TECHNICAL DOCUMENTATION

EarthDefine Building Footprints 3D: Technical Specification

How EarthDefine builds, verifies and delivers nationwide US building footprints: the three units the data counts in, where the imagery and LiDAR come from, what does and does not count as a building, how accuracy is checked, the full polygon and point schemas, and how a delivery ships.

Revised 11 November 2026

Current version https://www.earthdefine.com/evaluate/buildings-tech-spec/

This document is generated from the page above, which is the source of record.
Check that page for the current revision.

On this page

1. What is Building Footprints 3D?
2. How much of the US does it cover, and how many buildings are there?
3. Where does the data come from, and how is it produced?
4. How often is it updated, and do identifiers stay stable across releases?
5. What counts as a building, and what doesn't?
6. How is accuracy measured and verified?
7. What's in the schema?
8. How is it delivered, and what does it cost to license?
9. Need something the standard specification doesn't cover?

1. What is Building Footprints 3D?

EarthDefine Building Footprints 3D is a nationwide dataset of US building footprints delivered as two geometry layers: building polygons, and address points carrying the attribute table. Footprints are extracted by EarthDefine's AI classification models from LiDAR and 30cm to 60cm color-infrared aerial imagery, then split using parcel data, addressed from multiple sources, and enriched with LiDAR-measured 3D attributes.

Our dataset distinguishes three units:

- **Structure** or `str_UUID` identifies one contiguous physical building mass. Its mapped footprint follows the roof edge, including roof overhangs.
- **Building** or `bld_UUID` is a structure split at parcel boundaries: a townhome row of eight units on eight parcels is one structure and eight buildings.
- **Address** or `address` is an address record inside a building; multiple points per building are normal, and every point falls spatially within its polygon.



Figure 1: EarthDefine Building Footprints: structures, buildings, and address points

How these three units combine

Not all buildings, structures, and address points map one to one. Below are six illustrated cases that describe how they are handled:

1. A single-family home that sits on one parcel: one parcel, one structure, one building, and carries one address.
2. A duplex or small multi-unit building that sits on one parcel: one parcel, one structure, one building with two units, with one address per unit.
3. A large house that sits across two parcels: two parcels, one structure, one building, and carries one address.
4. A townhome row across three parcels with a mix of two units that have one address each, and a third which has two addresses because it includes a rented basement unit: three parcels, still one structure, three buildings, but four addresses (one + one + two).
5. A house and a shed that sit on one parcel: one parcel, two structures, two buildings, each carrying the same address.
6. A house and a separate studio that sit on one parcel: one parcel, two structures, two buildings, two addresses.
7. A multi-storied building in a single parcel: one parcel, one structure, four buildings, each carrying a different address.

8. A large rural parcel with a house, a shed, and two industrial units: one parcel, four structures, four buildings, the house and shed share one address and the industrial units have two have distinct addresses (one + one + one).

This explains why the buildings count and the structures count (see Section 2) diverge, and why per-parcel aggregation needs to be split. The variation is the direct result of buildings not always mapping one-to-one with addresses even when parcels do.

To view a building you are familiar with, try the Buildings API - it will return the str_UUID, bld_UUID and address records for any address or coordinates.

2. How much of the US does it cover, and how many buildings are there?

Coverage: contiguous US, Puerto Rico, US Virgin Islands, Hawaii, Mariana Islands, American Samoa, and Alaska (partial), covering a total of 3,227 out of the 3,244 counties.

	COUNT	DETAILS
Buildings	203,474,752	Includes parcel-level splits
Contiguous structures	194,555,029	Used for comparisons with datasets that do not split buildings at parcels
Structures with LiDAR-measured heights	179,167,746	92% of all contiguous structures

Context: The buildings figure is higher than the contiguous structures figure because of the parcel splits described in Section 1. Both figures are provided to support appropriate comparisons with datasets that do not split structures at parcel boundaries.

Check coverage for a specific address or region here: EarthDefine Buildings API

3. Where does the data come from, and how is it produced?

Footprints derive from three source types:

- USDA NAIP* color-infrared imagery at 30cm to 60cm resolution
- State and local orthoimagery
- LiDAR point clouds

*U.S. Department of Agriculture National Agriculture Imagery Program

Every address point records the most recent source that detected its building (`Source`) and the date of that source (`SourceDate`), making the provenance queryable per record rather than stated only at the dataset level.

Footprint source dates and height-data dates can differ. `Source` and `SourceDate` describe the source that most recently detected the building; they do not necessarily identify the source or acquisition date used for its height measurements. Consult `LiDAR_Year` separately when assessing the available LiDAR vintage.

Production runs as a staged pipeline: imagery is classified by the extraction models; classified output is tiled and regularized into clean polygon geometry; LiDAR enriches regularized polygons with 3D measurements and contributes new and refined footprints where imagery alone underperforms; polygons are split at parcel boundaries; and buildings are addressed from multiple sources. Data is processed at the county level, with each county processed separately. Record counts are reconciled at every stage.

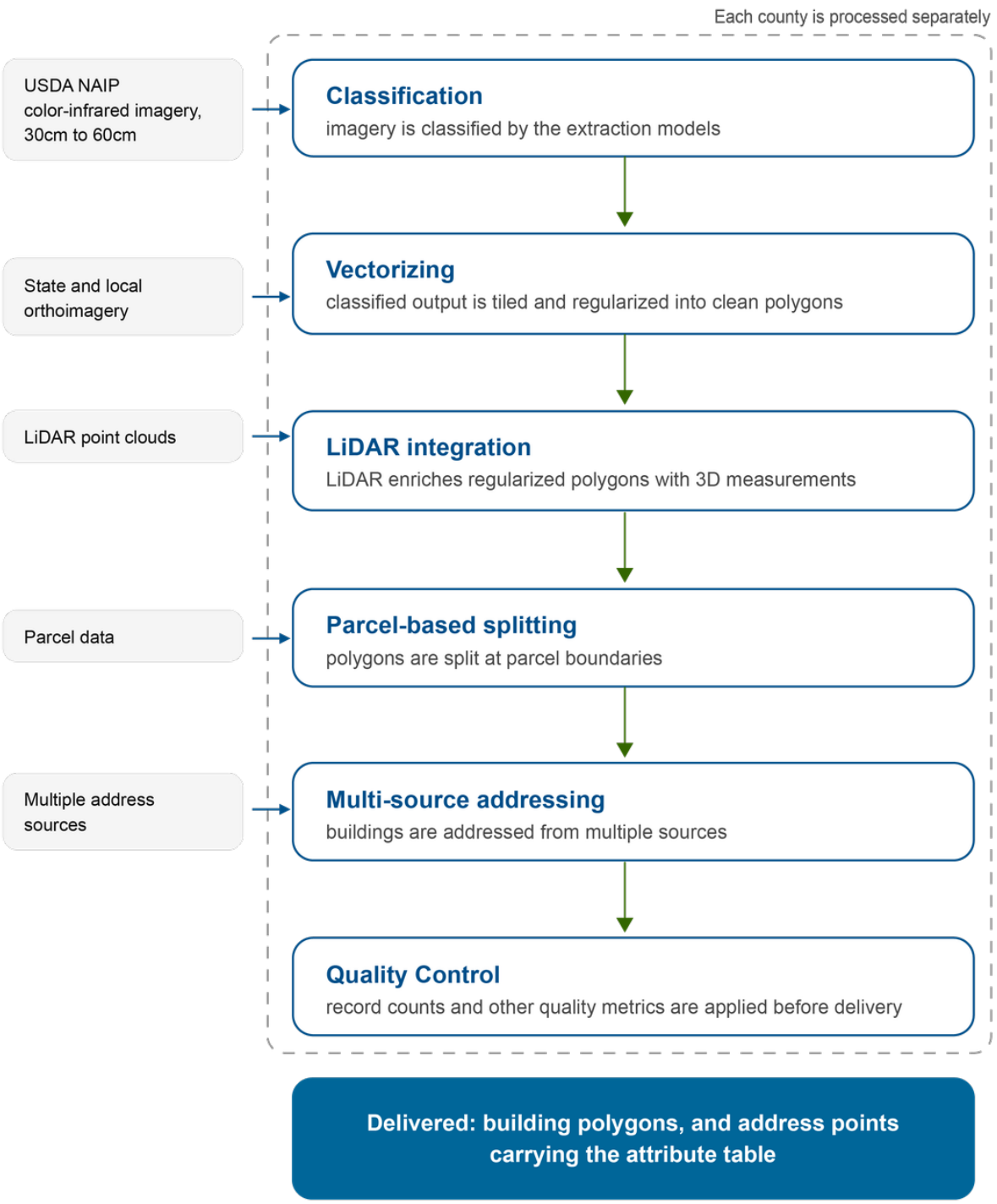


Figure 3: EarthDefine Building Footprints 3D production pipeline diagram.

For details about how output is reviewed and corrected before it ships, and how those corrections feed future releases, please refer to Section 6.

4. How often is it updated, and do identifiers stay stable across releases?

We release quarterly. Each release carries model improvements plus any new LiDAR and state or local imagery that has arrived since the last one, so footprints, heights, and parcel splits keep improving between imagery refreshes. The source imagery itself is refreshed on a fixed two-year cycle per state, about half the states each year. In the most recent quarter we added 1,192,508 structures and extended measured heights to a further 3,080,541.

Identifiers persist across releases: `str_UUID` and `bld_UUID` are stable. In cases when the building shifts or changes significantly, new UUIDs are assigned.

A quarterly product release does not mean every location has newly acquired imagery or LiDAR. Available source dates vary by location, including within a county. Buildings detected in newer imagery may have footprints but no height or volume values if suitable LiDAR measurements are not available.

5. What counts as a building, and what doesn't?

A building is a ***permanent structure with a roof***. That definition is applied consistently and decides most edge cases. Here are some common rules:

Captured as a building:
Houses, commercial buildings, and industrial buildings; parking garages; trailer homes and mobile homes sited in trailer parks; gazebos, covered baseball dugouts, and covered parking; glass sunrooms and pool houses; covered pool enclosures (usually attached to the main structure); greenhouses of clear glass or plastic; covered walkways connecting complexes of the same facility (hospitals, campuses); gas-station canopies; water towers; stadium seating with an alternate use beneath it (concessions, restrooms, locker rooms); solar panels serving a secondary structural use such as covered parking; silos and large cisterns; roofed structures at the end of docks; detached sheds and outbuildings, which are delivered as separate buildings.

Garages. Covered garages attached to the main building are included in its footprint and, where the required height data is available, its Volume calculation. Detached garages are captured as separate buildings associated with the same address or parcel. Uncovered parking areas are excluded. Open-sided structures with a permanent roof are treated as covered parking and are included.

For a plain-language introduction, read What counts as a building in a building footprint dataset?



Figure 5A: Examples of what is captured as buildings.
Top left: Covered walkways at an airport terminal | Top right: Gas station canopy as a building
Bottom left: Large cistern and silos as buildings | Bottom right: Water tower of a town as a building

Not captured as a building:

RVs, campers and towed trailers anywhere on residential property, on roads, or at dealerships; boats and vehicles; shipping containers and semi-truck trailers; tents, canvas or mesh temporary structures; sporting-event seating with no alternate use beneath it, including small stands at school fields; standalone solar arrays; transmission towers and wind turbines; statues and monuments; uncovered docks, pools, trampolines, patios and porches; pool-adjacent cabanas and pergolas unless roofed; construction-material piles and pallets; mesh agricultural covers; road segments; industrial piping.

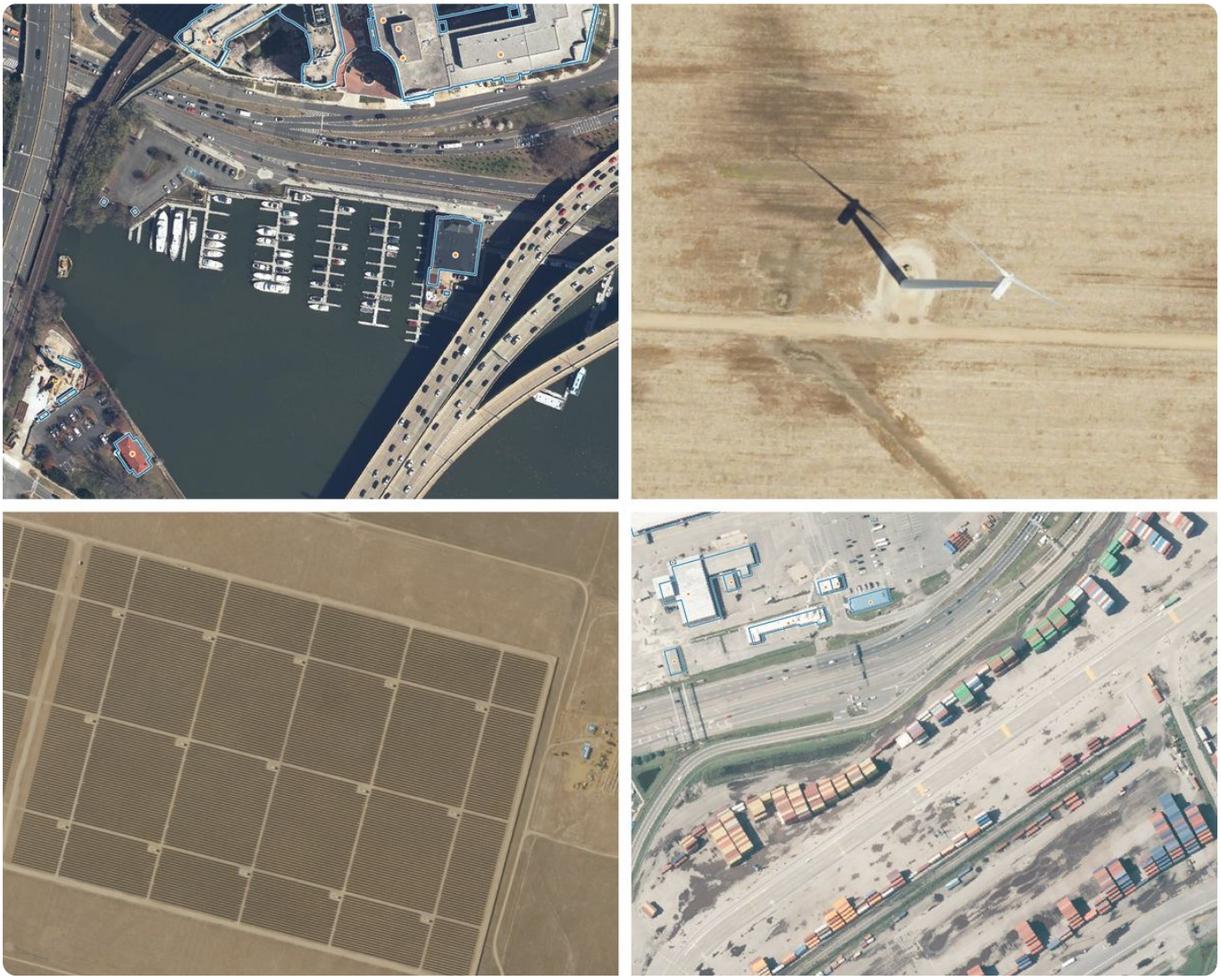


Figure 5B: Not captured examples.
Top left: Docks, boats, and vehicles | Top right: Wind turbines
Bottom left: Solar arrays | Bottom right: Shipping containers

Geometry rules.

Uncovered courtyards and patios inside a building's boundary are excluded from the footprint, as are adjacent pavement and bare soil. Above-ground decks and balconies are included, ground-level patios are not. The distinction is made from elevation evidence (stairs, furniture, cast shadow).

Building footprints follow the roof edge, including roof overhangs, rather than the exterior walls or foundation. The model attempts to correct for building lean caused by the viewing angle in the source imagery while preserving the roof-edge outline. Buildings partially obscured by tree canopy or shadow are captured on the estimated complete extent of the building rather than only the visible portion, and complex multi-material roofs are resolved as one footprint using shadow and context rather than split by roof color.

There is no minimum-size threshold published for capture.



Figure 5.1. Parcel split in practice.
Left: Aerial imagery of an eight-unit townhome row
Right: Delivered footprint layer for that imagery, showing one contiguous structure split into eight building polygons at the parcel lines.



Figure 5.2. Buildings under construction.
Left: A construction site in source imagery
Right: Delivered footprint layer for that imagery, correctly captured as a building despite being incomplete.



Figure 5.3. RV vs. trailer-park mobile home.
Left: An RV (intentionally not captured) on residential property
Right: Manufactured homes in a mobile home park (correctly captured as a buildings).

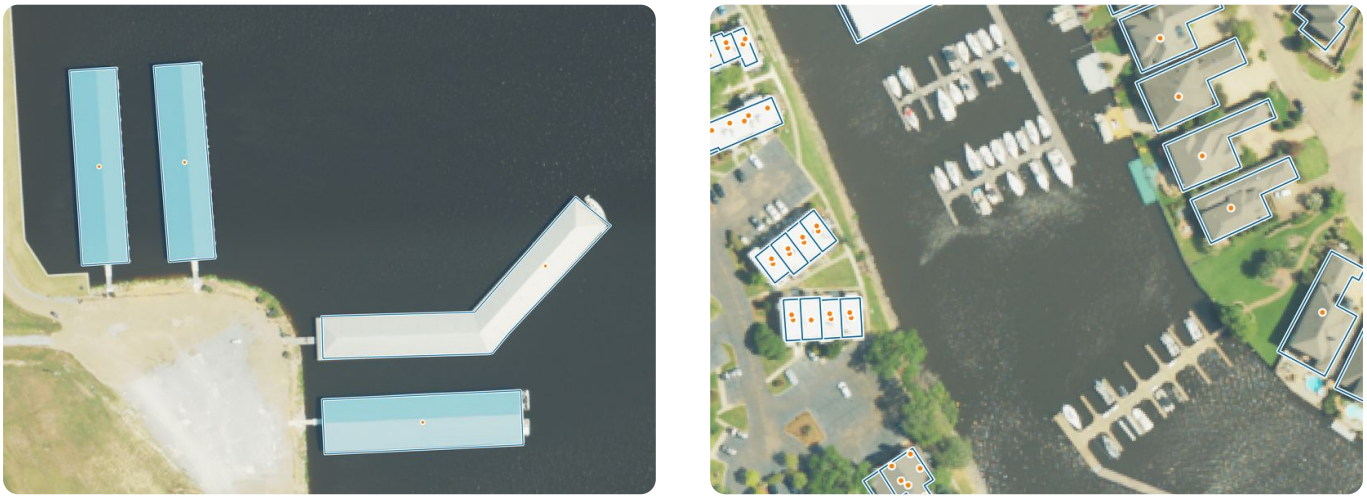


Figure 5.4. Roofed dock structure vs. uncovered docks.
Left: Roofed dock structures (correctly captured)
Right: Uncovered docks (intentionally not captured)



Figure 5.5. Solar panels serving covered parking vs. a standalone array.
Left: Solar panels over a parking lot (correctly captured)
Right: A standalone ground-mounted solar array (intentionally not captured)

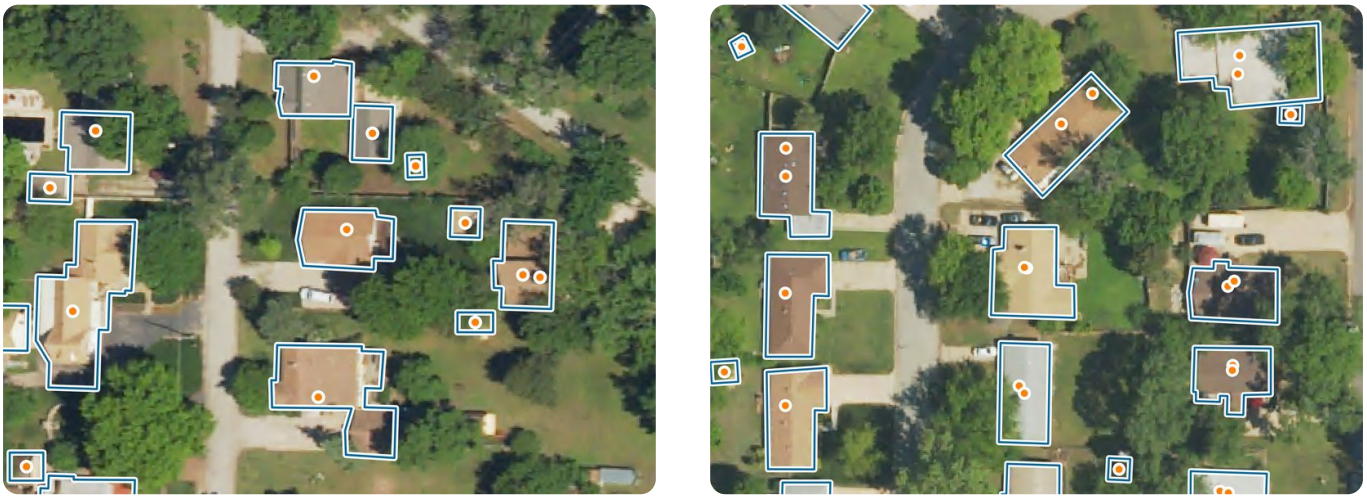


Figure 5.6. Above-ground deck vs. ground-level patio.
Left: A raised deck with visible support structure (correctly captured)
Right: A ground-level patio (intentionally excluded)
The distinction is made using elevation evidence.



Figure 5.7. Courtyard exclusion.
Left: Ariel imagery of building with open interior courtyards
Right: the delivered footprint with courtyard correctly excluded from the polygon

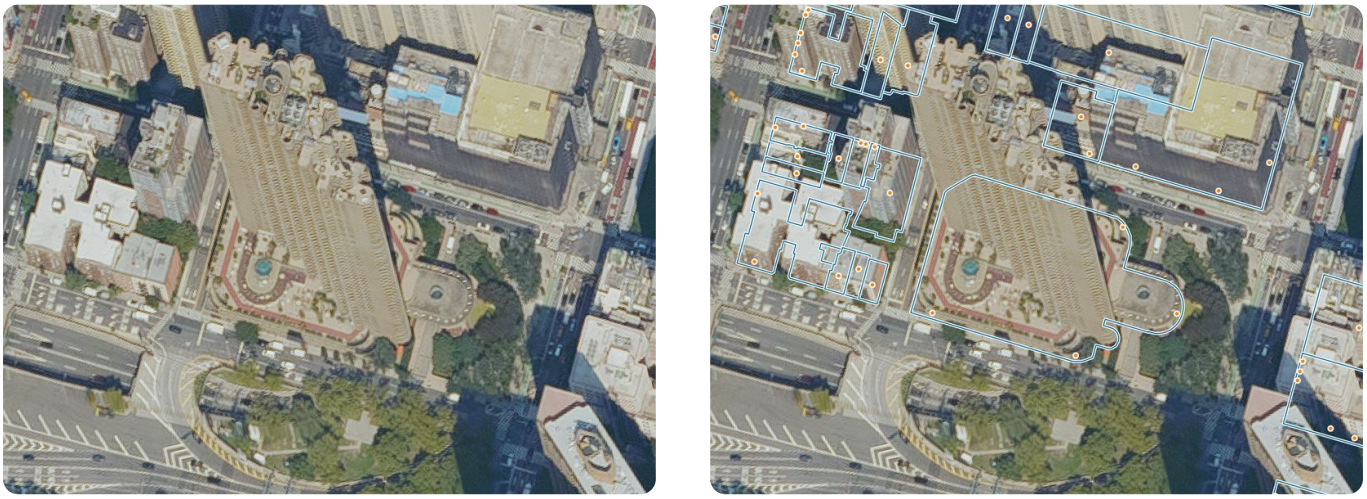


Figure 5.8. Multi-story building footprint.
Left: A multi-story building in oblique or angled imagery
Right: The delivered footprint, intended to follow the roof edge with correction for apparent building lean

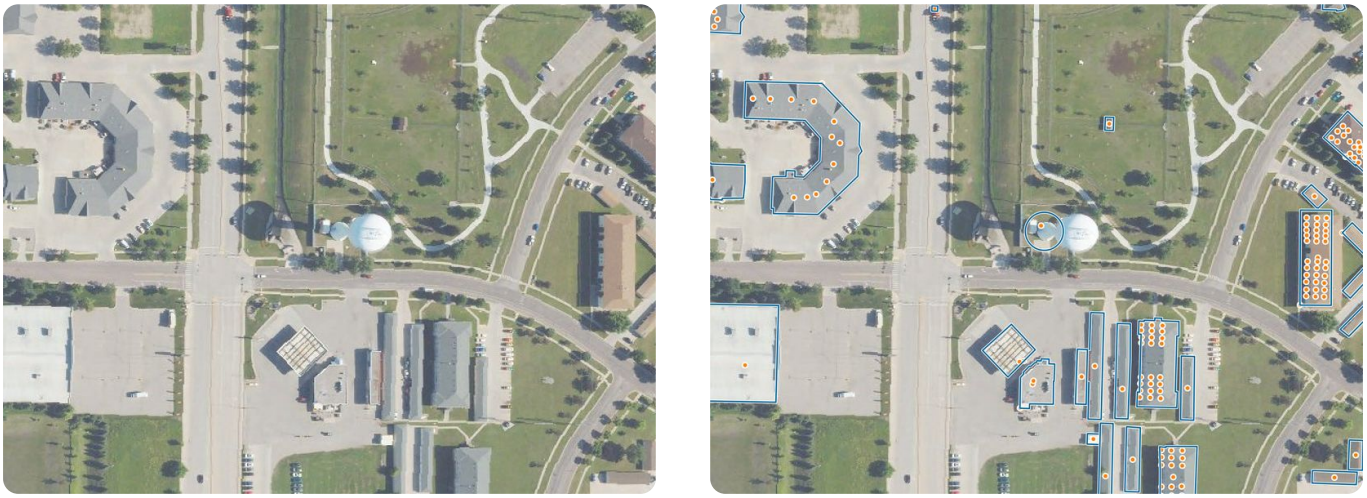


Figure 5.9. Building lean.
Left: A leaning building in oblique imagery
Right: The delivered footprint, with the model attempting to correct apparent building lean while preserving the roof-edge outline

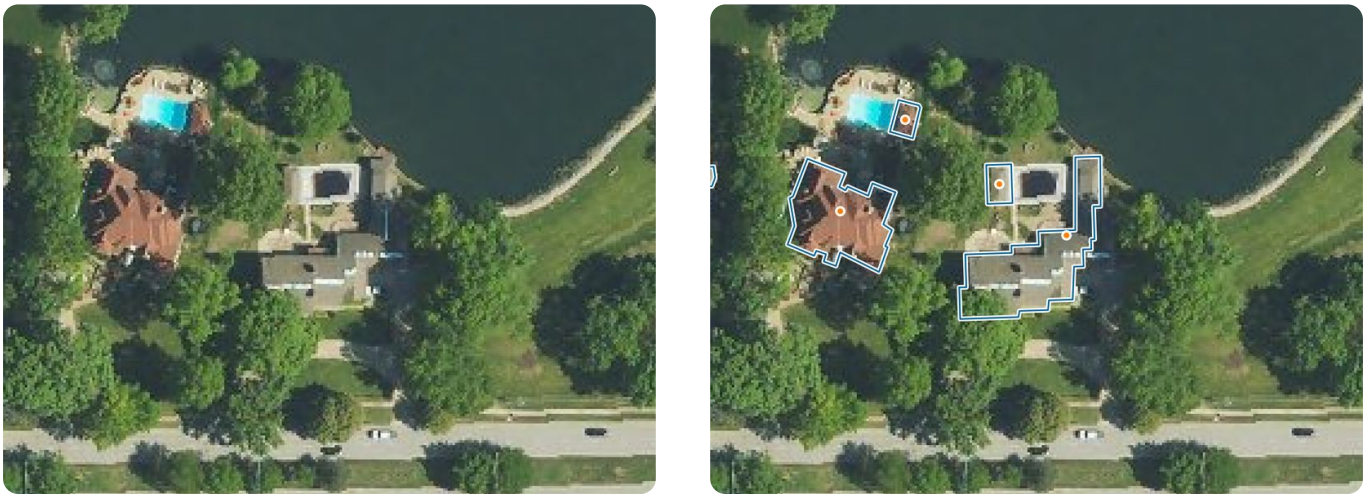


Figure 5.10. Canopy or shadow obstruction.
Left: A building partially hidden by tree canopy
Right: Delivered footprint, extended to the estimated complete extent rather than the visible portion.



Figure 5.11. Multi-material roof resolved as one footprint.
Left: A building with a roof in two or more materials or colors
Right: Delivered footprint, resolved as a single polygon rather than split by roof appearance

6. How is accuracy measured and verified?

Our quality assurance runs at three levels.

- **Pipeline stage gates.** Record counts are reviewed with every new data integration to ensure a net positive data improvement is achieved. Large fluctuations in building counts are investigated when new LiDAR or imagery is introduced.
- **Per-state release checks.** Before a state ships, a named checklist verifies, among other checks: UUID consistency between point and polygon layers (no extras, no missing); no null address fields; no zero values in ground-elevation fields (LAG, HAG, mean elevation); story-count outlier review; source-date correctness against the LiDAR and imagery inventories; gap investigation over the resolved layer; dissolution of overlapping polygons along tile edges; and spot checks of residential addresses against independent map services.
- **Human review.** County-level classifier output is reviewed and signed off by a named human data analyst; every county is hand-verified before it ships. Where output needs correction, human data analysts edit annotations against source imagery under the written guideline of section 5, and corrections feed the next model release. The review work is itself audited: annotation output is re-reviewed in structured feedback rounds that classify errors into four categories (omissions, commissions, incomplete annotations, mislabeling), with worked examples issued back to reviewers each round.

Accuracy is measured at different levels relevant to your project:

Height and volume accuracy. Height metrics are derived from LiDAR point clouds and delivered as multiple values per building, including maximum height, mean height, and maximum height including clutter, alongside ground elevations and roof slopes. We do not currently publish product-wide accuracy figures for MeanHeight or Volume. The vertical accuracy specification of the underlying LiDAR collection describes the

source measurements; it is not a validation result for the derived MeanHeight or Volume fields. Contact us for the source-LiDAR specification applicable to your area of interest.

Address confidence. Every address point carries `Add_Conf`, a low, medium or high confidence score based on the number of corroborating sources and the point's positional relationship to the building.

7. What's in the schema?

The polygon layer carries three fields (`str_UUID`, `bld_UUID`, `Area_SqFt`).

The point layer carries the full attribute table, addressing fields, `Add_Conf`, coordinates, 3D metrics (heights, ground elevation, roof slopes, stories, area, volume), Land Based Classification Standards (LBCS) land-use classification, a parcel identifier, source and source date, and a largest-structure-per-address flag, joined to the polygon layer on either UUID.

Coordinate system is geographic WGS84 (EPSG:4326), vertical datum NAVD88.

3D metrics require a 3D license and are computed from processed LiDAR point-cloud data; the LBCS land-use fields and parcel identifier (`ll_uuid`) derive from Regrid parcel data.

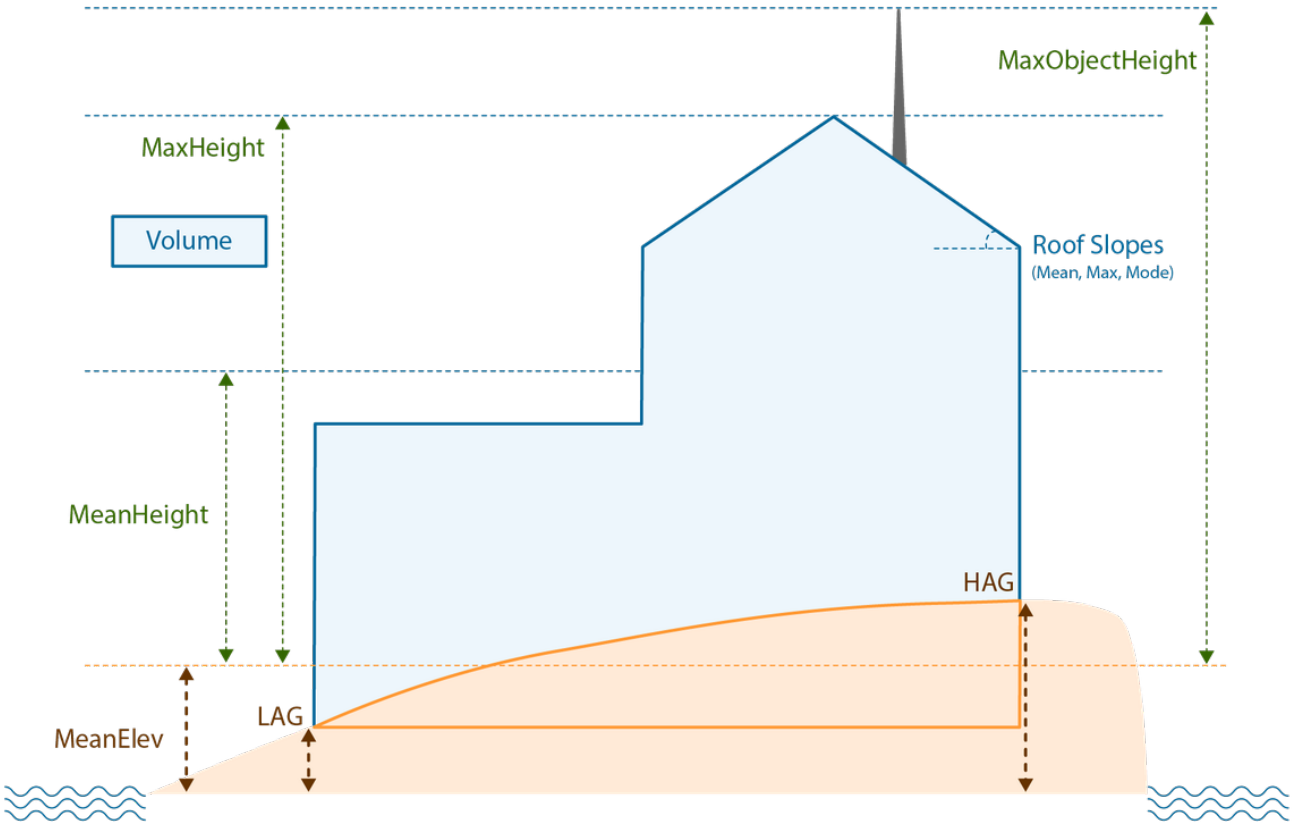


Figure 7: A visualization of building heights (Mean, Max, MaxObject), roof slopes (Mean, Max, Mode), volume, and ground elevations (Mean, LAG: Lowest Adjacent Grade, HAG: Highest Adjacent Grade)

For the authoritative field-by-field reference, definitions, and types, please refer to the data dictionary below.

Polygon Schema

FIELD	DESCRIPTION	EXAMPLE
str_UUID	Universal Unique ID of the entire structure	f985fa5b-9f2e-11eb-abd1-d89d67c66f28
bld_UUID	Universal Unique ID of the building from the parcel split	f9ea1bf5-9f2e-11eb-a32b-d89d67c66f28
Area_SqFt	Area of the building in square feet	24,052

Point Schema

FIELD	DESCRIPTION	EXAMPLE
str_UUID	Universal Unique ID of the entire structure associated with the Address	f985fa5b-9f2e-11eb-abd1-d89d67c66f28
bld_UUID	Universal Unique ID of the building from the parcel split associated with the Address	f9ea1bf5-9f2e-11eb-a32b-d89d67c66f28
Address	Full address of the building, including City, State, and ZIP Code	8875 NW 26th St, Doral, FL 33172
Address_No	Address number	8875
Street	Address street name	NW 26th St
SubAddress	Sub-unit Address	Unit C2
City	Address city	Doral
STPostal	Address USPS state abbreviation	FL
ZIP	Address Zip Code	33172
State	Address state name	Florida
County	Address county	Miami-Dade County
CTFIPS	County FIPS (Federal Information Processing Standard) Code (The first two digits are the State FIPS Code)	12086
Add_Conf	Address Confidence, dependent on the number of sources and positional relationship to sources, set as low, medium, or high	High
Latitude	Latitude in Decimal Degrees of the centroid of the building	25.798749
Longitude	Longitude in Decimal Degrees of the centroid of the building	-80.339941

FIELD	DESCRIPTION	EXAMPLE
MaxHeight*	Maximum height of the structure based on latest LiDAR data calculated in feet, excluding potential clutter like an antenna, overhanging tree, etc. (NULL denotes no elevation data)	14.48
MeanHeight*	Mean height of the structure based on latest LiDAR data calculated in feet (NULL denotes no elevation data)	11.58
MaxObjectH*	Maximum height of the structure based on latest LiDAR data calculated in feet, including potential clutter like an antenna, overhanging tree, etc. (NULL denotes no elevation data)	45.06
LAG*	Lowest Adjacent Grade: lowest observed ground elevation above sea level around the boundary of the building	249.77
HAG*	HAG - Highest Adjacent Grade: highest observed ground elevation above sea level around the boundary of the building	252.28
MeanElev*	Mean ground elevation in feet under the structure footprint	250.93
MeanSlope*	Mean roof slope in degrees 0-90 (NULL denotes no elevation data)	33
MaxSlope*	Max roof slope in degrees 0-90 (NULL denotes no elevation data)	45
ModeSlope*	Mode roof slope in degrees 0-90 (NULL denotes no elevation data)	30
ll_uuid**	Universal Unique ID of the parcel that intersects the building	9d79f331-9752-423a-b88e-15b5e9ed6b6a
lbcs_activity**	Land Use Code: Activity	3120
lbcs_activity_desc**	Land Use Code Description: Activity	Primarily goods storage or handling
lbcs_function**	Land Use Code: Function	3600
lbcs_function_desc**	Land Use Code Description: Function	Warehouse and storage services
lbcs_structure**	Land Use Code: Structure	2700
lbcs_structure_desc**	Land Use Code Description: Structure	Warehouse or storage facility
lbcs_site**	Land Use Code: Site	6000
lbcs_site_desc**	Land Use Code Description: Site	Developed site with buildings
lbcs_ownership**	Land Use Code: Site	4110

FIELD	DESCRIPTION	EXAMPLE
lbcs_ownership_desc**	Land Use Code Description: Site	City, Village, Township, etc.
Source***	Most modern source of the building footprint, either from aerial imagery or LiDAR	Aerial Imagery
SourceDate	Source data date formatted as MM/DD/YYYY or YYYY when exact capture date is unavailable, will match the source as the most modern date the building was detected	11/26/2019
LiDAR_Year	Acquisition year of the LiDAR data currently on file for the location. This may differ from SourceDate and does not by itself guarantee that height or volume values are available for a building.	2017
Area_SqFt	Area of the building in square feet	24,052
Stories*	Model estimated number of stories based on building attributes if elevation information is available	5
GrossArea*	Estimated square footage of the building (Area_SqFt * Stories)	8204.75
Volume*	Building volume in cubic feet, derived from height pixels within the mapped building footprint. Pixel size is 30cm or 60cm, depending on the source imagery.	1234.56
Largest	Identifies the largest structure for every address, indicated as 1 for largest structure or NULL	1

* Available with a 3D license only; 3D metrics are computed from processed LiDAR point-cloud data.

** LBCS codes and ll_uuid are derived from Regrid parcel data

*** Source may differ from where polygon is derived

To see the fields populated, try Buildings API. It returns the full point record for any address, coordinates, or region of interest.

8. How is it delivered, and what does it cost to license?

The data is available two ways: as a bulk delivery, or per lookup through the Buildings API. Bulk deliveries ship as standard GIS formats like ESRI Shapefiles and Geopackages. Depending on your project needs, we can deliver a joined version, only polygons with attributes (no points), repeating polygons for multiple addresses, etc.

The Buildings API returns the same records per lookup, queried by address, coordinate, or region of interest, with JSON responses. The response carries the point schema described in Section 7, so a single address lookup will show you the full attribute set on a building you already know, before you scope a bulk delivery.

Licensing: full commercial licensing, no share-alike obligations. The 3D license tier gates the 3D metrics; LBCS and parcel-identifier fields derive from Regrid parcel data. Our Buyer's Guide can help you understand what you can do with your license.

9. Need something the standard specification doesn't cover?

Section 5 describes the capture rules used for the standard, nationwide product. Those rules can be adapted for a specific project: custom capture criteria, delivery formats, or coverage scoped to what a project actually needs.

Get in touch to talk through what's possible for your project. Or try the API on your own data first.