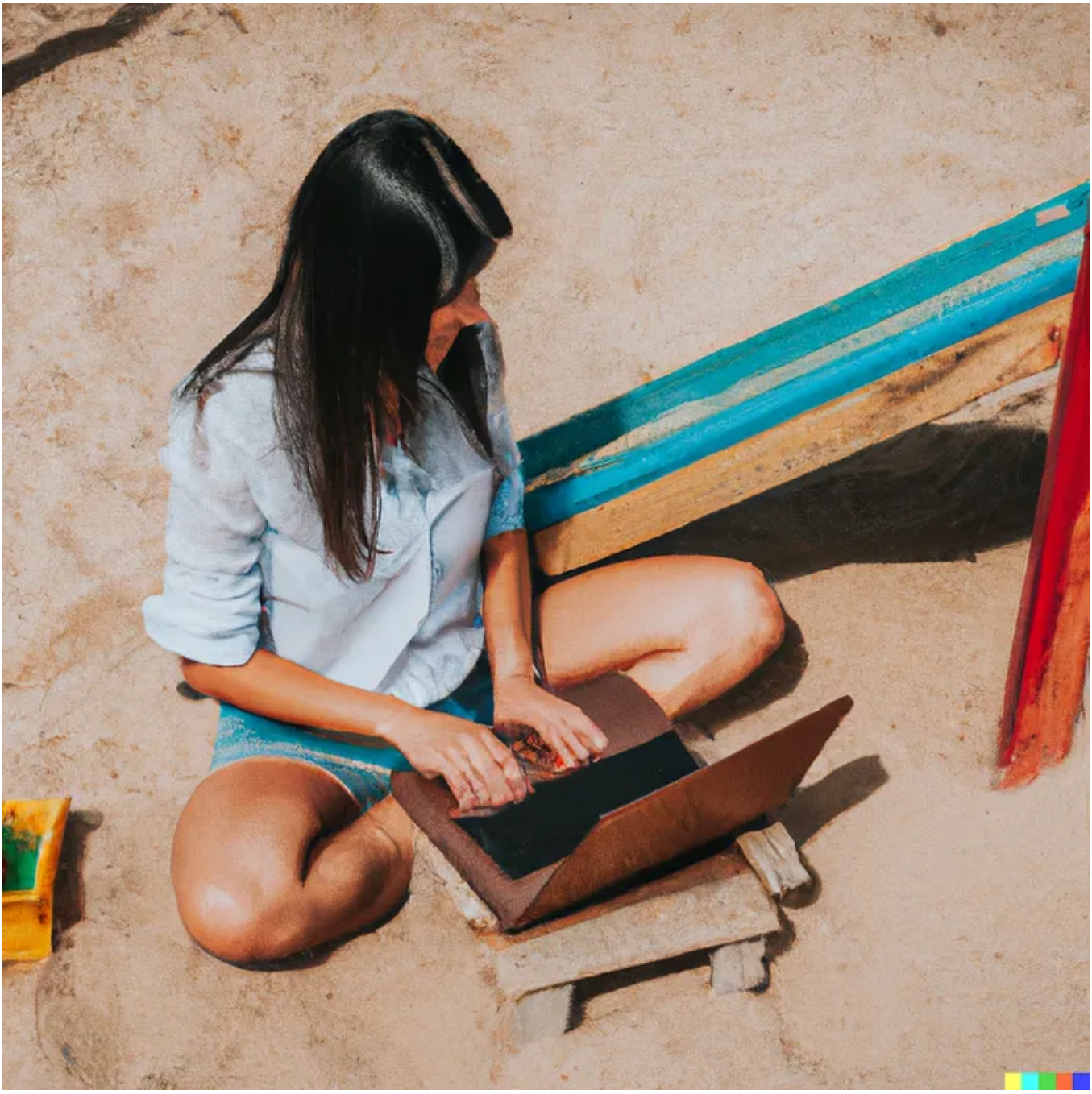
Build sandboxes and let them play

Roger Herger

5 min read · Mar 28



"A woman with laptop in a sandbox for children", generated using text-to-image model DALL-E2 by OpenAI.

Curiosity is one of the biggest drivers of humans. This is the third article in our series on industrial artificial intelligence. It is about machine learning engineers and why they should play in sandboxes.

Sooner or later (read sooner) on your path to industrial artificial intelligence (AI), you will get to the point of wanting to set up and train machine learning (ML) models yourself. This comes with some challenges and opportunities that we will now take a look at.

Your old IT infrastructure is not your new one, but part of it

One of the many challenges you may face is that your current IT infrastructure is not well-suited for the new AI world. Today, it is secure and meets all the requirements from the various quality guidelines. Your office applications work reliably, your engineers have powerful CAD machines, and your financial metrics are available on time. In addition, you have good IT support that helps your employees quickly and reliably.

But all of a sudden a bunch of wild-eyed data nerds come along and don't want to know anything about it. They expect you to be able to work with their own device, or at least in an environment where they are admin or "sudoers". Otherwise they are gone. Before you can understand what a sudoer actually does, the pain shoots into your head and your IT department gears up for battle.

Build sandbox environments

We cannot judge all the different legal, regulatory, or quality requirements for your products and your IT infrastructure for manufacturing. We assume that there are many, and that they are implemented according to best practice.

But we can share some ideas that we have already implemented in such cases. Specifically, it's about creating test environments that provide the necessary flexibility to try out new ML technologies. We like to call these environments "sandboxes".

At our company maXerial, sandboxes are virtual machines (VMs), hosted locally or in the cloud. We host the machines on servers that are compliant with our customers' data protection requirements. There are guidelines for handling different types of data. They regulate how data must be transferred, encrypted and stored, or which data may be processed in the cloud at all. For instance, every dataset is routinely encrypted and login mechanisms force 2-factor authentications.

We work almost exclusively with Linux VMs. These VMs are fast to create, extremely flexible, well-documented, and can be easily mirrored if necessary. On the VMs we

create virtual environments to try out our ML ideas.

You don't know how big or small your VM should be? No problem: Start small. In a few clicks your machine can be made bigger — much bigger. Modern cloud environments give you this flexibility and are easy to use. They give your team the ability to create their own sandboxes and try new things. And as in an ordinary sandbox, if you don't like it, flatten it and rebuild.

Sandboxes are not productive systems

One thing is key to remember — and you have to tell your data nerds now and then — sandboxes are not productive systems. That's also why I like to call them "sandboxes": By itself, sand is not weatherproof. It needs the cement and water to mix the concrete to build stable structures.

Hence, to move from a development environment to a productive system, you need clear rules and ways of doing it. Which tests have to be made, which libraries are needed, what versions, etc.? Only then can you safely transition to the real production system.

The good thing is that normally you can easily pack your environment out of the sandbox and bring it into your production environment. Today's container tools like Docker do an extremely good job of solving exactly such problems. Moreover, the cloud providers have these tools on offer to make the processes as simple as possible.

Your IT infrastructure becomes cheaper, more flexible — and more secure

True: Cloud is not everyone's cup of tea. We understand that. However, we think that the concepts at the large cloud providers with sophisticated role and rights management options surpass local security concepts. We see it this way: The most brilliant cyber security minds work there. They can do significantly more than we can do locally, and at a reasonable price because of the scaling.

Plus, you get access to fantastic computing power, right when you need it. You are flexible, depending on whether you need more GPU power in one case or significantly more CPU capacity in the other. Typically, you only pay for the machine when it's running.

We find another idea extremely exciting: The requirements for your office hardware per employee tend to decrease. For example, we use Chrome OS on a discarded 10-year old Macbook Air and access the cloud with it. Works.

Or even better: In Liechtenstein and Switzerland, we have beautiful mountain landscapes. When I was sitting in a mountain hut the other day, I had to briefly kick off the training of a new model. “Let it crunch while I enjoy life,” I thought to myself. So I logged in via my hotspot and the company VPN. I opened the SSH tunnel. What a feeling: In my fingers, the power of the most modern IT infrastructure, racing off with just a few commands in the shell, and in my view the perfect mountain panorama! That’s what I call a successful division of labor between me and the cloud!

Stay tuned to our next article, which will discuss typical problems that can be solved in your company using ML.

Further reading

This is the third article in our series on industrial artificial intelligence (AI). More articles in this series (list updated on release):

(1) How to bring AI to your manufacturing company

(2) Get machine-readable data for industrial AI

(3) Build sandboxes and let them play

(4) Problems you can solve with ML in your company

(5) Your route to success in industrial AI: Think big, start simple

(6) From pilot to maintainable AI technology stack

(7) What you can learn from your smartphone for industrial AI

Machine Learning | Artificial Intelligence | Cloud Computing | Manufacturing | Virtualization