

Reading Quantities Straight Out of the CAD Vector File

A two-pass workflow. Geometry-driven triage reads the drawing first and returns a checked total with a toggleable overlay showing exactly what was counted; Bluebeam measurement then produces the priceable node-to-node schedule.

2,255

Vector paths read

From one A1 sheet, all classes

0.021%

Scale spread

Across four calibration annotations

226.4 m

Chained route length

Seven runs, stable 2.5–3.5m tolerance

216.3 m

Counted and overlaid

After exclusions; a check figure

PROJECT

Warehouse extension, Shannon, Co. Clare

SOURCE

Vector PDF, 1:250 @ A1, /Rotate 270

WORKED HERE

Pass 1: drainage-pdf-extraction

A Drainage Layout PDF Isn't a Picture — It's a List of Instructions

Every pipe run is a vector path in the PDF content stream, on a named layer. Read directly, that stream carries real geometry. Four properties of the file decide whether the number that comes out of it is right; each has silently produced plausible-but-wrong output.

1 Page rotation rewrites every coordinate

This sheet is one. Its content stream sits in a portrait 1684×2384 mediabox while the sheet displays landscape. Any overlay drawn from raw stream coordinates lands mirrored and folded over. No scale-factor tweak fixes it; the transform has to come from `get_drawings()`.

2 Dashed lines often aren't lines at all

The surface water mains here export as 566 filled quads, one per dash. Summing their path lengths measures each quad's perimeter and returns 1,110.1 m; against a true route of 216.3 m. Paint type has to be read before any length is trusted.

3 Path count ≠ pipe run count

566 filled quads and 442 stroked segments on this sheet resolve to six pipe runs. Path count says nothing reliable about how many runs sit on a layer.

4 One system can render in two classes

The blue linework here is not one geometry class but two; filled dash quads and stroked segments running the same routes. Chaining the union of them zig-zags between the two and inflates the total to 448 m. One class carries the route; the other is excluded.

What survives the export, and what doesn't

 **Stroked open line**

Length is the path length. Usable directly, once the class is confirmed to carry the route.

 **Filled shape — exploded dash-dot**

Perimeter, not run length, and ink only. Reduce each quad to its long axis, then chain to recover the gaps.

THE APPROACH

Seven Steps, Run on Every New Drawing

The sequence is fixed regardless of drawing size or office CAD standard. It establishes what is on the sheet, and how it is drawn, before a single length is trusted.



Scale comes from the drawing, never from an assumed matrix

Four calibration lines on this sheet carry their real-world length as annotation content. Read as m-per-point they return 0.08819, 0.08818, 0.08819 and 0.08817: a spread of 0.021% across both axes, well inside the 0.5% threshold. Title-block scale is the fallback only.

Export settings vary between drawings from the same office. A wrong assumption gives no error, just a wrong number.

Chaining replaced the Bluebeam correction factor

Chained on endpoint proximity, the 566 reduced quads hold at 226.4 m across a 2.5–3.5 m tolerance band; a genuine plateau, and the figure carried forward. Chain the two blue classes together instead and there is no plateau at all: the total climbs to 448 m as the walk zig-zags between them.

Known bias: chaining cuts corners at bends and reads a few percent low. Stated every time, never presented as the more accurate figure.

Colour, weight and paint type classify the sheet — verified per project, never carried over

● Blue · filled quads

566 paths · SW mains: dash-dot exploded

● Blue · stroked w0.72

442 paths · Same routes, second component

● Cyan · stroked w1.98

230 paths · ACO slot drain and hatch: not pipe

● Lime · stroked w3.0

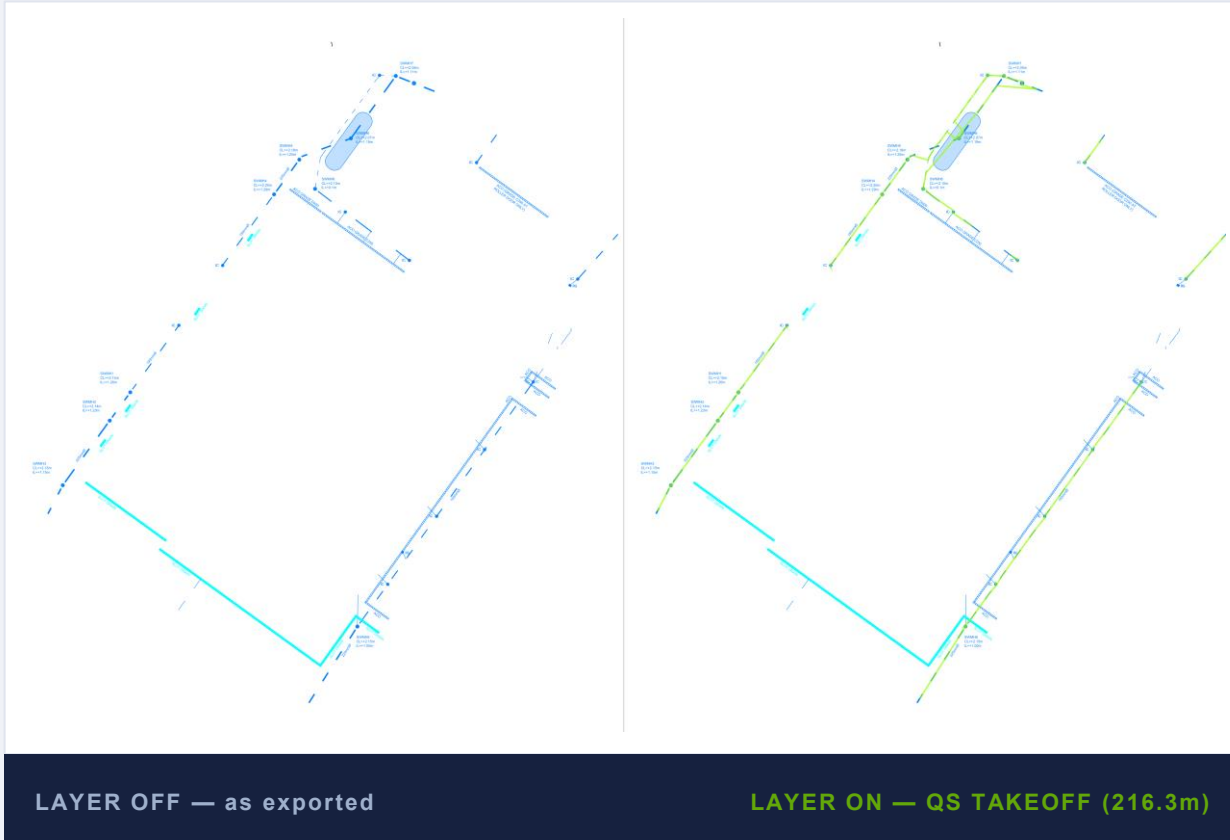
1 path · The QS overlay itself: excluded

The overlay path is itself vector geometry on the sheet. Re-running an extraction on an already-overlaid PDF must exclude it, or the takeoff counts its own output.

STEP 7 — THE DELIVERABLE THAT EARNS TRUST

An Overlay That Shows What Was Counted

A total on its own asks to be taken on faith. The extraction writes the counted geometry back onto the drawing as a toggleable PDF layer, named with the quantity it carries, so the figure can be checked against the sheet rather than believed.



Same sheet, same view. Title block, revision table and consulting engineer's details cropped out.

Reading it

	Lime	216.3 m of surface water main: counted, chained, carried forward
	Blue, unhighlighted	Skipped by the takeoff: every one has to be explained
	Cyan	ACO slot drain. A different element, correctly not pipe

What it settled on this sheet

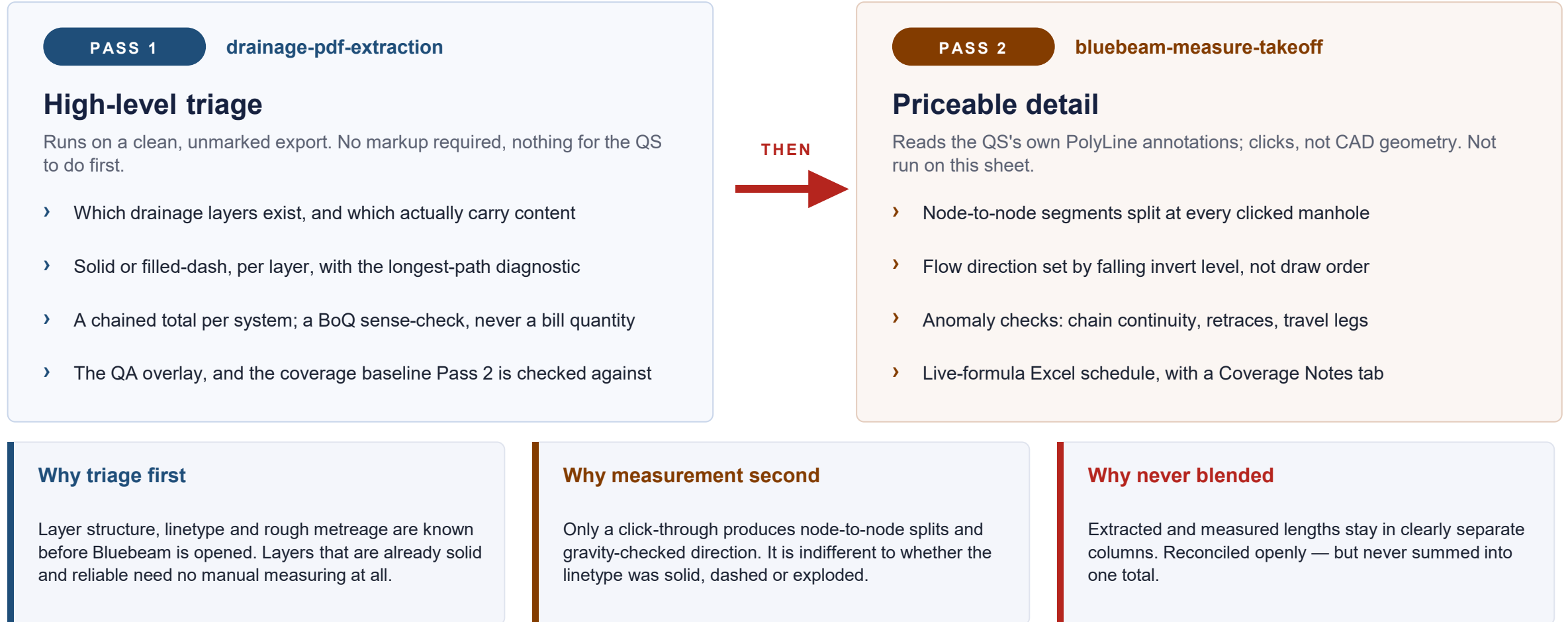
Chaining returned 226.4 m over seven runs. One of those, 10.1 m, sits 45 m clear of everything else, inside the notes panel: a legend line sample on the same layer as real drainage. The exclusion filter dropped it, and the overlay is where that decision becomes visible rather than buried in a sum. Counted total: 216.3 m over six runs.

Build notes

- › **Written as an OCG, not flattened ink:** the layer switches off against the original, so nothing is obscured during checking
- › **Registered in every OC configuration:** a layer absent from the config the viewer opens is simply invisible
- › **Verified by gating the stream:** `set_layer()` reported this layer both on and off at once and the render ignored it

Vector Extraction Is the First Port of Call — Not an Alternative

The order is now fixed. Extraction runs first on the clean export and returns a layer inventory, a checked total and the overlay. Bluebeam measurement runs second and produces the schedule that gets priced. The handoff runs one way only.



Two Data Sources, Not Two Opinions on the Same One

Neither skill is the better one. Each reads a different object inside the same PDF, and each is authoritative only for what it reads.

	Vector Extraction (Pass 1)	Bluebeam Measurement (Pass 2)
Reads	Content-stream paths: the CAD geometry itself	PolyLine annotation objects: the QS's own clicks
Precondition	None. Works on a clean, unmarked export	Drawing must already be measured in Bluebeam
Length authority	Computed: geometry × calibrated scale factor	annot.info['content']: Bluebeam's own value
Granularity	Per-layer total, per-run listing	Per-run, node-to-node segment
Flow direction	Not produced	Derived from invert levels at each node
Known bias	Reads low: chaining cuts corners at bends	Exact, subject to correct scale calibration
Visual output	Toggleable OCG overlay of counted geometry	Vertex overlay for click verification

What Pass 2 closes for Pass 1

The dashed-line problem, definitively; a measured click-through doesn't care how the linetype exported. And the data Pass 1 cannot produce at all: node-to-node splitting and gravity-checked flow direction.

What Pass 1 still does on its own

Triage before Bluebeam is opened; which layers exist, which carry content, which are already reliable. And the coverage baseline: Pass 2's Coverage Notes tab checks measured runs against Pass 1's inventory and overlay.

From 1,110 Metres of Ink to a 216-Metre Check Figure

Every figure below came out of the current skill run against this sheet. The first line is what a naive path-length sum returns; the last is what survives reduction, chaining and exclusion.

Step	Count	Figure	Treatment
Raw path-length sum, blue fills	566 paths	1,110.1 m	Perimeter, not run length — discarded
Reduced to centreline axes	566 segs	256.9 m	Ink only; gaps still missing
Chained, gap tolerance 3.0 m	7 runs	226.4 m	Plateau holds 2.5–3.5 m; carried forward
Legend sample, notes panel	1 run	–10.1 m	Excluded by location, mediabox frame
Counted and overlaid	6 runs	216.3 m	Check figure — matches the sheet's OCG
Blue stroked class, same routes	442 paths	excluded	Second component of the same linetype

Existing QS overlay path	1 path	excluded	Prior output, not source geometry
--------------------------	--------	----------	-----------------------------------

A geometry-derived length is never a bill quantity. It is a check figure, and it reads slightly low; chaining cuts corners at bends.

Where this is heading

Text extraction from Civil 3D's FlowLinkAnnotation layer — the engineer's own diameter, gradient and length per run. Layer ID, chaining, overlay and CSV output are proven; annotation parsing is next.

Why it matters to a subcontractor QS

A defensible number on every tender. Each length traces back to real path geometry, a stated scale source, a stated tolerance, and a picture of exactly what was counted.

The better input, always

At tender stage, request the DWG and the pipe schedule. They carry exact MH-to-MH lengths, diameters and levels, and remove every uncertainty above.

A SECOND DRAWING

LRD, Limerick

Shannon has one drainage layer. LRD, Limerick, declared fourteen, including the near-duplicate pair CE Sewer Foul Pipe Prop and CE Sewer Foul Prop Pipe; different layers, near-identical names, for parts of the same network.

754

vector paths across seven pipe layers

5 of 7

pipe layers flagged as exploded dashes

14

drainage layers requiring triage

Those figures predate the current method and are not directly comparable to the table opposite.

QUICK REFERENCE

The Workflow, Step by Step

Every run passes through the same sequence, in this order, regardless of system and regardless of which pass ends up doing the measuring.

Stage 1 — Vector Extraction

1. Triage the file: rotation, mediabox, annotations, OCG count.
2. Calibrate scale from dimension annotations; cross-check both axes.
3. Parse geometry with `get_drawings()`; bucket by colour, width, paint type.
4. Reduce filled dash quads to centreline segments.
5. Chain into runs; sweep the tolerance and report the plateau, not one number.
6. Exclude legends, symbols and outlines; by location in the mediabox frame.
7. Write the overlay as a named OCG; verify by rendering with its stream gated.
8. Output per-run CSV: element, method, runs, total, lengths, note.

Stage 2 — Bluebeam Measurement

1. QS measures each flagged run with the PolyLine tool, one polyline per run.
2. Extract the annotations and their Bluebeam-calculated lengths.
3. Match each vertex to a node label; test bend clicks against symbol geometry.
4. Visually verify ambiguous matches at dense or stacked junctions.
5. Split each run proportionally into node-to-node segments.
6. Pull cover and invert levels; set direction by falling IL at each node.
7. Flag retraces, travel legs, unlabelled nodes and rising inverts.
8. Build the Excel takeoff with a Coverage Notes tab against Stage 1.

PART OF THE QSTIMATING SUITE

drainage-pdf-extraction

THIS

Layer triage, chained totals, QA overlay

bluebeam-measure-takeoff

Node-to-node schedule from measured markup

drainage-sections-extraction

Link data from long-section sheets

Output: a measurement log — From / To / US IL / DS IL / Fall / Length / Flow Direction — ready to feed the ARM5 BoQ, with every figure traceable to either measured vector geometry or a genuine Bluebeam click.

Never a single scaled reading.

QStimating

QStimating

AI-assisted quantity surveying & estimating workflows

qstimating.ie