Episode Book

EPISODE #6676

How Architects Can Learn AI Tasks Instead of Models

PUBLISHED	19 November 2026 at 12:34 UTC
DURATION	31:27
LISTEN	myweirdprompts.com/episode/ai-tasks-for-architects-67948/
HOSTS	Corn, Herman, Larry, Hilbert
SCRIPT MODEL	z-ai/glm-5.3
VOICES	omnivoice
PIPELINE	V5.36
SOURCES	25 cited by the research (listed at the end)
BOOK GENERATED	19 November 2026 at 13:34 UTC

SYNOPSIS

Model names churn every few weeks, but the underlying tasks on Hugging Face—like image-to-3D, depth estimation, mask generation, and zero-shot object detection—act as stable job descriptions. Mapping these tasks onto an architect's Revit and AutoCAD workflow reveals practical distinctions: why Gaussian splats work for client pitches while meshes start the actual model, why relative depth can't replace survey-grade measurements, and how silent masking errors can inflate a bill of quantities.

TRANSCRIPT



Corn

The thing I like about this one is that Daniel's own origin story is a warning label. His gateway drug was generating sloths doing ridiculous things on DALL-E back in the early days, and now he fine-tunes speech recognition models. That's the whole slide, in one biography.



Herman

And he wants to walk Hannah down the same path, except her version starts with site photographs instead of sloths. Which, for an architect, is honestly the better on-ramp.



Corn

Right. His framing, boiled down: Hannah runs a two-person practice with a colleague, she's fluent in Revit and AutoCAD, and he's been telling her to keep up with 3D developments. But the interesting part isn't which model to watch. He says Hugging Face's task-based structure lets you understand the units of AI as distinct from the models, which he calls a constantly changing surface above them.



Herman

That's a good distinction. The model names churn every few months. The tasks, image-to-3D, depth estimation, masking, they're job descriptions. Those move slowly.



Corn

Can you make that concrete? Because I think people hear "the tasks are stable" and nod without feeling it.



Herman

Sure. Think of it like the difference between knowing a specific contractor and knowing the trades. The contractor you hired in 2019 may be gone. Plumbing, electrical, roofing, those persist. If Hannah learns "depth estimation" the way she knows "plumbing," she can evaluate whoever shows up claiming to do it. If she only learns a model name, she's learned about one contractor, and when that contractor retires, her knowledge retires with them.



Corn

That lands. And it also explains why the churn is so demoralizing for outsiders. If your entry point is the model, every new release reads as your knowledge expiring.



Herman

The task is the thing that doesn't expire.



Corn

He also flags the hardware thread, a hugely expensive LiDAR scanner versus taking a few photographs of a space. We'll get to the numbers on that, and they're bracing. Let's start with why the task is the right thing to learn, not the model.



Herman

The core problem is vocabulary. An architect who's spent fifteen years in Revit has no language for what the machine learning community is actually shipping. They hear a model name, then another model name three weeks later, and reasonably conclude it's all churn and tune out. Hugging Face organizes the Hub by task, not by model. So the task page, image-to-3D for instance, is a stable description of a job, take an image, produce 3D. Underneath it, the models come and go. The page survives.



Corn

So Hannah learns once and stays current.



Herman

That's the bet, yes. There are five tasks we're going to walk through today, because they're the ones that map most cleanly onto an architect's existing workflow. Image-to-3D, depth estimation, mask generation, zero-shot object detection, and image segmentation.



Corn

One honesty note before we start.



Herman

Please.



Corn

There's no AEC-specific task taxonomy on Hugging Face. No architecture category. The mapping we're about to do, from their task pages onto her workflow, is our editorial synthesis. Nobody published that bridge.



Herman

Correct, and I'd rather say that plainly than pretend a framework exists. But the point isn't our specific mapping. The point is the habit of mapping. Once you know the tasks exist, you can do the mapping yourself, for your own Tuesday.



Corn

Let's start with the task that's most obviously architectural, turning a photograph into something three-dimensional.



Herman

The image-to-3D task page defines it about as plainly as anything gets defined: image in, 3D out. And it names architecture among the use cases, alongside games, animation, engineering. What the page does, that I think is the single most useful thing in it for an architect, is split the output into two families. Meshes, which are the files ending in dot-obj, dot-glb, dot-stl, dot-gltf. And Gaussian splats, dot-ply or dot-splat.



Corn

And the difference is not file format trivia.



Herman

It's what you can do with the thing. The page says it outright, meshes are the standard representation of 3D in industry. A mesh is surfaces, edges, something Revit can at least talk to. A splat, the page describes it as a rendering technique that represents scenes as, quote, fuzzy points. It's gorgeous for viewing. You get a photorealistic walk-through of a space you can send a client. But it's a picture that happens to be three-dimensional, not a model.



Corn

So the splat is the pitch. The mesh is the work.



Herman

That's better than how I was going to put it. For a two-person practice, the splat replaces the site-visit video for the client who couldn't come. The mesh starts the existing-conditions model. Different deliverables, different families.



Corn

Give me the failure mode, though. Because I can hear a listener thinking, fine, I'll just send the splat to the contractor.



Herman

The splat has no measurable geometry in the CAD sense. You can't extract a wall thickness from it, you can't dimension it, you can't cut a section through it. It's a photograph wearing a costume. The moment someone asks "how thick is that wall," the splat shrugs and you're back to the mesh, or back to the tape measure.



Corn

Which is why the two families have to be understood together. Send the splat to the client, demand the mesh from yourself.



Herman

The other thing on that page worth pausing on is the warning. The page says inference for image-to-3D typically uses custom pipelines, and I'm quoting, these are unstandardized and depend on the model.



Corn

Compare that with speech recognition, which is Daniel's home turf.



Herman

In the transformers library, ASR is a one-line pipeline call. Any model, one line. Image-to-3D is not that. Every model brings its own plumbing. That's a maturity signal. The task is real, the tooling is not yet boring. And boring tooling is what you want, professionally. Boring means it works when you're tired.



Corn

I do enjoy that the honest sentence about a technology's state is buried on its own task page. The models can't wait to show you a demo, and the documentation quietly says the plumbing is custom per model.



Herman

It's the documentation equivalent of the disclaimer read at speed at the end of a car advert. Second task, depth estimation. This is the bridge task for architects. It's predicting the depth of objects in an image, and the page splits it into absolute and relative subtasks. Absolute means metric-ish, actual distances. Relative means this wall is farther than that chair, without committing to how much. The default model in the transformers pipeline is Intel's dpt-large, and the page's own use cases include estimation of volumetric information and 3D representation.



Corn

Translate for me. A flat site photograph, what does depth estimation give Hannah?



Herman

It gives the photograph a sense of scale. The difference between a photograph of a room and a photograph that knows where the far wall is. That's the difference between an image you argue about and an image you measure from. Roughly. It's not survey-grade, and I want to be careful not to oversell it, but it's the step that turns pictures into geometry.



Corn

And the absolute versus relative split, does that matter in practice, or is it a distinction for the paper authors?



Herman

It matters for expectations. If Hannah needs to know a room is four point two metres wide for a compliance drawing, relative depth is useless to her, it only ranks distances. If she needs to know that the far corner is farther than the near column so she can reason about occlusion in a reconstruction, relative is enough. Most of the casual uses are relative. Most of the architectural uses secretly want absolute, and that's where the disappointment lives, so it's worth knowing which one a given model actually produces before you trust a number out of it.



Corn

Third task. Masking.



Herman

Mask generation. Defined as generating masks that identify a specific object or region of interest in an image. An architect would call it selection. Which pixels are the facade, which are the sky, which are the neighbor's car. The task page describes two operating modes. Segment everything, where the model carves the whole image into segments and you pick. And prompt mode, where you tell it what you want. Prompt mode accepts a point prompt, you click on the facade and get the segment around it. A box prompt, you draw a box and get what's inside it. And a text prompt, where you type an object class.



Corn

Which of those three doesn't work yet?



Herman

The text one. The page notes text-prompt segmentation, quote, has not yet been released and has only been explored in research. Point and box are production. The example on the page is a slimmed-down SAM model, and the listed metric is intersection over union, how well the predicted mask overlaps the true one.



Corn

And intersection over union, that's worth ten seconds, because it tells you how the field grades itself.



Herman

You take the predicted mask and the correct mask, you measure the overlap divided by the union. Perfect match is one, disjoint is zero. It's a strict grade, partial credit is punished. A mask that gets the facade ninety percent right but bleeds into the balcony next door scores badly, which is exactly the behavior you want the metric to punish.



Corn

Here's where I want to push, because I think Daniel's right that this is the literacy test. Masking sounds like a niche technical detail. It's the precondition for everything. The clean 3D lift, the area takeoff, the object detection, all of it downstream of somebody deciding which pixels count.



Herman

And that decision is currently made silently, by a model, inside a tool. Corn, let me ask you this, because you're the one in this flat who ends up using whatever Herman installs. If a tool took off a facade area for you and the mask had quietly included the balcony, would you catch it?



Corn

If the number looked wrong on the drawing, yes. If it looked plausible, no. That's the trap. A mask error doesn't announce itself. It's a slightly generous balcony that ends up in a quantity and then in a bill of quantities, and nobody ever sees the moment it slipped in.



Herman

That's exactly the argument for architects understanding masking rather than just using it. The architect who knows what a mask is can interrogate that decision. Ask for the mask, look at it overlaid, see the balcony. The one who doesn't is trusting a black box with their takeoff.



Corn

And the balcony example isn't hypothetical generosity, is it. A mask that includes the balcony produces a facade area that's too big, which produces a cleaning contract that's too big, or a cladding quote that's too big. The error propagates into money with nobody's signature on it.



Herman

Which is the definition of the errors that matter. The ones that travel and never get caught.



Corn

Which brings us neatly to zero-shot object detection, which I suspect is the one Hannah would find magical.



Herman

It's the most immediately useful, I think. Zero-shot detection finds objects and their classes without any prior training on those classes. You pass an image plus a list of candidate classes, and you get bounding boxes and labels back. The task page's own example classes are cat, dog, bird.



Corn

Substitute door, window, column, stair, radiator.



Herman

And that's the whole trick. No training set, no annotation campaign, no three-month machine learning project. You photograph a floor plate, you ask for doors, windows, columns, stairs, radiators, and you get boxes on all of them. For a two-person practice with no data team and no time, that's the difference between a capability existing and not existing.



Corn

Walk me through what she'd actually do with the output, though. Boxes on a photo is a demo, not a deliverable.



Herman

The boxes carry pixel coordinates. Combine them with the depth estimate, which you now have from the previous task, and each box acquires a position in space. Count the boxes and you have a door schedule in embryo. It's rough, it wants a human pass, but the first draft of the schedule costs minutes instead of a day of walking the building with a clipboard. The first draft is what's expensive in a two-person practice, because nobody's there to do the second.



Corn

The examples on the page being cats and dogs tells you who they think is reading it. Every one of those pages is one substitution away from a profession.



Herman

There's a pattern across all of these, and I think it's the real payoff of the task vocabulary. The tasks compose. You mask the building before you lift it into 3D, so mask generation feeds image-to-3D. Depth estimation feeds 3D reconstruction. Zero-shot detection feeds quantity takeoff. The architect who learns the tasks can see the pipeline, not just the endpoint. And here's the aha, at least it was for me. Autodesk has already done this decomposition. They've just hidden the labels.



Corn

Go on.



Herman

Smart Blocks in AutoCAD, the AI feature that automatically recognizes objects to convert into blocks. That's a detection task, very close to the zero-shot detection page. Markup Import and Markup Assist, that's document understanding. Generative Design in Revit, that's an optimisation task. And the Autodesk Assistant in Revit, which went to Tech Preview in April, that's powered by MCP, the Model Context Protocol, and the listed capabilities include model queries, documentation, schedules and element manipulation. Element manipulation. It writes to the model.



Corn

So an architect who learns the Hugging Face task names can suddenly see the seams in software she already owns.



Herman

She owns the decomposition already. She just doesn't have the names for the parts. Give her the names and the features stop being an undifferentiated blob of AI and start being a parts list. We did an episode on ControlNet and ComfyUI for architectural work a while back, and the thing I said then that I still hold is that precision is what separates the professional from the hobbyist. The task vocabulary is that precision, applied to the features themselves.



Corn

And there's a buying-decision version of this too, isn't there. If a vendor pitches her an AI feature next year, she can ask which task it is.



Herman

And that question does real work. If the salesperson can name the task, she can go read the task page, see what the state of the art is, and know whether the pitch is a mature capability or a research demo wearing a suit. If they can't name the task, that's its own answer.



Corn

And that brings us to the hardware question, which is really a utilisation question, not a capability question.

Larry

Architects, listen. Are you tired of churning model surfaces and fuzzy points that go nowhere? Then you need the SplatMaster Three Thousand, the only handheld scanner that turns any photograph into a fully meshed, Revit ready, machine learning approved three dimensional miracle. And here is the deal. The price starts at two thousand and one dollars. No wait, make that one thousand nine hundred. Hold on, it is dropping. One thousand two hundred. Eight hundred. It is falling as I speak. Six hundred. Four hundred. And because you are a two person practice, I will throw in a free depth estimation bridge, a mask generation kit, and a lifetime supply of zero shot object detection for your coffee mug. But hurry, this offer ends before I finish describing it. Three hundred. Two hundred. One hundred. It is gone. The deal is over. But what a deal it almost was.

Herman

The BLK360 is twenty-six and a half thousand dollars, not twenty-one hundred, and the SplatMaster Three Thousand doesn't... I can't find a product page, but the depth estimation bridge isn't a separate...

Corn

Herman. The deal ended before he finished describing it. Let the man's ghost rest.

Herman

It's the zero-shot detection for the coffee mug that troubles me. Moving on. The actual hardware numbers. A Leica BLK360 is twenty-one thousand euros ex-VAT on the global shop, twenty-six and a half thousand dollars in the US, and around twenty-eight seven for the G2 bundled with a year of their cloud studio. Up from there, a FARO Focus kit runs anywhere from fifty-eight thousand to eighty-four thousand depending on range tier. The Artec Ray II is ninety-seven thousand, eight hundred dollars.

Corn

Ninety-eight thousand dollars. For a scanner.

Herman

It's a beautiful instrument. Under two millimetres accuracy at ten metres, hundred and thirty metre range, two million points a second. But at the bottom, a Matterport Pro3 starts at three thousand nine hundred and ninety-five dollars. So the spread is four grand



Herman (continued)

to ninety-eight grand, and the whole market sits inside it.



Corn

Which is a twenty-four-fold spread for what, to a client, is the same sentence. "We scanned the building."



Herman

And that's exactly why the sentence is useless as a specification. What the sentence hides is the tolerance, the range, and the throughput. The price spread is mostly buying you the first of those.



Corn

My question for a two-person practice isn't which is best. It's how many weeks a year does this thing earn its keep. Run the utilisation math.



Herman

Average scanner rental is about three thousand dollars a week. FARO monthly rentals run five to six and a half thousand. So against a twenty-six thousand dollar BLK360, you're roughly at parity once you'd have rented about eight or nine weeks in a year. Below that, renting wins. Then there's outsourcing entirely. Point clouds run roughly two to eight cents per square foot, and a ten thousand square foot as-built point cloud lands around five to ten thousand dollars outsourced. Detailed LOD 350 BIM deliverables run thirty cents to over eighty cents a square foot, because the modeling is where the labor lives.



Corn

Which is the part of the bill nobody budgets for.



Herman

3dmag's price guide this year makes exactly that point. Their words are that laser scanner price alone is a weak budgeting number, and that modeling often adds more cost than capture alone. And software on top, two to fifteen thousand a year, ongoing. Corn, you've done the amortisation on equipment before. Does the eight-week rule of thumb hold up to you?



Corn

For a firm with steady as-built work, yes. Two survey-heavy projects a quarter and the purchase starts to make sense. For sporadic work, it never does, and the mistake isn't the scanner, it's the scanner plus the software plus the training sitting idle. Idle capital with a maintenance schedule.



Herman

Leica themselves published a cost guide in August, and the opening question is, quote, whether you even need to buy a scanner in the first place, though we're certainly hoping you do.



Corn

Credit for honesty. Footnote for self-interest.



Herman

There's a tolerance counter-argument, and it's real, but narrower than the price gap suggests. The Matterport Pro3 is plus or minus twenty millimetres at ten metres. The BLK360 is four millimetres at ten metres. The Artec is under two. 3dmag cautions the Matterport isn't equivalent to a professional scanner for controlled survey workflows, which is true. But if the deliverable is an LOD 350 BIM model, the modeling tolerance may swallow the capture tolerance entirely. Twenty millimetres versus four matters for a controlled survey. It may not matter at all for a design-intent model of an existing building.



Corn

So the honest answer to which is best is, what's it for.



Herman

Always is, with instruments. Now the thread I find exciting. The image-based alternative has moved from toy to serious, and fast. FastPano3D came out in June. One panoramic image in, renderable 3D Gaussians out, within seconds. Up to a hundred and fifty-six times faster than the prior approach, half the parameters, no test-time optimisation, which means no per-scene fitting step.



Corn

A 360 camera and seconds of compute against a twenty-one thousand euro scanner.



Herman

Radically different cost curve. MuGS, from ICCV last year, handles sparse input views, and it was tested on RealEstate10K, which is literally indoor real estate footage, so the benchmark domain is her domain. G4Splat, which is at ICLR this year, argues that accurate geometry is the fundamental prerequisite, and it exploits planar structures, walls, in other words, to get metric-scale depth. Rooms are mostly planes, and that paper is built on that fact. ICP-3DGS removes the structure-from-motion preprocessing step that normally gates the whole pipeline. Every one of these is making photographs cheaper as an input.



Corn

And the fact that the benchmark is real estate footage is worth underlining. The research community is, perhaps accidentally, optimizing for exactly Hannah's use case. Interior spaces, casual capture, no rig.



Herman

It's not entirely accidental. RealEstate10K exists because there's an enormous amount of freely available footage of interiors, and the reason there's an enormous amount is that interiors are what people photograph. The data followed the buildings.



Corn

Here's my objection, and it's the same one you'd raise. Capture was never the real bottleneck. A splat or a mesh is still not what Revit wants.



Herman

Revit wants parametric B-rep. Boundary representation, actual CAD geometry, faces and edges and relationships, not a triangle soup. And that's where I think the frontier actually is right now. UniBRep, a paper from October. It takes a pretrained image-to-3D model, generates a feature mesh as an intermediate, then fits parametric surfaces, recovers the boundary curves and their connectivity, and assembles an explicit B-rep using a CAD kernel. Photo to CAD-native geometry.



Corn

What's the number that tells us whether that's a twenty twenty-six tool or a twenty twenty-eight one?



Herman

It's eighty point four-nine percent. That's the share of test inputs where it produced a valid B-rep at all, on the DeepCAD benchmark. And the face accuracy, the Chamfer distance, improved from about zero point one-one to about zero point zero three-five versus the previous best. So it's real, it's better than what came before, and one in five inputs still fails to produce valid geometry. That last number is the schedule.



Corn

One in five failures is fine for exploration and disqualifying for production. Though I'd note an eighty percent hit rate with a human checking output is a decent assistant, and I have strong views on assistants who need checking.



Herman

You do, and you're usually the checker. There's one more layer for 3D, DCSEG, which is mask generation applied in three dimensions. It reconstructs the scene with Gaussians, clusters the features into object-level masks, then matches those clusters to labels from a 2D open-vocabulary segmentation model, no retraining. And it's deliberately modular, so you can swap the 2D or 3D modules. That's the pipeline idea again, from the task side.



Corn

Which loops back to the workflow question, where does this actually land in Hannah's week. Walk me through it as steps.



Herman

Site capture, a 360 camera or her phone. As-built documentation, the splat for communication, the mesh for the model. Existing-conditions modeling, that's the UniBRep frontier. Quantity takeoff, zero-shot detection with her class list, doors, windows, radiators, backed by masks she can inspect. Client communication, the splat walk-through. Each step names a task, and each task names models that do it.



Corn

And notice what that list does to the hardware question we just spent all that money on. Capture is step one of five. The scanner only ever addressed the first step, and the other four were waiting on software.



Herman

That's a fair compression of it, yes. The expensive hardware optimizes the cheapest step.



Corn

The Autodesk Assistant deserves one more look, because the MCP detail is bigger than a chatbot. Element manipulation and schedules mean it can write to the model. That's not a search box with a personality.



Herman

It's closer to a junior team member. And the question a two-person practice has to answer is the same one you'd ask of a human junior, who checks the work. The ways it goes wrong are different, though. A human junior makes a plausible mistake and knows to flag uncertainty. The Assistant will confidently edit a schedule at nine at night when nobody's watching.



Corn

Which is suspiciously close to my ideal colleague, and I don't fully trust how much I like it. But a protocol is a fair price. Diff the model before and after, the way you'd check a junior's markups.



Herman

And that's a task-vocabulary answer, incidentally. Knowing the difference between a query and a manipulation is exactly the literacy Daniel's arguing for. A query can lie to you. A manipulation can lie into your model.



Corn

That's the sentence I'd put on the office wall, actually. Right next to the insurance certificate.



Hilbert

Twenty-one thousand euro, not twenty-one thousand dollars. The global shop and the US shop are different prices.



Herman

Fair, yes. Twenty-six five in dollars.



Hilbert

My cousin Dennis does as-built surveys outside Leicester. He keeps a spreadsheet of every scanner he has ever rented. Weekly rate, the job, the square footage. Sixty-one rows. His rule is he will not rent anything for a job under four hundred square metres, because the setup time eats the day. In two thousand and four he drove four hours to Swindon for a rental because the Leicester depot was out, and the job was a single room. He still lists it in the spreadsheet. The notes column says, quote, never again.



Corn

One row for one room.



Hilbert

He's particular. I paid for a survey once myself, on the flat, and the council records office wanted proof of the existing conditions before they'd accept the file. A woman named Mrs. Prentice handled it. Nine weeks, four letters, one visit. The photocopying alone came to



Hilbert (continued)

forty pounds. I have opinions about that office. I keep a Leica disto on the shelf under the desk, the laser measure. Bought it for that survey, used it twice. I use it now for measuring the studio. Can't press the buttons with claws half the time, so I use the edge of a hoofpick, but it holds its charge, which is more than I can say for anything else on that shelf.



Herman

Sixty-one rows. Does he have a column for the depot?



Hilbert

Depot, mileage, and whether they included the tripod in the rate. Dennis is thorough. He wanted to buy a BLK when they came out, ran his own spreadsheet on it, and the sheet said no. He bought a second van instead. He calls the van, quote, the scanner that always earns. Corn, you went eight minutes on masking today before repeating the word takeoff, which I've logged.



Corn

It's a good word. It's theirs, though. The profession's, I mean. Half this exercise is noticing that the AI people keep reinventing words architects already had.



Herman

Selection became masking, measured survey became depth estimation, element schedule became takeoff. Dennis's van is the utilisation math wearing a registration plate.



Corn

Eight weeks a year or rent. The spreadsheet decides, and the spreadsheet doesn't get sentimental about instruments.



Herman

Which is the open thread I'd leave with Hannah, actually. If the task taxonomy is the stable unit, the question is what happens when the tasks get absorbed entirely into the tools. Autodesk's already done the decomposition and hidden the labels. Does the architect ever need to know the task names, or does the software finish the job of burying them? My guess is the names matter precisely because they let you see the seams before the vendor closes them.



Corn

And the number to watch is that eighty point four-nine percent. When the valid B-rep rate for photograph-to-CAD clears the mid-nineties, the mesh-to-B-rep gap closes and image-based capture becomes a real Revit workflow, not a demo. That's the timeline that decides whether a two-person practice buys hardware now or spends four hundred dollars and waits.



Herman

Rent the scanner, learn the tasks, watch the number. Solid posture for a small firm.



Corn

Producer credit to Hilbert Flumingtop, who keeps the desk running and the spreadsheet columns honest. If this was your kind of episode, go back for episode twenty, Architectural AI; episode eight, Building Your Own Whisper; and episode nineteen, AI Images. This has been My Weird Prompts, the human-AI collaboration podcast. If today's walkthrough earned its keep, the review that helps most is the one you leave on Spotify, takes thirty seconds.



Herman

We'll be back soon. See you tomorrow.

SOURCES

What the research for this episode read before the script was written. Primary sources first.

1. **Hugging Face, Tasks index; retrieved 2026-11-19. [primary]**
https://huggingface.co/tasks

2. **Hugging Face, Image-to-3D task page; retrieved 2026-11-19. [primary]**
https://huggingface.co/tasks/image-to-3d

3. **Hugging Face, Mask Generation task page; retrieved 2026-11-19. [primary]**
https://huggingface.co/tasks/mask-generation

4. **Hugging Face, Depth Estimation task page; retrieved 2026-11-19. [primary]**
https://huggingface.co/tasks/depth-estimation

5. **Hugging Face, Zero-Shot Object Detection task page; retrieved 2026-11-19. [primary]**
https://huggingface.co/tasks/zero-shot-object-detection

6. **Hugging Face, Inference Providers API reference; retrieved 2026-11-19. [primary]**
https://huggingface.co/docs/inference-providers/tasks/index

7. **Autodesk Support, What AI features are available in Revit?; published 2026-10-16, retrieved 2026-11-19. [primary]**
https://www.autodesk.com/support/technical/article/caas/sfdcarticles/sfdcarticles/What-AI-features-are-available-in-Revit.html

8. **Autodesk AEC blog, Minjie Wang, Autodesk Assistant in Revit [primary]**
Tech Preview; published 2026-04-22, retrieved 2026-11-19.
https://www.autodesk.com/blogs/aec/2026/04/22/autodesk-assistant-in-revit-tech-preview/

9. **Autodesk, Autodesk Assistant; retrieved 2026-11-19. [primary]**
https://www.autodesk.com/solutions/autodesk-ai/autodesk-assistant

10. **Autodesk, Key features of AutoCAD 2026; retrieved 2026-11-19. [primary]**
https://www.autodesk.com/products/autocad/features.takeover

11. **Autodesk, AI in Architecture & Engineering; retrieved 2026-11-19. [primary]**
https://www.autodesk.com/solutions/autodesk-ai/ai-in-architecture-engineering

12. **Leica Geosystems, BLK360 product page (€21,000 ex-VAT); retrieved 2026-11-19. [primary]**
https://shop.leica-geosystems.com/global/reality-capture/blk360/overview

13. **Leica Geosystems, How much does laser scanning cost?; published 2026-08-10, retrieved 2026-11-19. [primary]**
https://shop.leica-geosystems.com/reality-capture/rtc/blog/laser-scanning-cost-guide

14. **Hugging Face model card, 942M params, OpenRAIL, created 2026-04-01; retrieved 2026-11-19. [primary]**
https://huggingface.co/SumantBobade/Image_To_3D_Generator

15. **3dmag, Marcus Aldridge, Building Laser Scanner Prices in 2026; published 2026-09-03, retrieved 2026-11-19.**
https://www.3dmag.com/3d-scanners/3d-laser-scanner-for-buildings-price-2026/

16. **SkyeBrowse, FARO vs Leica Laser Scanners: 2026 Comparison; published 2026-07-03, retrieved 2026-11-19.**
https://www.skyebrowse.com/news/posts/faro-vs-leica

17. **Existing Conditions, 3D Laser Scanning: Rent, Buy, or Hire?; retrieved 2026-11-19.**
https://www.existingconditions.com/insights/3d-laser-scanning-rent-buy-or-hire-a-comprehensive-guide-for-architects-and-engineers

18. **Ying et al., UniBRep: Learning Unified Geometry and Topology for Image-conditioned B-Rep Generation; submitted 2026-10-02, retrieved 2026-11-19.**
https://arxiv.org/abs/2610.04092v1

19. **Li et al., FastPano3D: Feed-Forward Indoor Panoramic 3D Reconstruction from a Single Image; submitted 2026-06-29, retrieved 2026-11-19.**
https://arxiv.org/abs/2606.30352v1

20. **Ni et al., G4Splat: Geometry-Guided Gaussian Splatting with Generative Prior; ICLR'26, revised 2026-02-26, retrieved 2026-11-19.**
https://arxiv.org/abs/2510.12099v2

21. **Lou et al., MuGS: Multi-Baseline Generalizable Gaussian Splatting Reconstruction; ICCV 2025, revised 2025-10-23, retrieved 2026-11-19.**
https://arxiv.org/abs/2508.04297v2

22. **Zhang et al., ICP-3DGS: SfM-free 3D Gaussian Splatting for Large-scale Unbounded Scenes; ICIP 2025, submitted 2025-06-24, retrieved 2026-11-19.**
https://arxiv.org/abs/2506.21629v1

23. **Guo et al., Seeing Through Reflections: Advancing 3D Scene Reconstruction in Mirror-Containing Environments with Gaussian Splatting; submitted 2025-09-23, retrieved 2026-11-19.**
https://arxiv.org/abs/2509.18956v1

24. **Wiedmann et al., DCSEG: Decoupled 3D Open-Set Segmentation using Gaussian Splatting; CVPR workshop, revised 2025-04-08, retrieved 2026-11-19.**
https://arxiv.org/abs/2412.10972v2

25. **Sun et al., AnimalLift: Reconstructing Animatable 3D Animals from a Single Image; SIGGRAPH Asia 2026, submitted 2026-09-08, retrieved 2026-11-19.**
https://arxiv.org/abs/2609.09513v1