2. Program Development

Parcel Value Impact Analysis

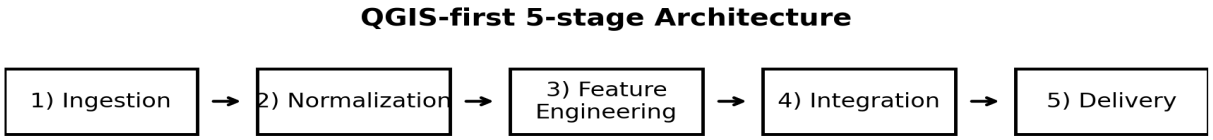
Product Documentation Methodology (QGIS-First, Implementation-Aligned)

Document type: Product Methodology / Implementation Specification

Version: v1.2 (validated)

Scope: Building a single parcel-centric analysis pipeline for England by combining parcel geometry with value-impacting environmental features

Tools: QGIS Desktop, GeoPackage/GeoJSON, HM Land Registry INSPIRE, OSM (Geofabrik), Streamlit delivery layer



1) Problem Statement

Valuation and investment teams need to assess parcel-level location strengths and weaknesses quickly, consistently, and with repeatable logic. In practice, parcel boundaries and environmental context are distributed across separate sources and formats, which increases preparation time and creates analyst-to-analyst variation. This directly reduces comparability of parcel decisions [9; 7; 34].

A second operational issue is performance: over-granular category taxonomies and large raw OSM layers can degrade desktop GIS and app responsiveness unless data is reduced early in the workflow [35; 36].

2) Product Objective and Success Criteria

2.1 Primary objective

Integrate INSPIRE parcel geometry with high-impact OSM-derived context layers to produce a parcel-centric decision-support dataset.

2.2 MVP success criteria

- Parcel geometry is available in one analysis project.
- Parcel-level proximity metrics are computed for: health, education, safety/security, transport access, water/coastal influence.
- Outputs are easy to map, filter, and export.
- The same workflow can be reused for additional regions.

3) System Design Principles and Architecture

3.1 QGIS-first architecture (5 stages)

- Ingestion - download and register data into project structure
- Normalization - CRS, geometry, and schema standardization
- Feature Engineering - parcel-level metric generation
- Integration - consolidate metrics via spatial joins

- Delivery - maps, tables, and export packages

This architecture provides low setup overhead, high repeatability, and a clear path to automation as scale grows.

3.2 Design principle: Minimal, high-impact category set

To maximize explainability and performance, keep category design concise:

- Health access (hospital/clinic)
- Education access (school)
- Safety access (police)
- Transport access (major roads + transit)
- Water/coastal influence (river/coast proximity)

4) Data Sources, Roles, and Reliability Framing

4.1 Parcel dataset (primary geometric reference)

- Source: HM Land Registry - INSPIRE Index Polygons
- Format: GML/ZIP
- Role: Parcel baseline (primary geometric reference)

INSPIRE provides positional representation for registered freehold properties in England and Wales; it is not a legal boundary determination and should be treated as indicative extent/position [10; 12; 9].

4.2 Spatial context dataset (value-driving layers)

- Source: Geofabrik - OSM England
- Format: OSM PBF and SHP-derived layers
- Role: Value-driving contextual layers (schools, hospitals, police, transport, etc.) [7]

4.3 Administrative boundaries (optional, recommended)

- Source: ONS - Local Authority Districts (May 2025) Boundaries UK BFE (V2)
- Role: Regional aggregation, reporting, benchmarking [18]

4.4 Reliability note

- INSPIRE geometry should be treated as indicative, not legal proof [9; 12].
- OSM-derived data quality and update depth vary by location because it is community-maintained [21].

4.5 Data freshness policy

- INSPIRE and OSM updates are checked on a regular cadence.
- Snapshots are versioned and date-stamped.
- The dataset version used in each run is recorded in metadata.
- Raw files are stored with monthly versioning (/data/raw/YYYY-MM).

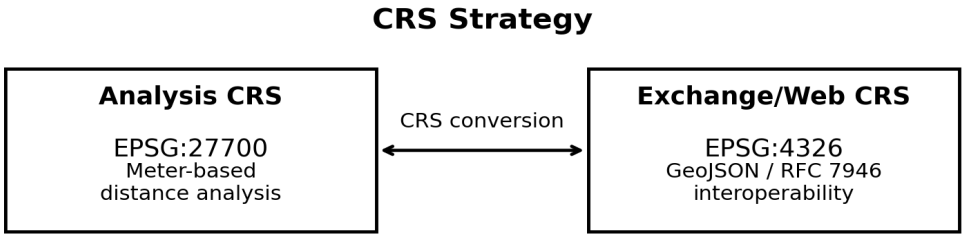
4.6 Licensing and attribution

Geofabrik distributions inherit ODbL obligations from OpenStreetMap. Attribution and derivative database obligations must be respected in outputs and documentation [6; 21].

5) CRS and Geometry Strategy

5.1 CRS strategy

- Analysis CRS: EPSG:27700 (meter-based distance analysis for Great Britain)
- Exchange/Web CRS: EPSG:4326 (GeoJSON/RFC 7946 interoperability) [24; 5; 2]



5.2 Critical geometry correction (mixed geometry handling)

In QGIS, Merge Vector Layers expects homogeneous geometry inputs per merge operation. Point/line/polygon layers should not be forced into one mixed merge step [39].

Correct pattern:

- keep separate geometry layers during processing,
- publish a multi-layer container (GeoPackage) as the master artifact.

6) End-to-End Workflow (Detailed)

6.1 Data acquisition / intake

- Download INSPIRE index polygon packages.
- Download Geofabrik OSM England extracts.
- Apply folder convention:
- /data/raw/inspire/
- /data/raw/osm/
- /data/processed/
- /exports/
- Maintain monthly raw-data versioning (/data/raw/YYYY-MM).

6.2 Load into QGIS

- Add INSPIRE (GML/ZIP) vector layers.
- Add OSM point/line/polygon layers.
- Run initial checks: geometry type, feature counts, field schema, null rates.

6.3 Preprocessing and quality

- Reproject all layers to EPSG:27700.
- Repair invalid geometries with Fix Geometries.
- Remove out-of-AOI features via Clip / Extract by extent [34; 35; 36].

6.4 Build parcel base layer

- Set INSPIRE as source of truth.
- Preserve unique parcel ID.

- Generate derived geometry fields:
- parcel_area_m2
- parcel_perimeter_m
- optional compactness/shape metric
- (Field Calculator / Add geometry attributes [40])

6.5 Category filtering (high-impact core set)

- Keep only health, education, safety, transport, and water/coastal categories.
- Apply AOI clipping and class reduction before distance operations for performance.

6.6 Generate value-driver features

Distance metrics

- MVP uses Euclidean (straight-line) distance.
- Later versions can add network travel-time.

Buffer-count metrics

- Counts within thresholds such as 500 m / 1 km.

Example fields

- dist_school_m
- dist_hospital_m
- dist_police_m
- cnt_busstop_500m
- cnt_rail_1km

Also core-category metric schema:

- dist_health_m
- dist_education_m
- dist_security_m
- dist_transport_m
- dist_water_m

Typical QGIS tools

- Join attributes by nearest
- Join attributes by location
- Distance matrix
- Distance to nearest hub
- [37; 38]

6.7 Integration

- Attach metrics to parcel ID using spatial joins.
- Null handling rules: no measured distance -> NULL; no count found -> 0
- Data typing standard: distance: float/int (meters); count: int; score/rate: float [37; 41]

6.8 Output and delivery

- Master artifact: analysis_output.gpkg (multi-layer container); parcel-centric integrated layer: parcels_enriched.gpkg / parcels_enriched
- Optional distribution: GeoJSON exports by category/geometry in EPSG:4326
- Map/reporting packages: project style + thematic legend; PNG/PDF map packages
- Table exports: CSV/Excel
- Metadata package: source, version, date, CRS, processing steps, confidence flags (recommended)

7) User-Facing Outputs

7.1 Map layers

- Base parcel layer (INSPIRE)
- Thematic OSM layers
- Enriched parcel layer (single integrated layer)

7.2 Decision-support views

- Parcel classification by proximity score
- Suitability heatmap (graduated/choropleth)
- Rule-based filtering (e.g., school <1 km, hospital <2 km, main transport <800 m)

7.3 Exports by audience

- GeoPackage: GIS teams
- CSV/Excel: business teams
- PNG/PDF: reporting packages

7.4 Recommended parcel-centric minimum schema

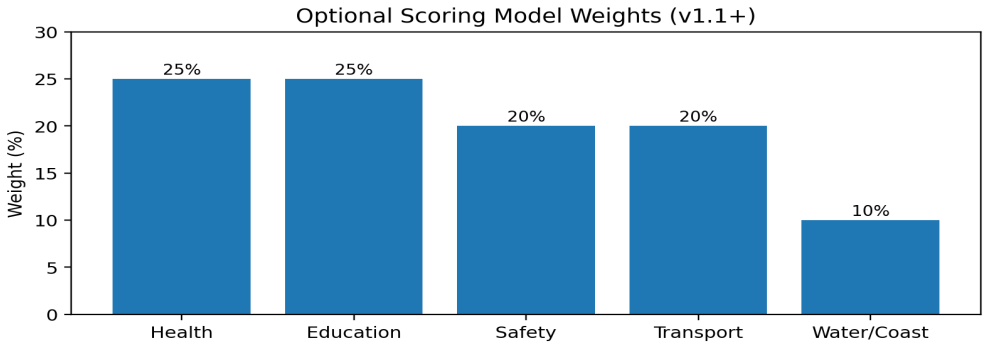
- parcel_id
- area_m2
- perimeter_m
- centroid_lat
- centroid_lon
- dist_health_m
- dist_education_m
- dist_security_m
- dist_transport_m
- dist_water_m
- score_location (optional)
- updated_at
- source_version

8) Optional Scoring Model (v1.1+)

Normalize distance metrics to 0-100 (nearer = better), then apply interpretable weights:

- Health: 25%
- Education: 25%
- Safety: 20%
- Transport: 20%
- Water/coast: 10%

score_location = sum_i (w_i x normalized_i)



This keeps the decision model interpretable and auditable.

9) Performance Levers

- Keep the category system limited to high-impact classes.
- Clip by AOI as early as possible.
- Remove irrelevant OSM classes before distance operations.
- Use GeoPackage as the master working/output format.
- Simplify display geometries where analytically acceptable.
- Persist intermediate layers to disk (avoid temporary-only chains).
- Apply viewport/extent-based rendering for very large layers.
- Use prepared/simplified Parquet and viewport filtering where applicable (as observed in MVP practice).

10) Quality Assurance (QA) and Definition of Done

10.1 QA checklist

- Correct CRS at each stage? (analysis: EPSG:27700, exchange: EPSG:4326)
- Geometry validity fixed?
- Null/duplicate parcel IDs?
- Any record loss after joins?
- Null rates within accepted thresholds for critical fields?
- Any impossible distance values (negative/extreme anomalies)?
- Naming schema consistent?
- Same input produces same output?
- Export files reopen correctly in target tools?
- Symbology and labels readable?

10.2 Definition of Done

- Enriched parcel layer is produced.
- At least 5 core spatial features are integrated.
- Export files and metadata are complete.
- Pipeline steps are documented and reproducible.

11) Risks and Mitigation

- OSM coverage/freshness variance - Impact: underrepresentation in some areas - Mitigation: coverage reporting + low-confidence flag
- Geometry/topology issues - Impact: join errors, incorrect metrics - Mitigation: mandatory geometry-repair QA checkpoint
- Distance metric interpretation risk - Impact: inconsistent decision interpretation - Mitigation: explicit note ("MVP uses straight-line distance"); add network-time in later versions
- Legal interpretation risk - Impact: indicative parcel extents mistaken as legal boundaries - Mitigation: explicit methodological disclaimer [12]

12) Implementation Alignment Check (MVP Validation)

12.1 Confirmed in current implementation

- File-driven parcel ingestion (GPKG/Parquet workflows)
- Parcel baseline preference using NATIONALCADASTRALREFERENCE when available
- CRS conversion between EPSG:4326 (display) and EPSG:27700 (metric calculations)
- Parcel geometry metrics (area/perimeter/centroid)
- Nearest-unit distance computation and map-based parcel interaction
- Performance levers: prepared/simplified Parquet, geometry simplification, viewport filtering

12.2 Partial or not yet implemented

- Formal, always-on QGIS preprocessing chain (Fix Geometries + AOI Clip as mandatory gates)
- Fully materialized parcel-wide dist_* feature store for all parcels in one persisted analytical layer
- Canonical analysis_output.gpkg as the default delivery artifact
- Formal source freshness/quality metadata contract and confidence flags in every output

12.3 Validation scope note

Validation was performed against packaged application behavior and exposed workflow logic. Internal helper modules referenced by the app are not fully packaged in the archive, so some deep internals are inferred from callable flow and UI behavior.

13) Roadmap

V1

- QGIS-based manual/semi-automated pipeline
- Parcel-level feature integration
- Single enriched output layer

V1.1

- Semi-automation with Model Builder
- Regional batch processing
- Standard scoring templates

V2

- Python + PostGIS pipeline
- API-driven updates
- Self-service query interface

V3

- Hybrid spatial + market signal scoring
- Explainable scoring engine
- If/then scenario analysis

14) Documentation Standard

- Unit suffixes are mandatory (_m, _km, _cnt, _score)
- Source/version/date are mandatory in metadata
- Release changes are logged in changelog
- Pipeline must be quickly reproducible by another analyst

15) Executive Methodology Summary

This product integrates HM Land Registry INSPIRE parcel geometries with Geofabrik OSM England thematic layers in QGIS. The process includes data acquisition, CRS standardization (EPSG:27700), geometry repair, AOI clipping, parcel-level feature engineering, and spatial joins. The result is a single parcel-level analytical layer with accessibility/proximity indicators for decision support [10; 12; 7; 24; 34; 35; 36; 37].

The key methodological correction is proper mixed-geometry handling during processing, while keeping GeoPackage as the master artifact for stable multi-layer delivery [19; 39].

References

[1] Bedford Borough Council. (n.d.). Brownfield land register. Bedford Borough Council. Retrieved February 9, 2026, from https://www.bedford.gov.uk/planning-and-building-control/planning-policy/brownfield-land-register

[2] Butler, H., Daly, M., Doyle, A., Gillies, S., Hagen, S., & Schaub, T. (2016). The GeoJSON format (RFC 7946). Internet Engineering Task Force. https://www.rfc-editor.org/rfc/rfc7946

[3] Department for Levelling Up, Housing and Communities. (n.d.-a). CSVW table schemas. Energy Performance of Buildings Open Data. Retrieved February 9, 2026, from https://epc.opendatacommunities.org/docs/csvw

[4] Department for Levelling Up, Housing and Communities. (n.d.-b). Energy Performance of Buildings documentation: Guidance. Energy Performance of Buildings Open Data. Retrieved February 9, 2026, from https://epc.opendatacommunities.org/docs/guidance

[5] EPSG Registry. (2021). OSGB36 / British National Grid (EPSG:27700). International Association of Oil & Gas Producers. https://epsg.org/crs_27700/OSGB36-British-National-Grid.html

[6] Geofabrik GmbH. (n.d.). Downloads [Service information].

[7] Geofabrik GmbH. (n.d.). England: OpenStreetMap data extracts. Retrieved February 9, 2026, from https://download.geofabrik.de/europe/united-kingdom/england.html

[8] HM Land Registry. (2016, June 14). How to access HM Land Registry Price Paid Data. GOV.UK. https://www.gov.uk/guidance/about-the-price-paid-data

[9] HM Land Registry. (2020, July 3). INSPIRE index polygons spatial data [Guidance]. GOV.UK. https://www.gov.uk/guidance/inspire-index-polygons-spatial-data

[10] HM Land Registry. (2021, September 16). INSPIRE Index Polygons spatial data [Dataset]. data.gov.uk.

[11] HM Land Registry. (2026, January 29). Price Paid Data. GOV.UK. https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads

[12] HM Land Registry. (n.d.). Download Index polygons spatial data (INSPIRE) [Data service]. Use Land and Property Data Service. https://use-land-property-data.service.gov.uk/datasets/inspire/download

[13] HM Land Registry Open Data. (n.d.). Search the price paid dataset. Retrieved February 9, 2026, from https://landregistry.data.gov.uk/app/ppd/

[14] Homes England. (2023, April 4). Using the Homes England Land Hub. GOV.UK. https://www.gov.uk/guidance/using-the-homes-england-land-hub

[15] Homes England. (n.d.). Land Hub sites [Dataset]. data.gov.uk. Retrieved February 9, 2026, from https://www.data.gov.uk/dataset/630b2ebc-157d-44e2-9a6e-2cf9cc8ea424/land-hub-sites

[16] Ministry of Housing, Communities and Local Government. (2017, July 28). Brownfield land registers. GOV.UK. https://www.gov.uk/guidance/brownfield-land-registers

[17] Office for National Statistics. (2025). ONS Postcode Directory (November 2025) for the UK (Hosted Table). data.gov.uk. https://www.data.gov.uk/dataset/56a898b9-8a3f-4c5f-8932-a03d958e0a43/ons-postcode-directory-november-2025-for-the-uk-hosted-table

[18] Office for National Statistics. (2025). Local Authority Districts (May 2025) Boundaries UK BFE (V2) [Dataset]. data.gov.uk.

[19] Open Geospatial Consortium. (2023). OGC GeoPackage encoding standard (Version 1.4.0). https://www.ogc.org/standard/geopackage/

[20] Open Geospatial Consortium. (n.d.). GeoPackage. Retrieved February 9, 2026, from https://www.ogc.org/standards/geopackage/

[21] OpenStreetMap Foundation. (n.d.). Copyright and license. Retrieved February 9, 2026, from https://www.openstreetmap.org/copyright

[22] OpenStreetMap Wiki contributors. (n.d.-a). Map features. OpenStreetMap Wiki. Retrieved February 9, 2026, from https://wiki.openstreetmap.org/wiki/Map_features

[23] OpenStreetMap Wiki contributors. (n.d.-b). Overpass turbo/GeoJSON. OpenStreetMap Wiki. Retrieved February 9, 2026, from https://wiki.openstreetmap.org/wiki/Overpass_turbo/GeoJSON

[24] Ordnance Survey. (2024, May 14). Coordinate reference systems. OS National Geographic Database documentation.

[25] Ordnance Survey. (2025). OS Open UPRN. data.gov.uk. https://www.data.gov.uk/dataset/c4f80d19-8cfa-4bf6-a283-83183842f876/os-open-uprn

[26] Ordnance Survey. (n.d.-a). OS Open Roads. OS Data Hub. Retrieved February 9, 2026, from https://osdatahub.os.uk/downloads/open/OpenRoads

[27] Ordnance Survey. (n.d.-b). OS Open Rivers. OS Data Hub. Retrieved February 9, 2026, from https://osdatahub.os.uk/downloads/open/OpenRivers

[28] Planning Data. (n.d.-a). Brownfield land [Dataset]. Retrieved February 9, 2026, from https://www.planning.data.gov.uk/dataset/brownfield-land

[29] Planning Data. (n.d.-b). Brownfield site [Dataset]. Retrieved February 9, 2026, from https://www.planning.data.gov.uk/dataset/brownfield-site

[30] Planning Data. (n.d.-c). Public transport access node [Dataset]. Retrieved February 9, 2026, from https://www.planning.data.gov.uk/dataset/transport-access-node

[31] Planning Data. (n.d.-d). Title boundary [Dataset]. Retrieved February 9, 2026, from https://www.planning.data.gov.uk/dataset/title-boundary

[32] Planning Data. (n.d.-e). Datasets. Retrieved February 9, 2026, from https://www.planning.data.gov.uk/dataset/

[33] PostGIS Project Steering Committee. (n.d.). ST_Distance. PostGIS Documentation. Retrieved February 9, 2026, from https://postgis.net/docs/ST_Distance.html

[34] QGIS Development Team. (n.d.-a). Working with projections. QGIS Documentation. https://docs.qgis.org/latest/en/docs/user_manual/working_with_projections/working_with_projections.html

[35] QGIS Development Team. (n.d.-b). Vector overlay: Clip / Extract by extent. QGIS Documentation. https://docs.qgis.org/latest/en/docs/user_manual/processing_algs/qgis/vectoroverlay.html

[36] QGIS Development Team. (n.d.-c). Vector geometry: Fix geometries. QGIS Documentation. https://docs.qgis.org/latest/en/docs/user_manual/processing_algs/qgis/vectorgeometry.html

[37] QGIS Development Team. (n.d.-d). Vector general: Join attributes by nearest / Join attributes by location. QGIS Documentation. https://docs.qgis.org/latest/en/docs/user_manual/processing_algs/qgis/vectorgeneral.html

[38] QGIS Development Team. (n.d.-e). Vector analysis: Distance matrix / Distance to nearest hub. QGIS Documentation. https://docs.qgis.org/latest/en/docs/user_manual/processing_algs/qgis/vectoranalysis.html

[39] QGIS Development Team. (n.d.-f). Vector general: Merge vector layers. QGIS Documentation. https://docs.qgis.org/latest/en/docs/user_manual/processing_algs/qgis/vectorgeneral.html

[40] QGIS.org. (n.d.-e). Vector geometry: Add geometry attributes. QGIS Documentation.

[41] QGIS.org. (n.d.-f). Working with joined tables and relations. QGIS Documentation.

[42] Streamlit. (n.d.-a). st.cache_data. Streamlit Docs. Retrieved February 9, 2026, from https://docs.streamlit.io/develop/api-reference/caching-and-state/st.cache_data

[43] Streamlit. (n.d.-b). Understanding widget behavior. Streamlit Docs. Retrieved February 9, 2026, from https://docs.streamlit.io/develop/concepts/architecture/widget-behavior

[44] Streamlit. (n.d.-c). Working with fragments. Streamlit Docs. Retrieved February 9, 2026, from https://docs.streamlit.io/develop/concepts/architecture/fragments

[45] Valuation Office Agency. (n.d.). How non-domestic property, including plant and machinery, is valued. GOV.UK. Retrieved February 9, 2026, from https://www.gov.uk/guidance/how-non-domestic-property-including-plant-and-machinery-is-valued