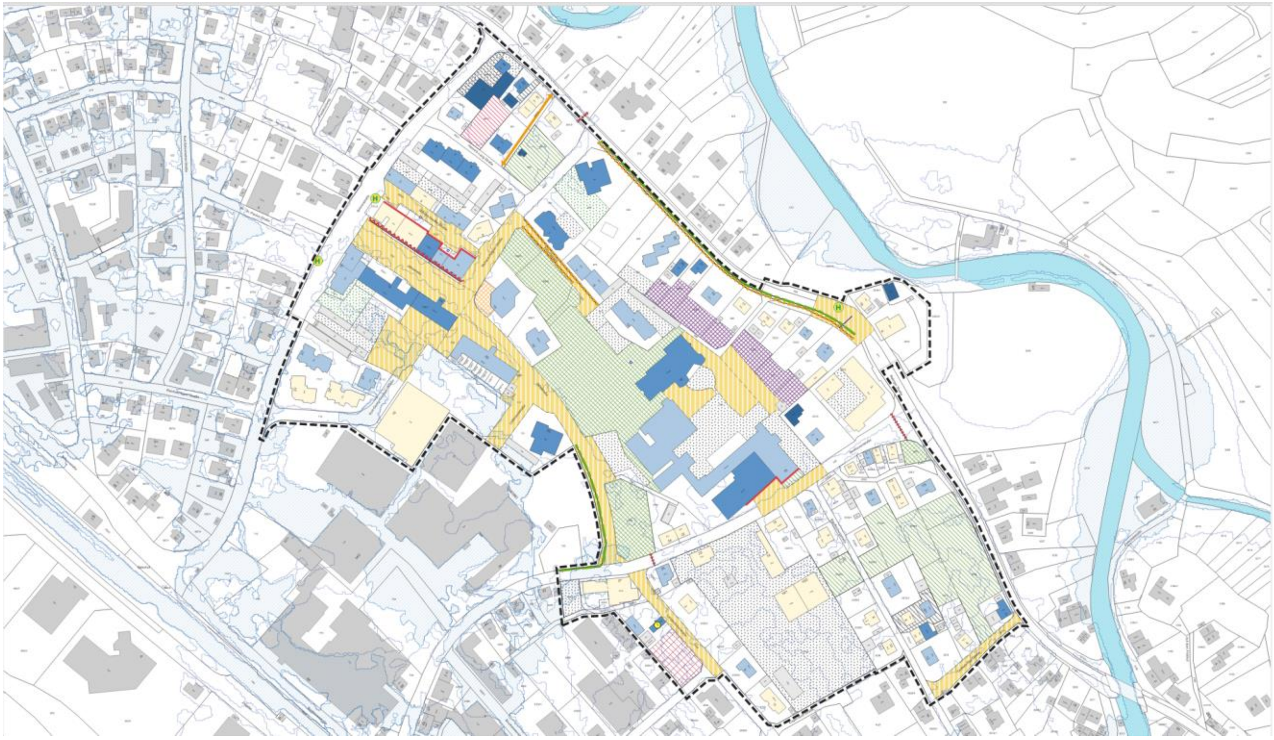

Figure 1. Illustrative Missstände (deficiencies) map produced from the harmonized views. Categories are encoded via the controlled dictionary and symbolized in QGIS

- Round-trip editing: Offline captured packages are prepared and merged back under a defined conflict policy with edit history tracking.
- Transparency Compliance: Licensing, attribution, CRS, and extraction date are recorded to support transparent reuse.

5.2. Process improvements (qualitative)

The following improvements are observed in practice, relative to manual desktop workflows:

- Fewer manual transfers: All clients consume the views directly; hence few export-copy steps are needed.
- Lower data errors: Dictionary-driven forms reduce typos and inconsistent codes.
- Faster iteration: Scripted ingestion and view refresh shorten the path from open data to planning-ready layers.
- Traceability: Edits and merges are auditable, supporting accountability in planning decisions.

6. References

[1] Masik G, Sagan I, Scott JW. Smart City strategies and new urban development policies in the Polish context. *Cities*. 2021;108:102970. https://doi.org/10.1016/j.cities.2020.102970

[2] Schieferdecker I, Tcholtchev N, Lämmel P, Scholz R, Lapi E. Towards an Open Data Based ICT Reference Architecture for Smart Cities. *CeDEM 2017: Conference for E-Democracy and Open Government*; 2017:184–193. https://doi.org/10.1109/CeDEM.2017.18

[3] Costa DG, Bittencourt JCN, Oliveira F, Peixoto JPJ, Jesus TC. Achieving Sustainable Smart Cities through Geospatial Data-Driven Approaches. *Sustainability*. 2024;16(2):640. https://doi.org/10.3390/su16020640

[4] Krauze-Maślankowska P. Open Data and Smart City Initiatives for Digital Transformation in Public Sector in Poland: A Survey. In: Wrycza S, Maślankowski J, eds. *Digital Transformation: 13th PLAIS EuroSymposium on Digital Transformation, Sopot, Poland, 23 September 2021*. Lecture Notes in Business Information Processing, vol. 429. Cham: Springer; 2021:71–81. https://doi.org/10.1007/978-3-030-85893-3_5

[5] Skiba Ł. Open Data as a Fundamental Element of the Development of the Smart City. *Scientific Papers of Silesian University of Technology: Organization and Management Series*. 2023;(180). https://doi.org/10.29119/1641-3466.2023.180.29

[6] Mazzetto S. A Review of Urban Digital Twins Integration, Challenges, and Future Directions in Smart City Development. *Sustainability*. 2024;16(19):8337. https://doi.org/10.3390/su16198337