

AI Packs

AI Packs control the content you have access to in the online experience of Nearmap AI. Adding AI Packs to your subscription will determine what AI Layers are made visible in MapBrowser, and what data is present in your online exports.





- Available Packs
 - Statistical Performance
 - Precision (Correctness)
 - Recall (Completeness)
 - Confidence
 - Minimum Object Sizes
 - AI Parcel Data Specifications
 - Gen 1 & 2 Spreadsheet Data
 - .csv file
 - Rich Geospatial Data
 - Point Objects
 - Polygon Objects

Nearmap AI customers can subscribe to any of the available AI Packs. The AI Packs control which Nearmap AI content you have access to. An AI Pack allows you to view specific AI Layer attributes and export the AI Parcel data, both from within MapBrowser.

Available Packs

Currently available AI Packs on our [latest generation](#) of content are listed below. The configuration below defines availability as at [Generation 4 AI Content](#).

AI Pack	Offline Delivery	AI Feature API	MapBrowser Visualisation (All Generations)	MapBrowser Export* (Gen 1 & 2 Data)
Building Footprints	✓	✓ <ul style="list-style-type: none"> • "Building" feature class (vector polygons) 	✓ Building	✓ Gen 1 / 2 - Building Footprint presence; area estimate and polygon for each building footprint.
Building Characteristics **	✓	✓ Metadata attached to Building Feature Class <ul style="list-style-type: none"> • Storeys (1, 2, 3+) with confidences • Peak Height 	ⓘ No applicable AI Layers	ⓘ Not available in Gen 1 / 2 Data
Roof Characteristics **	✓	✓ Metadata attached to Roof Feature Class <ul style="list-style-type: none"> • Roof Material Component (ratio and area of each, dominant material) • Roof Type Component (ratio and area of each) • Tree Overhang Component (ratio and area) • Dominant Roof Pitch (degrees) • Feature classes for each roof material, roof type, and tree overhang (vector polygons) 	Roof Materials: <ul style="list-style-type: none"> ✓ Tile ✓ Metal ✓ Shingle Roof Types: <ul style="list-style-type: none"> ✓ Hip ✓ Turret ✓ Flat ✓ Gable ✓ Dutch Gable ✓ Tree Overhang 	Attributes of each Building Footprint: Dominant Roof Material, presence of each roof shape. Tree Overhang presence; area estimate and polygon of each element of overhang.

Roof Condition***	✓	<ul style="list-style-type: none"> ✓ Metadata attached to Roof Feature Class <ul style="list-style-type: none"> • Roof Condition Component (ratio and area of each) • Feature classes for each roof condition (vector polygons) 	Roof Conditions: <ul style="list-style-type: none"> ✓ Structural Damage ✓ Temporary Repairs ✓ Roof Ponding ✓ Roof Rusting ✓ Tile or Shingle Staining 	 Not available in Gen 1 / 2 Data
Poles	✓	<ul style="list-style-type: none"> ✓ Metadata for power and light presence attached to pole centroids. 	<ul style="list-style-type: none"> ✓ Power Pole ✓ Light Pole 	 Not available in Gen 1 / 2 Data
Pool	✓	<ul style="list-style-type: none"> ✓ "Swimming Pool" feature class (vector polygon). 	<ul style="list-style-type: none"> ✓ Swimming Pool 	Swimming pool presence; area estimate and centroid of each pool.
Solar	✓	<ul style="list-style-type: none"> ✓ "Solar Panel" feature class (vector polygon). 	<ul style="list-style-type: none"> ✓ Solar Panel 	Solar panel presence; area estimate and centroid of each solar array.
Trampolines	✓	<ul style="list-style-type: none"> ✓ "Trampoline" feature class (vector polygon). 	<ul style="list-style-type: none"> ✓ Trampoline 	Trampoline presence; area estimate and centroid of each trampoline.
Construction	✓	<ul style="list-style-type: none"> ✓ Feature class for each included AI Layer 	<ul style="list-style-type: none"> ✓ Construction Site ✓ Construction Vehicle ✓ Construction Crane ✓ Building Under Construction 	Construction presence in a parcel.
Vegetation	✓	<ul style="list-style-type: none"> ✓ Feature class for each included AI Layer. 	<ul style="list-style-type: none"> ✓ Medium & High Vegetation (>2m) Low Vegetation (0.5 m-2m) Very Low Vegetation (<0.5m) 	 Not available in Gen 1 / 2 Data
Surfaces	✓	<ul style="list-style-type: none"> ✓ Feature class for each included AI Layer (vector polygons). 	<ul style="list-style-type: none"> ✓ Lawn Grass ✓ Water Body ✓ Asphalt ✓ Concrete Slab ✓ Natural (soft) 	 Not available in Gen 1 / 2 Data

*All data are based on WGS84 (EPSG : 4326)

**If you subscribe to Roof Characteristics or Building Characteristics, you must also subscribe to Building Footprints.

***If you subscribe to Roof Condition, you must also subscribe to Roof Characteristics and Building Footprints.



Statistical Performance

The [table below](#) is a summary of the statistical performance for each attribute with a "Y"/"N" flag in the spreadsheet output of an AI Parcel export. These numbers refer to whether at least one object of this type is present in an AI Parcel. The AI Packs also define what data will be delivered if you order a bulk offline delivery of data.

To ensure that our performance scores are as objective as possible, our examples are drawn from a statistically determined sample across our coverage regions that is weighted towards populated areas. We have deliberately chosen a significant portion of our examples as challenging cases where our models are least certain.

Our "source of truth" is a highly trained team of human expert labellers, who use a custom version of MapBrowser to check multiple dates, multiple angles, and even our 3D models to determine whether they believe an object is present. Our brief is "on the assessment date, mark what you judge to be present using all MapBrowser tools available to you". This means that a swimming pool missed on a leaf-on survey will be marked as incorrect, if the labeller can see the pool before or after that point in our capture history.

As a customer, ***you will notice three practical causes of error*** that stand out above all others:

- **Inconsistencies in third party parcel boundaries**, causing an object to be misidentified as belonging to a neighbouring property.
- **Definitional differences**, where your working definition is subtly different from ours (such as our current Building Footprint definition excluding rooftop carparks).
- What we like to term "**forgivable errors**" where, on clicking the MapBrowser link provided with every AI Parcel output, you may think to yourself "I appreciate why that would have happened". This is most noticeable for the Construction Site class, where the Precision is decreased by picking up things such as landscaper's yards full of rubble and trucks, and the Recall is decreased mostly by examples of the first stage of construction, which is usually just an area of dirt with no obvious construction occurring.

Our models do of course make other kinds of errors; however this is one of the primary reasons we have provided you with the ability to view AI Layers in MapBrowser. Nearmap AI gives you a unique perspective of the truth on the ground based on visual data, with different strengths and weaknesses to other sources (such as paper records of construction or solar installations). The consistent behaviour makes it an excellent source of truth as a standalone data product. If you have an existing data set representing a different perspective, we encourage you to combine it with ours to achieve a much more accurate picture than either perspective could alone.

Precision (Correctness)

We are correct approximately $< P \text{recision } \%$ of the time, in cases where we said "Yes". For example, in [Gen 2](#) we were correct 99% of the time in cases where we said "Yes, this is a swimming pool".

Precision is about whether the model accidentally flags parcels as "Yes". e.g. if we incorrectly flag a landscaper's yard as construction, it reduces the Precision.

Recall (Completeness)

We find approximately <Recall %> percent of the parcels which should actually say "Yes". For example in [Gen 2](#) we found approximately 94% of the parcels containing a swimming pool which actually do contain a swimming pool.

Recall is about whether the model accidentally misses parcels which should be picked up as "Yes". e.g. if we miss flagging a swimming pool because it is hidden completely under trees on the assessed imagery date, it reduces the Recall.

Precision and recall summary by attribute (Gen 2 content)

Class (object presence in a parcel, or on a building)	Precision	Recall
Building	>99.9%	>99.5%
Tile Roof	97%	95%
Shingle Roof	97%	97%
Metal Roof	91%	89%
Hip Roof	99%	94%
Gable Roof	98%	91%
Flat Roof	92%	91%
Turret Roof	93%	* 33%
Construction Site	75%	** 81%
Swimming Pool	99%	*** 94%
Solar Panel	97%	98%
Trampoline	98%	94%

* 2/3 of the examples of missed turrets in the test set were partial round turrets (see [AI Pack: Roof Characteristics](#)).

**The majority of missed construction is bare earth and slab down, before obvious signs of construction have commenced. The majority of false positive construction is "construction like" such as landscaper yards, or parcels with a lot of rubble and exposed dirt.

***The majority of missed swimming pools are very small above ground pools in a poor state of repair, or with a winter cover on.

Confidence

We help you make sense of the data and how you can use it by giving each attribute a Confidence score. This score is a representation of how much we believe the classification to be true.

Our confidence at detecting the different attributes varies, as some are more challenging than others.

In a technical sense, these represent a reasonably well-calibrated probability that the Y/N assignment is correct. For example, taking all the swimming pools with "99%" confidence listed, approximately 99% of them will actually be swimming pools.

This is a powerful tool for deciding whether you want to include all possible examples (e.g. including swimming pools above 70% confidence will capture empty or heavily shaded pools, but will include more false positives). In other cases, a more strict high threshold may be suitable (e.g. ignoring buildings below 90% confidence will remove many of the small garden sheds in shadow, but also remove the vast majority of false-positive roofs).

Minimum Object Sizes

For Gen3+ data, no minimum object size filters are applied. The smallest features tend to be the most likely to be false positives, but a customer filtering for a minimum "confidence" of e.g. 70% or 90% is a better filter for false positives than size unless there is a meaningful minimum size that applies to a use case. e.g. filtering out swimming pools less than 4 sqm will reduce the number of false positives, but also remove all small spas and paddling pools.

The table below defines the minimum attribute area for Gen 1 and Gen 2 data. Any object smaller than this minimum area is excluded from the AI data set. You may wish to impose additional minimum thresholds, for example to rule out garden sheds, or small paddling pools.

	m ²	ft ²
Solar Panel	3	32

Swimming Pool	4	43
Building Footprint	9	97
Construction Site	9	97
Trampoline	3	32
Tree Overhang	0	0

AI Parcel Data Specifications

Gen 1 & 2 Spreadsheet Data

This represents a summary of information, where each row is a property parcel, and columns are summarised facts about that parcel. It often contains less information than the geospatial data, but is very convenient to use if you do not have a GIS background. We often refer to it as a "parcel rollup", because it rolls up information about all objects detected in that AI Parcel to a single row in a database, a little like a geospatial version of a "GROUP BY" aggregation over parcels.

Most of the .csv files (comma separated value files) have common elements:

.csv file

Common elements of .csv files

This applies to Trampoline.csv, Swimming_Pool.csv, Solar_Panel.csv, Construction_Site.csv

Column	Description	Format
property_id	Unique parcel identifier from third party provider that best matches the parcel	GNAF (AU) / Parcel APN (US)
survey_date	Capture date of imagery used for the object detection	DD/MM /YYYY
address	Address data as supplied by third party that best matches the parcel NB: We still process parcels that have a blank address field from third party providers, as they often represent a legitimate property parcel.	
longitude	Longitude coordinate of parcel centroid (EPSG : 4326 WGS84)	
latitude	Latitude coordinate of parcel centroid (EPSG : 4326 WGS84)	
attribute	Name of the attribute represented by the file, e.g. "Trampoline"	
present	Yes/No value representing whether at least one example of the object has been detected "Y" or there has been insufficient evidence that any examples exist in the parcel "N".	Y/N
confidence	A percentage measure (represented as e.g. "100%" in the file) of the likelihood we believe the "present" decision is correct. "Y" cases are typically well-calibrated, such that the decision should be correct in 9/10 cases among all examples with a confidence of 90%. Note that "N" cases are typically set to 100% confidence for technical reasons (when there is no evidence at all of object presence, it is not possible to separate out different confidence levels meaningfully. 100% is approximately well-calibrated, given the true negative rates are typically in excess of 99%).	%

area_estimate_sqm OR area_estimate_sqft	An estimate of the total summed area of all instances of this type within the parcel, e.g. aggregate all sections of solar panel array. Units are selected automatically based on region, and reflected in the column name. Given the variation in areas due to occlusion from tree overhang etc., no statistical guarantees are made about the accuracy of the areas. Observing the AI Layers give a good visual indication of how well the area of each object is likely to be captured. Estimated area is in the horizontal plane and does not take slope and 3D structure into account.	area_estimate_sqm - m ² (AU) area_estimate_sqft - ft ² (US)
map_browser	Direct link to the parcel centroid at the specified survey date in MapBrowser. This can be used for examining why a particular decision may have been made.	
wkt	Well Known Text representation of parcel centroid for easy use within GIS software.	

Additional elements of .csv files

Some files have the above columns, plus additional ones representing further information. These include:

Building_And_Roof_Characteristic.csv

Additional Column	Description
building_confidence	Represents the likelihood that the "present" field is correct for "Roof", indicating the likelihood that there is at least one building present in the parcel.
dom_roof_material	Text field set to: "Tile", "Shingle", "Metal", "Other" or "NA" The Dominant Roof Material of the primary (largest) 'Roof/Building' within the parcel . "NA" is indicated when no buildings at all are detected in the parcel.
dom_roof_material_confidence	The percentage likelihood that the dominant roof material has been chosen correctly.
roof_hip	As per "present" field, applied to hip elements on the largest roof in the parcel.
roof_hip_confidence	As per "confidence" field, applied to roof_hip.
roof_gable	As per "present" field, applied to gable elements on the largest building in the parcel.
roof_gable_confidence	As per "confidence" field, applied to roof_gable.
roof_flat	As per "present" field, applied to flat elements on the largest roof in the parcel.
roof_flat_confidence	As per "confidence" field, applied to roof_flat.
roof_turret	As per "present" field, applied to turret elements on the largest roof in the parcel.
roof_turret_confidence	As per "confidence" field, applied to roof_turret.
tree_overhang	As per "present" field, applied to presence of non-zero area of tree overhang on the largest building in the parcel.
tree_overhang_confidence	As per "confidence" field, applied to tree_overhang.
tree_overhang_area_sqm OR tree_overhang_area_sqft	As per "area_estimate", except referring to the summed area of tree overhangs on the largest building in the parcel.

Rich Geospatial Data

The GeoPackage files provided contain all the information in the spreadsheet files, but have some key differences, and represent more fine-grained information about geospatial positions, areas and shapes of individual objects. It has been tested to open in QGIS, Esri's suite of ArcGIS products, and using the GeoPandas python package.

The geopackage data are all based on WGS84 (EPSG:4326).

Point Objects

- The point represents the centroid of the object rather than the centroid of the parcel.
- The area estimate represents the area estimate of that object (not the aggregated area of the object type within the parcel).
- If multiple objects are present in a parcel, multiple points per parcel may be returned, each with their own area estimates. Solar Panels, for example, typically return the centroid of each group of panels on a roof.

Polygon Objects

If only the "Residential Building Footprint" AI Pack is enabled, the file Building.gpkg will be present. The polygons represent every Building Footprint detected in each parcel.

If the "Residential Building Characteristics" AI Pack is also enabled, this file is replaced by Building_Footprint_With_Roof_Characteristic.gpkg. It contains the same Building Footprint polygons, with additional metadata for dominant roof material and roof shapes of each building. In that case, the file "Tree_Overhang.gpkg" will also be present, which contains polygons of all overhanging sections of tree detected.