

A fee-free GUGiK data chain: from the LiDAR point cloud to a list of changes in the field

QGIS manual and scripts · stage 2 of the AI implementation plan · tested on data from Ełk, 3-5 October 2026, revised after audits no. 1, 3, 4 and 5

This kit implements stage 2 of the plan for introducing AI in a land surveying practice. From free data published by the Head Office of Geodesy and Cartography (Główny Urząd Geodezji i Kartografii, GUGiK) - the LiDAR point cloud, the orthophotomap and buildings from the land and building register (ewidencja gruntów i budynków, EGiB) obtained from the aggregated service - QGIS produces a **map of differences**: a list of places where the situation on the ground most probably differs from the register. The practice uses it before going out to survey for a design base map (mapa do celów projektowych, MDCP), when updating the EGiB and when modernising the register. The tools and data involve no fees: QGIS, the plugins and the data are free, and the two scripts supplied with this manual compute everything locally, without sending any data to the cloud. The chain was tested on 3-4 October 2026 in QGIS 3.44.15 LTR on a LiDAR sheet from Ełk (2025, nominal density 12 pts/m²): the scripts' own computation from the point cloud to the CSV list takes about 10 seconds (excluding data download and work at the computer), and the result contains 123 places to check, including 49 objects that are not in the public EGiB. Script 1 was corrected after conformity audit no. 3 (4 October 2026): since then QGIS and the Łańcuch GUGiK (GUGiK Chain) portal give the same result on this sheet. **The result is a hint, not a survey**: no object from this layer goes directly into the survey documentation.

Quotations from Polish legislation are unofficial translations.

● What the chain produces and what it is for

The chain has five links:

1. **Data download** with the GUGiK plugins: LiDAR point cloud (LAZ), orthophotomap, EGiB buildings from the aggregated WFS service.
2. **Building outlines from LiDAR** (script 1): class 6 points (buildings) converted to polygons.
3. **Optionally, AI on the orthophotomap** (the Deepness plugin for the whole sheet, SamGeo for individual objects), when the LiDAR point cloud is older than the orthophotomap.
4. **Comparison with the EGiB** (script 2): four categories of differences, eave strip subtracted, CSV list with coordinates.
5. **Field check** with a tablet or phone (e.g. the free QField app) and entry of the result in the list.

The result is the layer "To check in the field" and a CSV file for Excel with Polish regional settings (semicolon separator, decimal comma). The category codes are fixed in the data and remain in Polish: NOWY (new), ROZBUDOWA (extension), BRAK (missing), ZMNIEJSZENIE (reduction). Categories:

Category	Meaning	Priority	Typical cause
NOWY	detected object with almost no EGiB coverage (below 20%)	1	building, garage, hall or open-sided shed not entered in the EGiB
ROZBUDOWA	part of the detected object extends beyond the EGiB outline	2	lean-to extension, added garage, canopy
BRAK	EGiB building that was not detected (below 20% coverage)	2	demolition, building hidden by trees, classification error
ZMNIEJSZENIE	part of the EGiB building was not detected	3	partial demolition, low roof outside class 6



Figure 1. EtK, 250 × 250 m extract. Blue: EGiB buildings (aggregated WFS service, as of 3 October 2026). Yellow: outlines from LiDAR 2025 (script 1). Red: NOWY. Orange: ROZBUDOWA. Green: ZMNIEJSZENIE. Background: GUGiK orthophotomap (WMS). Figure from the first trial (script 1 before the correction of 4 October 2026, outlines about 0,5 m wider).

● Legal limits in six points

1. **The data are free by operation of law.** Art. 40a(2)(1) of the Geodetic and Cartographic Law (Prawo geodezyjne i kartograficzne, PGiK) exempts from fees, among others, the digital terrain model (letter d), the orthophotomap (letter e), BDOT10k (letter h), parcel identifiers and geometry (letter i) and building identifiers, geometry and type (letter j). GUGiK makes the LAZ point clouds available free of charge in the open data module of the geoportal.
2. **The source must be stated.** Art. 40c(3) PGiK requires information on the origin of the material in published works, e.g. "Data: GUGiK, LiDAR 2025, EGiB (aggregated WFS service)".
3. **Data from the county geodetic and cartographic documentation centre (powiatowy ośrodek dokumentacji geodezyjnej i kartograficznej, PODGiK) only for the notified work.** Art. 12(8) PGiK: materials from the resource made available in connection with a notification of survey work (zgłoszenie) may be used solely for the work covered by that notification. The public data from step 3 can be used freely; files from the PODGiK (EGiB with attributes, BDOT500, GESUT) must not be accumulated in your own database for other commissions.
4. **This is not a survey.** The technical standards regulation (consolidated text, Journal of Laws (Dz.U.) 2022 item 1670) classifies buildings as group I terrain details (§ 6(1)(c)) and requires for them an

accuracy of 0,10 m relative to the control network (§ 16), while the choice of method and responsibility for accuracy rest with the survey work manager (§ 3). An outline from LiDAR is the roof with its eaves, determined with an accuracy of the order of several tens of centimetres. It serves to plan the survey, not to replace it.

5. **No personal data in the cloud.** The aggregated WFS service returns building geometry and identifiers, without owners. The scripts compute locally. Data from the PODGiK containing personal data are not entered into external AI tools (GDPR (RODO), see the main report).
6. **Responsibility stays with the surveyor.** The output of the chain is working material. If it was used to plan the survey, it is worth recording this in the technical report as an analysis of materials, and every change should be entered on the basis of a survey.

● What to prepare

Item	Version (as of 3 October 2026)	Cost	Notes
QGIS	3.44.15 LTR "Solothurn"	PLN 0	long-term release; the "Pobierz dane GUGiK" plugin works only in QGIS 3.x
Pobieracz danych GUGiK (GUGiK data downloader)	1.3.9 (23 September 2026), EnviroSolutions	PLN 0	LAZ, NMT, NMPT, orthophoto, BDOT10k, 3D models, EGIB
Pobierz dane GUGiK (Download GUGiK data)	1.2.12 (16 September 2026), GUGiK	PLN 0	parcels and buildings from the aggregated EGIB WFS service
NEW PROJECT scripts	np_1 and np_2 (folder "skrypty")	PLN 0	tested in QGIS 3.44.15; withheld-point filter in np_1 (6 October 2026) not yet re-run in QGIS
Deepness (optional)	0.6.5, Poznań University of Technology (Politechnika Poznańska), Apache 2.0 licence	PLN 0	ONNX models, including buildings and roads; on QGIS 3.44 requires a workaround (step 4b)
SamGeo (optional)	1.2.0 (stable), Qiusheng Wu	PLN 0	Segment Anything; in the plugin, prompting works only with SAM 3 (step 4b)
QField (optional)	app for Android and iOS	PLN 0	field work on the results layer
Computer	16 GB RAM, SSD	already owned	a LAZ sheet is about 60-90 MB; a graphics card is needed only for SamGeo (test: 1,7 GB of graphics memory)

● Step 1. Installing QGIS and the plugins

1. Download the **QGIS 3.44.15 LTR** installer from qgis.org (Windows, "Long Term Release", about 580 MB) and install it with the default settings.
2. In QGIS: **Plugins > Manage and Install Plugins...**, search for and install **Pobieracz danych GUGiK** (GUGiK data downloader) and **Pobierz dane GUGiK** (Download GUGiK data).
3. Copy the **skrypty** (scripts) folder from this kit to your disk (e.g. to the practice's documents).
4. Open **Processing > Toolbox** (Ctrl+Alt+T; the panel is titled "Processing Toolbox"), click the Python icon at the top of the panel and choose **Add Script to Toolbox....** Select the files `np_1_budynki_z_lidar.py` and `np_2_porownanie_z_egib.py` in turn. The group **Scripts > NEW PROJECT - GUGiK chain** with two algorithms will appear in the panel. The scripts are bilingual: when QGIS runs in Polish, the names, parameters, messages and CSV headers are in Polish (group "NEW PROJECT - łańcuch GUGiK"); in any other language they are in English, while the category codes NOWY, ROZBUDOWA, BRAK and ZMNIEJSZENIE are the same in both versions.
5. Optionally install the **Deepness** and **SamGeo** plugins (step 4b) and the **QField** app on a phone or tablet.

QGIS 4.x is already available (4.2.3 of 25 September 2026), but the GUGiK plugin for the aggregated EGIB service declares compatibility only up to 3.99. The scripts were written to work in QGIS 4 as well, but they

have been tested only in 3.44.

● Step 2. Project and coordinate reference system

Create a new project and set its CRS (bottom right corner of the window) to **EPSG:2180** (PL-1992), in which GUGiK publishes national data. For field work the result can be transformed straight away to the PL-2000 system of the appropriate zone by specifying it in script 2: EPSG:2176 (zone 5), 2177 (6), 2178 (7, e.g. EtK) or 2179 (8). Save the project in the job folder and keep all downloaded and output files next to it.

● Step 3. Downloading the data

Data	Plugin	Format	What to watch out for
LiDAR point cloud	Pobieracz danych GUGiK	LAZ, sheets of about 540 × 600 m	choose the most recent year; for EtK the 2023 and 2025 flights are available (nominal 12 pts/m ² , actual average about 39 pts/m ² of all returns)
Orthophotomap	Pobieracz danych GUGiK or the geoportal WMS	GeoTIFF / WMS	check the flight date; it determines what is not yet visible
EGiB buildings	Pobierz dane GUGiK (buildings)	vector layer	download an area slightly larger than the LAZ sheet
NMT, NMPT, 3D LoD models	Pobieracz danych GUGiK	ASC / GML	auxiliary, e.g. for assessing heights and storeys

Order of work: in Pobieracz select LIDAR measurement data, indicate a point on the map or a layer with the boundary of the job area, and download the sheet or sheets. Then, in "Pobierz dane GUGiK", download the buildings for the same area (the plugin lets you define the area from the map, from a layer, by TERYT identifier or by a point). The buildings layer from the aggregated service has the field **ID_BUDYNKU** (e.g.

280501_1.0002.653_BUD), which script 2 copies into the list.

● Step 4. Building outlines from LiDAR (script 1)

Add the downloaded sheet to the project: **Layer > Add Layer > Add Point Cloud Layer...** and select the LAZ file. When the file is opened for the first time, QGIS creates a COPC index next to it (a `.copc.laz` file); this happens once for each sheet.

Run **1. Building footprints from GUGiK LiDAR (class 6)**. GUGiK point clouds are classified according to the ASPRS standard: class 2 is ground, 3-5 are low, medium and high vegetation, and **class 6 is buildings and structures**. The script selects class 6 points, converts them to a helper raster (each point goes only to the cell in which it lies; for this the script calls the PDAL program installed with QGIS), closes small gaps with a positive and negative buffer, removes holes smaller than 4 m², simplifies stepped edges and discards objects smaller than the specified area.

Parameter	Default	When to change
Building class (APRS)	6	only for point clouds from sources other than GUGiK with a different classification
Helper raster resolution [m]	0,5 m	1,0 m for point clouds of 4-6 pts/m ² , 0,3 m for 20+ pts/m ² (with a value other than 0,5 m check items at the sheet edge manually, because the margin of script 2 is 0,51 m)
Gap closing (buffer + and -) [m]	0,5 m	more, when roofs break up into pieces
Minimum building area [m ²]	10 m ²	6 m ² , when small open-sided sheds and outbuildings are of interest

Parameter	Default	When to change
Edge simplification [m]	0,3 m	0, when the outline is to be used for visual checking of the shape

Test result: from the Etk 2025 sheet (12,9 million points) the script created 171 outlines in about 7 seconds, with the correct EPSG:2180 CRS. An outline from LiDAR is systematically larger than the outline in the EGiB because it includes the roof eaves. Script 2 takes this into account.

The output of script 1 ("Buildings from LiDAR") is by default a temporary layer, which disappears when QGIS is closed. Save it: before running, specify a GeoPackage file in the output field, or after running, choose **Export > Save Features As...** on the layer.

● Step 4b. Optionally: AI on the orthophotomap

When the LiDAR point cloud is several years older than the orthophotomap, or when you are interested in objects not visible in class 6, detection from imagery can be added. Both approaches were checked on 4 October 2026 on the same sheet from Etk and the GUGiK orthophotomap from the 2025 flight. Because the point cloud and the orthophotomap are from the same year, the test measures the method itself, not the difference in dates.

Deepness (Poznań University of Technology) runs ONNX models from the plugin's model zoo in QGIS: building segmentation, land cover (building, forest, water, road), roads, tree tops and cars. Workflow:

1. Install the Deepness plugin. On first run the plugin offers to install Python packages; accept.
2. **Workaround for QGIS 3.44.** The installer of plugin 0.6.5 downloads OpenCV 4.9 and onnxruntime 1.17, built for NumPy 1.x, whereas QGIS 3.44 has NumPy 2. The plugin then fails with the error "numpy.core.multiarray failed to import". This is a known bug (issue no. 235 with the authors, open). Close QGIS, open the **OSGeo4W Shell** from the Start menu and enter a single command (all on one line): `python -m pip install -U --target="%APPDATA%\QGIS\QGIS3\profiles\default\python\plugins\deepness\python3.12" opencv-python-headless onnxruntime` In the test, OpenCV 5.0 and onnxruntime 1.30 were installed, after which the plugin worked.
3. Download **Buildings Segmentation** from the model zoo (ONNX file, about 80 MB) and select it in the plugin. The plugin takes the resolution, tile size and threshold from the file metadata. Note: the model zoo states 40 cm/px, while the file metadata state 50 cm/px; the value from the file applies.
4. As the input layer, select the orthophotomap in EPSG:2180 and run the segmentation. The result is a temporary layer "Building". Save it to GeoPackage, split it into single polygons (**Multipart to singleparts**) and discard small fragments with the expression `$area >= 10` (**Extract by expression**).
5. Run script 2 as for LiDAR. Outlines from Deepness are sometimes geometrically invalid; script 2 (version of 4 October 2026) repairs them itself, takes the polygon parts from geometry collections created during the repair and reports in the log the number of repaired and rejected objects (in the test 3 repaired, 0 rejected).

Deepness test result: 36 tiles in 169 s on the CPU alone, 162 outlines after filtering, a list of 107 items (NOWY 24, ROZBUDOWA 36, BRAK 19, ZMNIEJSZENIE 28). On this sheet Deepness agrees less with EGiB and detects fewer EGiB buildings. This compares the two methods with each other and with EGiB, not with a field survey:

Measure (Etk sheet, 268 EGiB buildings)	LiDAR class 6	Deepness
EGiB buildings detected in at least half	254	235
Of which small, below 50 m ² (105 buildings)	101	87
BRAK items on the list	12	19, of which 15 are in fact seen by LiDAR
Outline area confirmed by the second source	-	about 93%

LiDAR column after the correction of script 1 (4 October 2026): detections of EGİB buildings unchanged (254 and 101), 12 BRAK items instead of 10; the Deepness area confirmed by LiDAR fell by about 1 percentage point, because LiDAR outlines are no longer artificially widened.

These figures describe the agreement between two methods on one sheet, not the accuracy of detecting changes on the ground: the second source can also be wrong (LiDAR classified a viaduct as a building, Figure 3). The accuracy of each category (how many indications are true and how many changes are missed) can only be assessed on an independently verified sample, preferably in the field.

Conclusion: Deepness is suitable as a **second source** for an older point cloud or for a quick review of a cadastral district (obręb). On this sheet most of its BRAK items disagree with LiDAR (mostly small and dark roofs that LiDAR sees), so treat BRAK from Deepness with caution and NOWY as a hint to check. Stronger statements about the accuracy of either method will be possible only after the field check (sample Z4).

SamGeo (Segment Anything, the SAM model by Meta) extracts a single object indicated by a box or a point. It is useful for individual objects, e.g. a new open-sided shed on the client's plot. The test revealed three limitations of the plugin itself:

- **Prompting works only with SAM 3.** Plugin 1.2.0 (stable in the QGIS plugin repository; the experimental 1.4.0 is also there, and the latest 1.4.2 only on GitHub) creates SAM 2 and SAM 1 models in automatic mode. When a point or a box is indicated, this ends in an error ("object has no attribute predictor"), and saving the results is also tailored to SAM 3 only. SAM 3 requires a free Hugging Face account and Meta's consent to download the model; this path was not tested.
- **Installation is heavy.** SamGeo needs PyTorch: with graphics card support this is about 5 GB of packages plus 0,9 GB for the SAM 2 model. The author recommends a conda environment with QGIS; in the test the packages were installed into the QGIS Python (OSGeo4W) in a separate directory.
- **The image must be a file.** The plugin does not read a WMS layer, only a GeoTIFF from disk, and passes the entire image to the model. Instead of the whole sheet, use a clip around the object (in the test 40-140 m with a 10 cm pixel), because the model reduces the image to 1024 pixels anyway.

The SAM 2 model itself, called directly from the samgeo library (bypassing the plugin), was checked on 12 objects: the 6 largest NOWY items from the LiDAR list and 6 EGİB buildings from 81 to 4 744 m², with the LiDAR outline as the reference. Prompting with a **box** gave agreement with LiDAR of at least 0,74 (ratio of intersection to union) for 8 of the 12 objects, median 0,80. Prompting with a **point** is unreliable: median 0,44. Time on an RTX 3050 card: about 0,4 s per object. This is a demonstration on 12 selected objects, not an assessment of effectiveness for other areas, point densities or types of development.

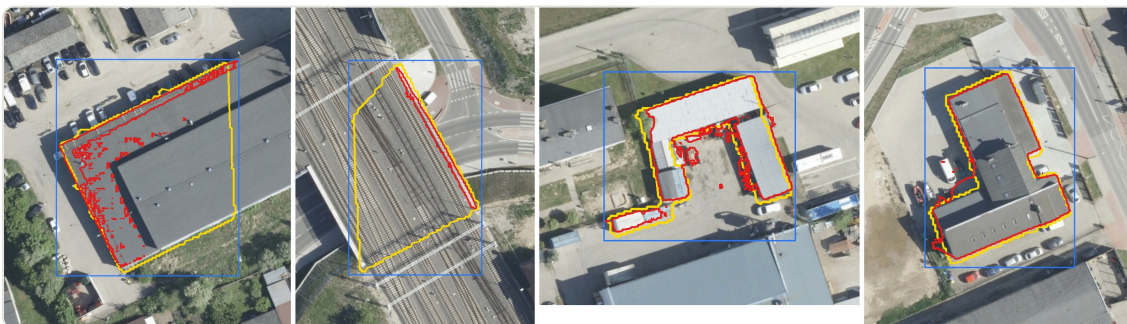


Figure 3. SamGeo (SAM 2, box) on a 10 cm orthophotomap. Yellow: outline from LiDAR. Blue: the indicated box. Red: SAM 2 result. From the left: a hall with a dark roof, whose roof on the orthophotomap is displaced relative to LiDAR (relief displacement) and which the model split into fragments; a LiDAR "building" that is a viaduct over railway tracks (class 6 also includes structures), so SAM rightly found nothing; a group of farm buildings and an EGİB building, where both outlines almost coincide.

Outlines from the orthophotomap have an additional source of error: on a standard orthophotomap the roof of a tall building is displaced relative to its ground floor (relief displacement), so differences of the order of a metre for tall buildings are normal.

● Step 5. Comparison with the EGiB (script 2)

Run 2. Comparison with EGiB - map of differences for field check. Input A is the outlines from step 4 (or 4b), input B is the EGiB buildings.

Parameter	Default	Meaning
Field with EGiB building identifier	ID_BUDYNKU	goes to the "EGiB id" column
Output coordinate reference system (projected, metres)	EPSG:2180	specify the PL-2000 zone if the list goes to the field
Analysis area (e.g. LAZ sheet extent); empty = extent of detected footprints	extent of detected footprints	set the extent of the LAZ sheet ("Calculate from Layer" button); only objects lying entirely inside the sheet (with a 0,51 m margin) are then listed
Eave strip and edge tolerance [m]	1,5 m	strip along the other layer ignored in the differences (eaves, stepped edge); the part outside the strip must be at least 1 m wide
Minimum change area [m ²]	15 m ²	smaller differences are not listed
NOWY when covered by EGiB less than [%]	20%	an object covered by the EGiB in less than 20% is NOWY
BRAC when less than [%] of EGiB building detected	20%	an EGiB building detected in less than 20% is BRAC
EGiB snapshot (TSV file)	temporary file	specify a file next to the CSV list; it stores the EGiB buildings used in the comparison

The script first checks each detected object against the EGiB (NOWY, ROZBUDOWA), then each EGiB building against the detected objects (BRAC, ZMNIEJSZENIE). The eave strip (1,5 m by default) along the other layer is ignored. The part extending beyond the strip must be at least 1 m wide (0,5 m opening), and it is then restored to its full depth, together with the fragment lying within the eave strip. As a result, the rim around roofs disappears, even where the roof overhangs by about 2 m, and extensions from about 2,5 m deep, eaves included, are listed. Finally, a light 0,5 m opening removes slivers narrower than 1 m, and parts with an area of at least 15 m² remain. The same rule works in both directions (ROZBUDOWA and ZMNIEJSZENIE) and in the "Łańcuch GUGiK" (GUGiK chain) portal. The list is sorted by priority and change area.

The analysis area determines only what is listed: objects cut by the area boundary are not listed, but they count in the comparison. An EGiB building right at the edge of the sheet is therefore not reported as BRAC merely because its LiDAR outline extends beyond the sheet.

Script 2 has three outputs. The layer "To check in the field" is temporary by default: save it to GeoPackage (specify a file in the output field or use **Export > Save Features As...**), otherwise it will disappear when QGIS is closed. It has the empty fields `sprawdzono` (checked) and `uwagi` (notes), to be filled in in the field, e.g. in QField (layer field names are the same in both language versions); the EGiB identifiers are stored in full, and an overflow is indicated by the suffix "(+n)". The CSV list is created only if you specify a file in the "CSV list for field work" field. In the log the script reports the number of objects at the area boundary and the number of repaired geometries.

The third output is the EGiB snapshot: a text file with the EGiB buildings reaching the analysis area (identifier, type, WKT geometry to 1 cm, one building per line). In the log the script reports its SHA-256 hash; the saved file has the same hash. The WFS service always returns the current state, so without a snapshot the trial cannot later be repeated on the same data. Save the snapshot next to the CSV list and keep it with the result. The Łańcuch GUGiK portal saves the same snapshot with every analysis (from version 1.5): for the same area and the same EGiB data the file and the hash are identical in QGIS and in the portal. In the trial of 5 October 2026 script 2 gave the hash 80dc4b3e...e435 for the Elk sheet, equal to the portal hash, both from a layer downloaded from WFS for an area 100 m larger than the sheet and from the saved snapshot.

The choice of eave strip still affects the result, but no longer so strongly. In the test, a 1 m strip gave 60 ROZBUDOWA items, 1,5 m gave 56 and 2 m gave 55. The earlier rule (opening by the full strip width) lost extensions narrower than 3 m; the current one reports them, and the ROZBUDOWA items correspond on the

orthophotomap to lean-to extensions, garages, canopies and places where the EGiB outline is clearly smaller than the building.

● Step 6. Desk review and field check

1. **Desk review (15-30 minutes per sheet).** Against the orthophotomap background, review the items starting from priority 1. Mark obviously wrong items (e.g. a lorry classified as a building, a container) in the `sprawdzono` (checked) field as "rejected at desk". Remember that the orthophotomap may be older than the point cloud.
2. **Restricting to the job.** For a design base map, keep the items within the boundaries of the survey area plus a surrounding strip; for EGiB modernisation, whole cadastral districts.
3. **Field.** Transfer the results layer and the orthophotomap to a tablet (e.g. QField via the QFieldSync plugin or as a GeoPackage file). The CSV list has the columns "X north" and "Y east" in the surveying convention and the empty columns "checked in the field" and "notes".
4. **Survey.** Each confirmed change is surveyed with a method ensuring the accuracy required by the standard, and the results are submitted to the PODGiK under the normal notification procedure.

● The Łańcuch GUGiK portal

The same chain runs without QGIS in the portal <https://lancuch-gugik.newproject360.pl/> (accounts are created by the administrator). The portal user guide is on its Help page: searching for a parcel or address, marking the area, the analysis queue, reviewing items (statuses and notes with change history), PDF, CSV, DXF, GeoJSON (PL-1992 and WGS84), GeoPackage and EGiB snapshot export. Audit reports of the portal are on the Audits page, after signing in. The portal can be installed as an app on a phone and a computer (the App page in the footer). This manual, both scripts, the data of the Elk example and the study on AI in a surveying practice can be downloaded without signing in from the Materials page: <https://lancuch-gugik.newproject360.pl/materialy>

Differences between the portal and the QGIS scripts (the comparison rule and parameters are the same):

- **Comparison area.** The portal compares only within the part of the selection covered by the point cloud (also for several sheets with non-rectangular coverage); script 2 uses the "Analysis area" rectangle. With several sheets and gaps in coverage, EGiB buildings in the gaps give false BRAK in QGIS.
- **Grid origin.** The portal starts the grid at the corner of the selection and merges several sheets into one grid; script 1 starts at the corner of the point cloud of one layer. Results are identical when the selection equals the sheet extent (as in the test).
- **Edge margin.** In the portal it equals the grid resolution + 0,01 m; in script 2 it is a fixed 0,51 m (for a 0,5 m grid). With a 1,0 m grid, check items at the sheet edge manually.
- **Warnings.** The portal stops the analysis when the EGiB service returns no buildings also for an area 300 m wider or sends a building without a readable geometry. It warns about a point cloud without class 6 in the whole sheet and, separately, about no class 6 in the selected area where EGiB has buildings. BRAK and ZMNIEJSZENIE items where the cloud has few points (less than half of the grid cells with any point) are marked "sparse data": LiDAR cannot resolve them. On the Elk sheet 9 of 12 BRAK items and 4 of 6 ZMNIEJSZENIE are marked this way; they lie at the edge of the sheet, where the cloud has no points (the index extent can be wider than the actual points). The portal also stops the analysis when an EGiB building lies far outside the area (e.g. swapped axes), and reports EGiB buildings without area and withheld points with separate warnings. Warnings, marks, parameters, version and source hashes are in all exports (GeoJSON: the "metadata" member, GeoPackage: the "metadata" table, CSV: the "assessment" and "analysis warnings" columns, DXF: a line in the title block). The QGIS scripts do not do this.
- **Categories in English.** The English portal shows NEW, EXTENSION, MISSING, REDUCTION; the QGIS script keeps the Polish codes.
- **List of EGiB identifiers.** The portal fits up to 200 characters, script 2 up to 100; both mark the rest as "(+n)".

● Step 7. The same on PODGiK data in notified work

After notifying the work and receiving the materials (EGiB and BDOT500 in GML), the same script 2 can be run on the buildings layer from the PODGiK instead of the public service. This gives a comparison with the complete, up-to-date database of the centre. In accordance with Art. 12(8) PGiK, the result and the input data relate only to that work: after its completion they are not carried over to other commissions. Files from the EGiB containing personal data stay on the practice's computer.

● Test results: Elk, sheet N-34-81-D-a-2-2-2-1

Indicator	Value
Input data	LiDAR 2025, 12,9 million points, nominal 12 pts/m ² (actual average about 39 pts/m ² of all returns), sheet about 540 × 600 m; 296 EGiB buildings from the aggregated service (snapshot of 4 October 2026; 294 in the first trial)
At the sheet edge (not listed, but in the comparison)	28 EGiB buildings, 1 LiDAR outline (script 2 counter; the portal shows 26 buildings, because it counts only objects reaching into the area)
Outlines from LiDAR	171 (time about 7 s)
Comparison	170 outlines with 268 EGiB buildings (time about 3 s)
Result	123 items: NOWY 49, ROZBUDOWA 56, BRAK 12, ZMNIEJSZENIE 6
Hardware	ordinary office computer with Windows 11, no graphics card

A review of selected items on the orthophotomap showed that the NOWY category covers halls and open-sided sheds in the industrial zone, and garages and farm buildings in the single-family housing area. The BRAK item with an area of 724 m² (Figure 2) is an EGiB building in whose place the orthophotomap shows a car park. Without a field inspection it is impossible to determine whether this is a demolition or a register error. The differences relate to the public WFS service and may partly result from its completeness, not solely from the situation on the ground. This is precisely why the result is a list to be checked.

The Łańcuch GUGiK portal computes with the same method and gives the same result on this sheet: 172 outlines (171 in QGIS, a difference of one object with an area at the 10 m² threshold) and the same 123 items, each with a counterpart of the same category, median shape agreement (IoU) 1,0. Until 4 October 2026 script 1 used the QGIS algorithm "pdal:exportraster", which sets every cell within 0,71 m of a point; outlines were therefore on average about 0,5 m wider, and the first trial gave 161 outlines and 124 items (NOWY 43, ROZBUDOWA 63, BRAK 10, ZMNIEJSZENIE 8). Conformity audit no. 3 established this cause, and the corrected script 1 was checked in QGIS 3.44.15 on the same sheet.

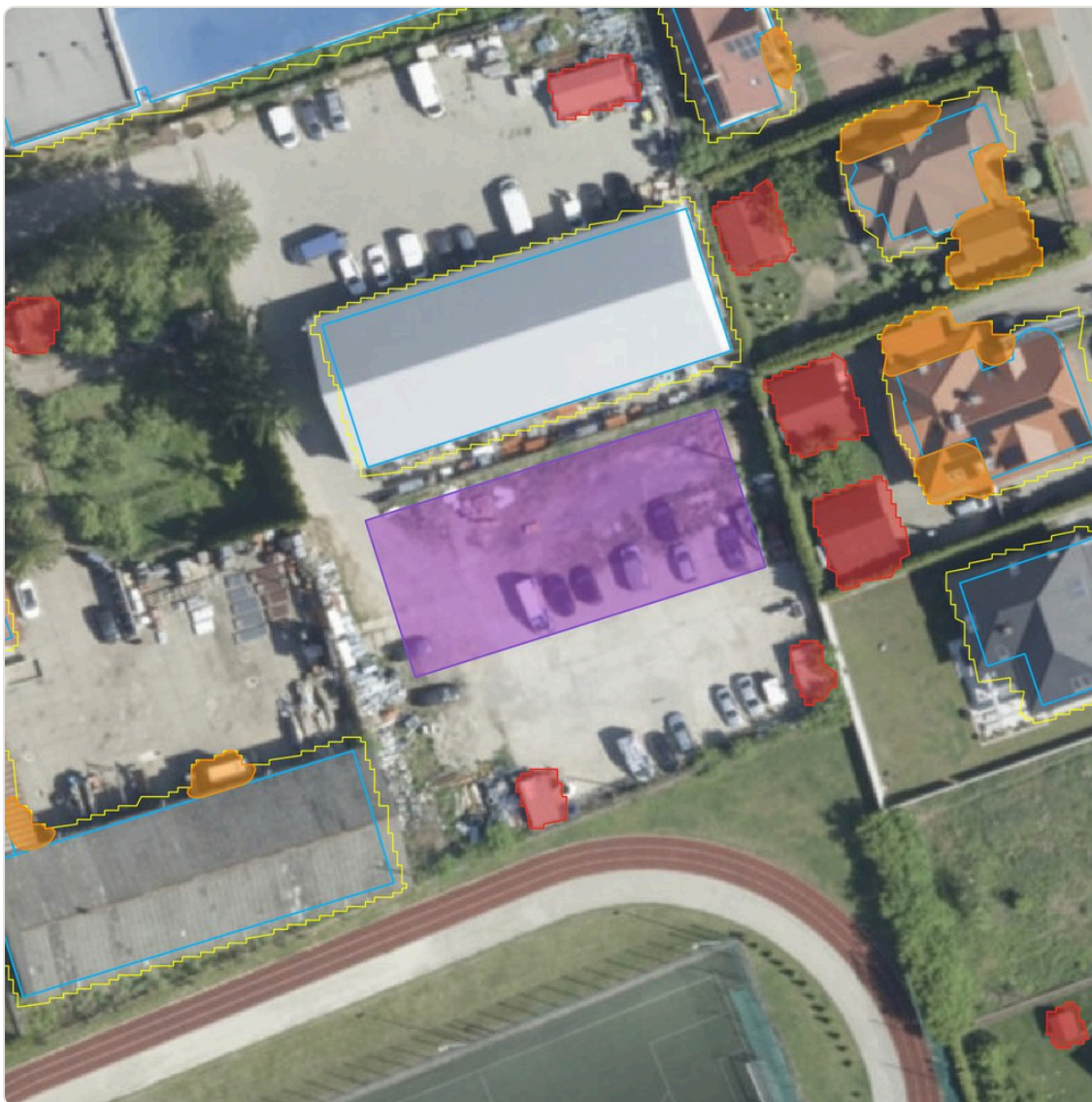


Figure 2. EfK, 120 × 120 m extract. Purple: BRAK, an EGiB building (724 m²) in whose place the orthophotomap shows a car park. Red: NOWY (garages, farm buildings). Orange: ROZBUDOWA at single-family houses. Figure from the first trial (script 1 before the correction of 4 October 2026, outlines about 0,5 m wider).

● Pitfalls and limitations

- **Data dates.** The point cloud, the orthophotomap and the EGiB come from different years. An object visible in LiDAR 2025 may not be visible on an older orthophotomap and vice versa. Always record the year of each dataset.
- **Eaves and relief displacement.** An outline from LiDAR includes the eaves; an outline from the orthophotomap additionally includes the relief displacement of the roof. The outline in the EGiB is the projection of the ground floor. Differences of up to about 1,5 m at the walls are normal.
- **Trees and low roofs.** A building under tree crowns may be classified as vegetation (false BRAK), and very low open-sided sheds as ground. Sometimes class 6 also includes objects that are not buildings (e.g. structures, containers).
- **Adjoining buildings.** Dense development gives in LiDAR a single mass covering several EGiB buildings. The script computes coverage for each building separately, but the boundaries between buildings on the ground must be established by survey.
- **Sheet edge.** Objects cut by the analysis area boundary are not listed, but they count in the comparison. Without the "Analysis area" parameter the script does not know where the sheet ends: an outline cut off by the sheet edge is listed as a whole object, and EGiB buildings partly outside the point cloud may give false BRAK and ZMNIEJSZENIE. Always set the extent of the sheet (or of several sheets).

- **A difference between datasets is not a change over time.** The result compares the cloud from the flight date with EGiB from the download date. A building completed after the flight gives BRAK, and a building not entered in EGiB gives NOWY, even though nothing has changed on the ground.
- **A cloud without building classification.** When the cloud has no class 6 points, every EGiB building comes out as BRAK. The portal then warns about limited reliability; in QGIS, check the number of class 6 points before the comparison. Class 6 present somewhere in the sheet does not prove that the cloud is classified at the selected place, and no class 6 in the area need not mean a classification error (the building may not have existed on the flight date): the portal therefore reports it as an uncertainty to check, not as a failure.
- **Withheld points.** The LAS specification allows points to be flagged for exclusion from normal processing. The portal (from version 1.7) and script 1 do not use them; the Elk sheet has no such points.
- **Detection thresholds.** The 0,5 m grid, the 1,5 m eaves band and the 15 m² minimum change are filters: narrow extensions (narrower than about 1 m beyond the eaves) and small changes are deliberately left off the list. No item on the map does not mean no change.
- **Coordinate system.** In QGIS "x" is east, whereas in surveying "X" is north. The CSV list uses the surveying convention and column names that make this unambiguous.
- **Test environment.** The tests were carried out in QGIS 3.44.15 unpacked without a system installation. In this setup the GDAL helper scripts had to be added manually; the standard installer does this itself. Script 1 uses PDAL and GDAL algorithms (including the raster vectorisation gdal:polygonize), so it needs a complete QGIS installation. Both scripts were run with the Polish and the English interface; the output data were identical. The Deepness and SamGeo plugins were run with the same calls that their code uses, but without the QGIS window; the plugins' interface itself (panels, buttons) was not clicked through.
- **AI plugins change.** The bugs described apply to Deepness 0.6.5 and SamGeo 1.2.0 (as of 4 October 2026). Before deployment, check whether newer versions have removed them.

● Time and costs

Item	Time	Fees
One-off: installation of QGIS, plugins, scripts	1-2 hours	PLN 0
One-off: learning on your own example	3-4 hours	PLN 0
Per job: data download (1-4 sheets)	10-15 minutes	PLN 0
Per job: scripts 1 and 2	under 1 minute per sheet	PLN 0
Per job: desk review	15-30 minutes per sheet	PLN 0

The "Fees" column covers licences and data. It does not include working time or hardware, nor, for the online portal, the server, backups and updates.

The gain comes not from the price of a service but from a better-planned field trip: before leaving, the surveyor knows where, within the area of the design base map, there are objects not shown in the register, and does not have to return to the field after the first verification. In EGiB modernisation, the list structures the field inspection for the whole cadastral district. The metric from the implementation plan: hours of office work and the number of additional field trips per job, measured before and after.

● Contents of the kit

File	Contents
GUGiK chain - manual.pdf	this manual (English version)
Lancuch GUGiK - instrukcja.pdf	the same manual in Polish
skrypty/np_1_budynki_z_lidar.py	script 1: building outlines from LiDAR (class 6)

File	Contents
skrypty/np_2_porownanie_z_egib.py	script 2: comparison with the EGIB, layer and CSV list
przyklad Elk/budynki_lidar.gpkg	output of script 1 for the test sheet
przyklad Elk/do_sprawdzenia.gpkg and .csv	output of script 2 (123 items)
przyklad Elk/egib-migawka.tsv	EGiB snapshot from the trial (296 buildings, SHA-256 80dc4b3e...e435)
przyklad Elk/deepness_budynki.gpkg	building outlines from Deepness (2025 orthophotomap)
przyklad Elk/deepness_do_sprawdzenia.gpkg and .csv	output of script 2 for the Deepness outlines (107 items)
przyklad Elk/rys1.jpg, rys2.jpg, rys3.jpg	figures from this manual
Z4 proba terenowa/	sample of 78 objects for field checking (audit no. 3, recommendation Z4): PDF cards, GeoPackage, results workbook

● Glossary

Term	Meaning
ASPRS	American Society for Photogrammetry and Remote Sensing; the American society whose point class standard (2 ground, 3-5 vegetation, 6 buildings) is used by GUGiK.
BDOT10k	Baza danych obiektów topograficznych 1:10 000 (topographic objects database at the level of detail of a 1:10 000 map); maintained by the voivodeship marshals (Art. 7c(1)(3) PGiK), while the Surveyor General of Poland (Główny Geodeta Kraju, assisted by GUGiK) maintains its integrated copy (Art. 7a(1)(14)(b) PGiK).
BDOT500	Baza danych obiektów topograficznych (topographic objects database) at the level of detail of the basic map (mapa zasadnicza), maintained by the county head (starosta).
conda	Python environment manager; the SamGeo author recommends installing QGIS together with SamGeo in it.
CSV	Text file with data in columns separated by a delimiter (here a semicolon), opened in Excel.
EGiB	Ewidencja gruntów i budynków (land and building register, the cadastre), maintained by the county head (starosta).
EPSG:2180	Code of the PL-1992 coordinate system; EPSG:2176-2179 are zones 5-8 of the PL-2000 system.
EVRF2007	European Vertical Reference Frame 2007, in which GUGiK publishes LiDAR data.
Footprint agreement	Area of the intersection of two outlines divided by the area of their union; 1 means identical outlines, 0 disjoint ones.
GeoPackage (GPKG)	Open file format for spatial layers, read by QGIS and QField.
GeoTIFF	Image file (TIFF) with stored position and coordinate reference system.
GESUT	Geodezyjna ewidencja sieci uzbrojenia terenu (geodetic register of utility networks).
GML	Geography Markup Language; the format for exchanging EGiB, BDOT500 and GESUT data with the documentation centre.
GUGiK	Główny Urząd Geodezji i Kartografii (Head Office of Geodesy and Cartography).
Hugging Face	Service providing AI models; some models (e.g. SAM 3) can be downloaded only after creating an account and obtaining the author's consent.
LAS	Open format for storing point clouds developed by ASPRS; class 6 in it is Building.
IoU	Intersection over Union: the area of the intersection of two outlines divided by the area of their union; 1 means an identical shape.
LAZ	Losslessly compressed LAS point cloud format.
PDAL	Point Data Abstraction Library: a point cloud library and program installed with QGIS; script 1 uses it to create the helper raster.
LiDAR	Airborne laser scanning; the result is a point cloud with height and class.

Term	Meaning
LoD	Level of Detail; the level of detail of a 3D building model (LoD1 block, LoD2 with roof).
LTR	Long Term Release; a QGIS version with long-term support, recommended for production work.
Morphological opening	Negative buffer followed by a positive buffer of the same width; removes narrow strips, preserves solid shapes.
NMT, NMPT	Numeryczny model terenu (digital terrain model, DTM: ground) and numeryczny model pokrycia terenu (digital surface model, DSM: with buildings and vegetation).
NumPy	Python library for array computations; QGIS 3.44 has version 2, with which older packages do not work.
ONNX	Open Neural Network Exchange; open format for neural network models used by Deepness.
onnxruntime, OpenCV	Python libraries that Deepness uses to run the model and process images.
OSGeo4W Shell	Command line installed together with QGIS on Windows, with the QGIS Python.
PGiK	Prawo geodezyjne i kartograficzne (Geodetic and Cartographic Law), consolidated text Journal of Laws (Dz.U.) 2024 item 1151, as amended.
PODGiK	Powiatowy ośrodek dokumentacji geodezyjnej i kartograficznej (county geodetic and cartographic documentation centre).
PyTorch	Library for running neural networks required by SamGeo; the version for NVIDIA graphics cards takes up several GB.
QField	Free mobile app for field work on QGIS projects.
QFieldSync	QGIS plugin that prepares the project and data for work in QField and transfers changes from the field back to the project.
Relief displacement	Displacement of the roof of a tall object relative to its base on a standard orthophotomap (Polish: przechył).
RODO	Rozporządzenie o ochronie danych osobowych (General Data Protection Regulation, GDPR): Regulation (EU) 2016/679 on the protection of personal data.
SAM	Segment Anything Model; Meta's AI model for extracting objects from images (versions 1, 2 and 3).
SHA-256	Cryptographic hash of a file or text; the same hash means the same content.
TERYT	Krajowy rejestr urzędowy podziału terytorialnego kraju (national official register of the territorial division of the country); among other things, identifiers of voivodeships, counties and municipalities.
TSV	Text file with columns separated by a tab.
WKT	Well-Known Text; text notation of a geometry, e.g. POLYGON ((x y, x y, ...)).
WFS, WMS	Web services: WFS returns vector features, WMS a map image.

● Sources

- Consolidated text of the PGiK, Art. 7a(1)(14)(b), Art. 7c(1)(3), Art. 12(8), Art. 40a(2)(1), Art. 40c(3) (checked 3-4 October 2026): <https://api.sejm.gov.pl/eli/acts/DU/2024/1151/text.html>
- Technical standards regulation, § 3, § 6(1)(c) (buildings in group I of terrain details) and § 16 (0,10 m accuracy for group I): <https://api.sejm.gov.pl/eli/acts/DU/2022/1670/text.pdf>
- Pobieracz danych GUGiK plugin, version 1.3.9 and scope of data: https://plugins.qgis.org/plugins/pobieracz_danych_gugik/
- Pobieracz source code with the addresses of the LiDAR and orthophotomap index services: https://github.com/envirosolutionspl/pobieracz_danych_gugik
- Pobierz dane GUGiK plugin, version 1.2.12, parcels and buildings from the aggregated service: https://plugins.qgis.org/plugins/pobierz_dane_GUGiK/
- GUGiK plugin source code with the address of the aggregated EGiB WFS service: https://github.com/GlownyUrzadGeodezjiIKartografii/pobierz_dane_GUGiK

- Aggregated EGiB WFS service, from which the 296 test buildings (snapshot of 4 October 2026) were downloaded (service address to be added in QGIS): <https://mapy.geoportal.gov.pl/wss/service/PZGIK/EGIB/WFS/UslugaZbiorcza>
- Index of LiDAR measurement data (EVRF2007), years 2018-2026: <https://mapy.geoportal.gov.pl/wss/service/PZGIK/DanePomiaroweLidarEVRF2007/WFS/Skorowidze>
- LAZ sheet used in the test (Etk, 2025): https://opendata.geoportal.gov.pl/NumDaneWys/DanePomiaroweLAZ/82292/82292_1699329_N-34-81-D-a-2-2-2-1.laz
- GUGiK orthophotomap (WMS) as the figure background; this is a service address to be added in QGIS (Layer > Add Layer > Add WMS/WMTS Layer...), not a page to open in a browser: <https://mapy.geoportal.gov.pl/wss/service/PZGIK/ORTO/WMS/HighResolution>
- Geoforum: all counties make EGiB buildings and parcels available via WFS (6 May 2022): <https://geoforum.pl/news/32332/juz-wszystkie-powiaty-umozliwiaja-pobieranie-danych-egib-przez-wfs>
- GUGiK technical specifications: ASPRS classes in LiDAR point clouds, class 6 Building according to the LAS format: http://www.gugik.gov.pl/__data/assets/pdf_file/0018/24840/Zalacznik-nr-1-do-SIWZ-Szegolowy-Opis-Przedmiotu-Zamowienia.pdf
- QGIS: versions 3.44.15 LTR and 4.2.3, Windows installers: <https://qgis.org/download/>
- Deepness, description and licence: <https://github.com/PUTvision/qgis-plugin-deepness>
- Deepness, model zoo (buildings, land cover, roads, tree tops): https://qgis-plugin-deepness.readthedocs.io/en/latest/main/model_zoo/MODEL_ZOO.html
- Deepness, version 0.6.5: <https://plugins.qgis.org/plugins/deepness/>
- Deepness, package requirements (OpenCV up to 4.9, onnxruntime up to 1.17, NumPy below 2): https://github.com/PUTvision/qgis-plugin-deepness/blob/master/src/deepness/python_requirements/requirements.txt
- Deepness, issue no. 235: bug in 0.6.5 with NumPy 2 in QGIS 3.44: <https://github.com/PUTvision/qgis-plugin-deepness/issues/235>
- Buildings Segmentation model used in the test (ONNX file): <https://chmura.put.poznan.pl/s/MwhgQNhyQF3fuBs>
- SamGeo, stable version 1.2.0 and experimental 1.4.0, prompting methods: <https://plugins.qgis.org/plugins/samgeo/>
- SamGeo plugin source code (creation of SAM 1/2/3 models, SAM 3 and Hugging Face requirements; release 1.4.2 only here): <https://github.com/opengeos/qgis-samgeo-plugin>
- segment-geospatial library 1.4.2, calling SAM 2 with a box and a point: <https://samgeo.gishub.org>
- SAM 2 model (sam2-hiera-large) downloaded by SamGeo: <https://huggingface.co/facebook/sam2-hiera-large>
- GUGiK orthophotomap index, 2025 flight (RGB, 5 cm) for the test sheet: <https://mapy.geoportal.gov.pl/wss/service/PZGIK/ORTO/WFS/Skorowidze>
- Report "AI w pracowni geodezyjnej JDG" (AI in a sole-trader surveying practice, 3 October 2026), stage 2 of the implementation plan: NEW PROJECT internal document, no URL

Sources accessed: 3-5 October 2026.