The World Should Not Have to Be Rediscovered by Every Robot

A public essay on shared perception infrastructure for physical AI

Spatial Fabric concept paper

The central idea: instead of forcing every robot to build its own private picture of the world, we can let the environment maintain a shared picture that any authorized machine or application can use.

The problem is not that robots cannot see

Imagine an autonomous robot moving through a warehouse. It has cameras, LiDAR, perhaps radar, and enough onboard compute to identify people, map walls, avoid obstacles, and plan a route. From the robot's perspective, this is impressive. From the perspective of the building, however, something odd is happening: every robot is independently rebuilding a model of the same place.

A worker walks behind a shelf. The robot cannot see that worker, so for a moment the worker effectively disappears from the robot's world. Another robot approaches from the opposite direction and performs the same reconstruction. A third robot does it again. The warehouse may already have cameras, fixed sensors, access points, phones, and other machines observing the same space, yet each stream of information remains largely isolated.

That is the deeper problem. We have become very good at making individual machines perceive. We have not yet created a common perception layer for the environment itself.

A simple example

Consider a robot approaching the end of a warehouse aisle. A tall shelf blocks its view. A person is walking toward the same intersection from the other side. The robot's onboard perception may not detect the person until both are close to the corner. The limitation is not intelligence. The information simply never reached the robot.

Now add two fixed sensors mounted above the aisle. One can see the worker. The other can see the robot. A shared system combines those observations and knows that two moving objects are on intersecting paths. Before the robot has direct line of sight, it can already be told: a person is three meters beyond the corner, moving toward you.

Nothing magical happened. The environment simply knew something the robot did not.

Why this matters

Robotics is moving out of controlled cages and into warehouses, hospitals, factories, stores, sidewalks, homes, and eventually many other public and private spaces. The more machines share space with people, the more wasteful it becomes for every machine to perceive the world alone.

This duplication has practical consequences. Each robot needs expensive sensors and compute. Each vendor maintains its own perception stack. Each machine has blind spots. A sensor upgrade must often be repeated across a fleet. And information that is obvious from one part of a building may remain unavailable to a machine only a few meters away.

Researchers and large technology companies are already moving toward this idea. NVIDIA's Halos Outside-In Safety work uses external infrastructure cameras to extend robot perception beyond onboard sensors. Oxford researchers have explored what they call 'robotic inversion': shifting sensing and compute from individual robots into the surrounding infrastructure. These efforts are important because they validate the premise that the environment can become part of the robot's perception system.

The proposed solution: a shared perception layer

The proposed solution is a platform I call Spatial Fabric: a neutral software layer that turns many spatial sensors into one continuously updated model of a place.

A building might contain radar sensors, cameras, LiDAR, phones, door sensors, robot-mounted sensors, or future devices that do not exist yet. Spatial Fabric would not require every sensor to speak the same native language. Instead, each device would be connected through an adapter that describes what it can observe, where it is located, when an observation occurred, and how confident that observation is.

The system would then align those observations into a common coordinate system, synchronize them in time, decide when multiple sensors are seeing the same object, and maintain persistent tracks. The output would not be 'camera frame 184' or 'radar point cloud 73.' The output would be a useful statement about the world: there is a person here, a forklift there, this robot is moving north, this area is free, and these two trajectories may intersect.

From sensor data to world state

This distinction is important. Raw sensor data is useful to perception engineers. Most applications should not need to understand it. A robot developer should be able to ask the environment for nearby people or free space. A warehouse system should be able to subscribe to collision risks. A phone should be able to display the same shared map. An augmented-reality application should be able to place information into the same coordinate system.

The platform therefore becomes less like a radar product and more like infrastructure. Radar may be an excellent starting sensor because it works in darkness, measures motion directly, preserves

more privacy than cameras in many settings, and can observe some movement that optical systems struggle with. But the long-term system should be sensor-agnostic.

Why a common layer is valuable

The value of a common layer is that perception becomes reusable. One ceiling sensor can help ten robots instead of being duplicated ten times. A new sensor can improve the shared model without redesigning every machine. A robot entering a facility can immediately benefit from information gathered by infrastructure it did not bring with it. A phone can become both a temporary sensor and a viewer. Different manufacturers can participate without agreeing on one physical sensor.

There is also a powerful safety advantage. Onboard perception remains important and should not disappear. But an outside view can reveal hazards beyond the robot's current field of view. The two perspectives can reinforce each other: the robot sees what is immediately around it, while the environment provides context beyond that horizon.

What has to be solved

The hard part is not drawing a shared map on a screen. A useful platform has to solve several difficult engineering problems reliably.

First, sensors must be registered and calibrated. The system must know where each device is and how its local coordinates relate to the building. Second, observations must be synchronized so that two sensors reporting the same moving person at slightly different times do not create two people. Third, the system must estimate uncertainty rather than pretending every measurement is exact. Fourth, identity must persist as an object moves from one sensor's field of view to another. Fifth, the system must continue working when a sensor fails, a network connection drops, or the environment changes.

Machine learning becomes valuable above this foundation. Classical geometry, tracking, filtering, and sensor fusion can establish where things are and how they move. AI can then help answer richer questions: Is this a person or a cart? Is this movement unusual? Is someone likely to enter a robot's path? Are observations from different sensors likely to belong to the same individual? What changed in the environment?

A realistic first step

The first prototype does not need a smart city or a hundred robots. It can fit inside one room.

Place two or three millimeter-wave radar units around a space and connect them to one edge computer. Calibrate them into a shared coordinate system. Have a person walk through overlapping fields of view. The first milestone is simple but fundamental: the system should maintain one person, one identity, and one trajectory even as different radars alternately see that person.

Then add a small mobile robot. Put a visual obstruction between the robot and a person. The fixed sensors should see the person, the shared world model should publish that person's position, and

the robot should change its route before its own onboard sensors have direct visibility. A mobile application can display the same world state in real time.

That demonstration contains the entire idea in miniature: the environment perceives, the platform combines, and the machine benefits.

The long-term vision

If this works, the long-term opportunity is much larger than warehouse safety. Shared spatial perception could support humanoid robots, delivery machines, drones, hospitals, factories, assisted-living environments, smart homes, augmented-reality systems, and autonomous infrastructure.

Today, when a laptop joins a network, it does not need to understand the physical wiring of every machine on the internet. Standard layers hide that complexity. Physical AI may eventually need a similar abstraction for space.

A machine entering an instrumented environment should be able to ask simple questions: Where am I? What objects are here? What is moving? Where are the people? Which areas are free? What changed? What is likely to happen next?

The goal is not to eliminate the robot's own senses. The goal is to stop treating every machine as if it were alone.

The broader idea is straightforward: as machines become more capable, the spaces around them can become capable too. Instead of building intelligence only into robots, we can build a shared perception layer into the physical world itself.

Further reading

1. NVIDIA, "Inside NVIDIA Halos for Robotics: A Full-Stack Functional Safety System for Physical AI," June 2026. Source
2. NVIDIA Isaac ROS - modular perception, mapping, localization and ROS 2 integration. Source
3. University of Oxford, "Rethinking Robotics: From Intelligent Machines to Intelligent Environments," 2026. Source