Privacy, Data Protection in PerCV.ai platform

**How computer vision and PerCV.ai ensures my privacy?**

Artificial intelligence, of which computer vision is a part, has become increasingly important. There are a lot of discussions and arguments for AI and data issues related with large tech companies like Google, Facebook, Amazon etc. Also, governments across Europe and worldwide are investing in AI and discussing the legal and business framework.

In Irida Labs we are developing Vision AI solutions, using our PerCV.ai platform that integrates computer vision, machine learning and a proprietary data engine. PerCV.ai turns video and images into metadata and actionable insights without storing any video or image data. There are usually two stages in developing a Vision AI solution with PerCV.ai, namely a) the evaluation or development stage and b) the operational stage or runtime stage. We will explain how data are used in both stages.

PerCV.ai privacy and data protection at the operational stage

Deploying a Vision AI solution at operational stage means that the vision system or vision sensor receives as input image(s) or video(s) and produces metadata like the number of people or cars, or an event related with specific objects (like product recognition). In this stage PerCV.ai ensures privacy in two ways, firstly with edge AI and secondly with the vision sensor design where possible. Edge AI means that all videos and image data are processed locally at the vision sensor in real-time, the metadata are extracted at the edge device and only these metadata may be passed to a cloud or other infrastructure. No image or video data are stored at the edge AI during the operational stage. This is a strong point of PerCV.ai since we can deploy Vision AI solutions in a variety of edge platforms ranging from powerful GPUs and VPUs to tiny MCUs and DSPs. More details can be found in our partnership section of Irida Labs web site.

Another point for ensuring privacy and data protection at the operational stage has to do with the vision system design component of the PerCV.ai platform. The vision sensor resolution and optics, along with the installation geometry, are always chosen in a way to avoid covering or "looking" at areas that are not important for the deployed Vision AI solution. For example, if PerCV.ai operates in a factory, we focus on a production line or area of interest and not in a panoramic view of the factory; or if we operate in a parking scenario, the focus is on the parking slots and not on any road or sidewalk that is out of scope. See below typical deployment examples.
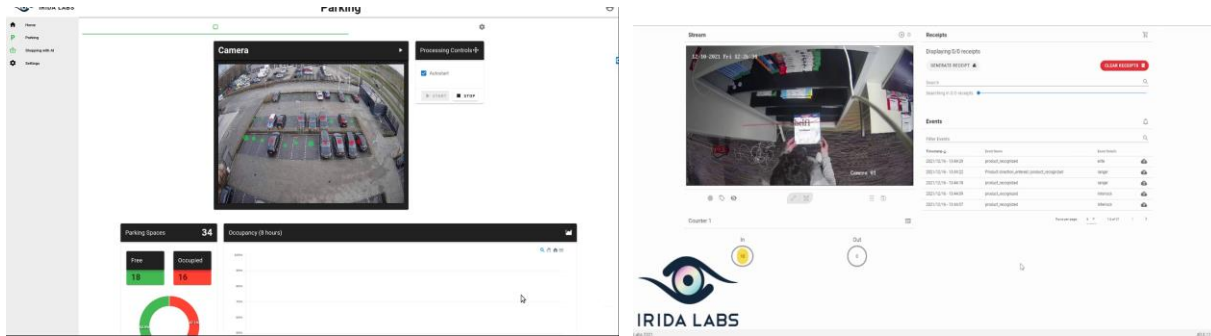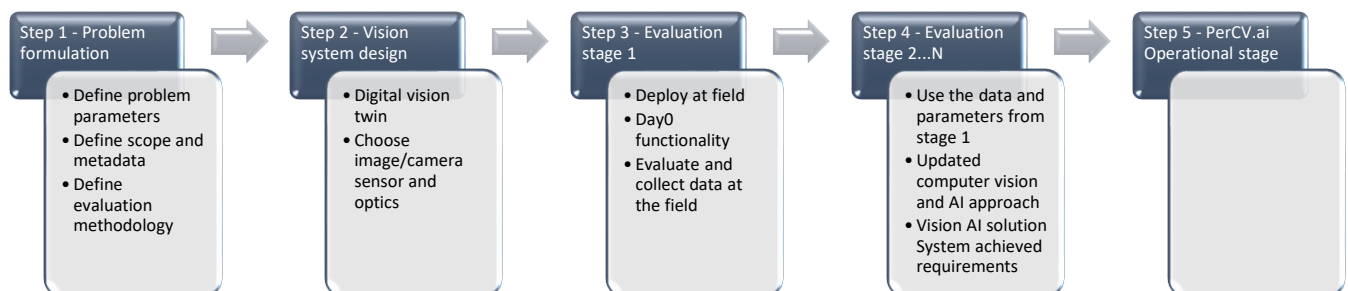
PerCV.ai privacy and data protection at the evaluation or development stage

Developing a Vision AI solution for a real -world environment is a complex task. In Irida Labs we have developed PerCV.ai that brings together all the necessary components for successfully powering Vision AI at the edge. The PerCV.ai components are presented in a high- level diagram below.

But then, how you can you start using PerCV.ai? Let's discuss the PerCV.ai workflow below and how data are treated.

Step 1 - Problem formulation
- Define problem parameters
- Define scope and metadata
- Define evaluation methodology

Step 2 - Vision system design
- Digital vision twin
- Choose image/camera sensor and optics

Step 3 - Evaluation stage 1
- Deploy at field
- Day0 functionality
- Evaluate and collect data at the field

Step 4 - Evaluation stage 2...N
- Use the data and parameters from stage 1
- Updated computer vision and AI approach
- Vision AI solution System achieved requirements

Step 5 - PerCV.ai Operational stage

In this workflow there is data flow from PerCV.ai users (our customers) to Irida Labs. We will call them *PerCV.ai external data*. That raises questions like:

1. Why are these data needed?
2. What other types of data is Irida Labs using?
3. Where are PerCV.ai external data stored? In which format and type?
4. How are PerCV.ai data transferred?
5. Who has access to PerCV.ai external data?
6. How PerCV.ai external data are used for the specific customers and other customers?
7. What happens to PerCV.ai external data after we reach the Step 5 ?

Why are these data needed?

Irida Labs has a long and strong experience in computer vision (before the AI and machine learning buzz wave). As a result, we know that real world efficient Vision AI solutions must be adaptable. Of course, machine learning and deep learning is far more efficient than "traditional" computer vision for most applications, but still the efficiency depends on the data and now more and more on the smart data. We need to make a distinction here with very generic problems where mass amount of data, collected by Google or Tesla etc., is the only way. PerCV.ai is not targeting these types of problems. Therefore, the PerCV.ai external data are needed to provide the adaptation of the computer vision and deep learning algorithms to the exact real-world conditions that are described at the Problem formulation stage of PerCV.ai workflow. External data are also used for reaching the extremely high accuracy often needed by our partners, which cannot be achieved without their input.

What other types of data is Irida Labs using?

Irida Labs has developed a proprietary Data Engine as part of the PerCV.ai platform. This Data Engine includes various sources of image and video data like a) open datasets, b) simulated or synthetic data and c) data generated by Irida Labs in our testing facilities. These types of data are used in the initial Evaluation Stage and in what we call Day-0 performance of the PerCV.ai vision solution. Most of the times we achieve a high accuracy based on these data in terms of ML, which is improved further with PerCV.ai external data as explained above. Following this first deployment, the solution is further improved by tuning the performance with respect to the target environment. Afterwards, the performance if further improved by incorporating customer feedback as well as corner cases.
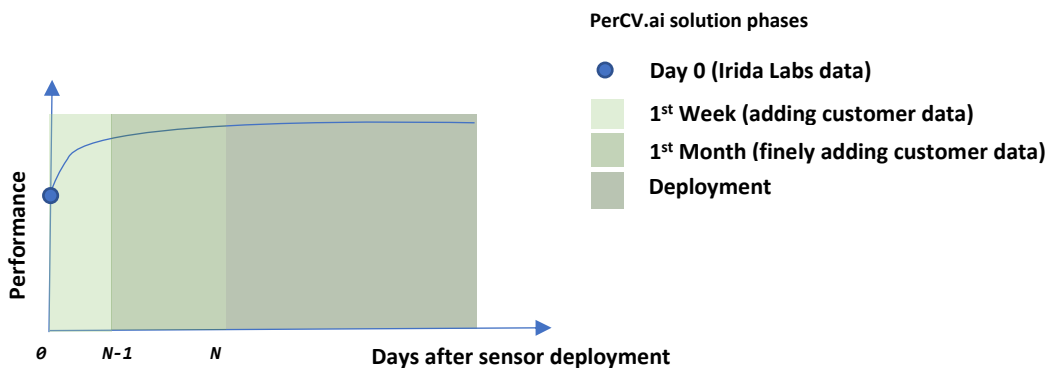
Where are PerCV.ai external data stored? In which format and type?

PerCV.ai external data are stored on a dedicated data server at Irida Labs. Only the required Irida Labs staff has access to each project data and this access is granted by the Project manager that Irida Labs assigns to each customer. On Irida Labs server, PerCV.ai external data are a) encrypted, b) accessible only with permissions by the data administrator, c) accessible only from the local network.

How do we use data at Irida Labs ?

Irida Labs is working with in house data, in house synthetic data and customer data. In house data are collected by Irida Labs, synthetic data are data generated with Computer Graphics and Machine Learning algorithms at Irida Labs.

How are PerCV.ai data transferred?

PerCV.ai external data can be sent to Irida Labs via:

    a. PerCV.ai platform data exchange tool.

When uploading the data on PerCV.ai platform, data are sent to PerCV.ai cloud storage where they are kept for a limited amount of time and in encrypted format. On PerCV.ai cloud storage access is granted only by the company's Project manager via API and Machine Learning Operations (MLOPs).

    b. Sensors connected to PerCV.ai platform

By attaching sensors in PerCV.ai platform and allowing data acquisition.

    c. Physical transfer via hard drives by the customer

In this case, the customer sends the data stored on hard disks with a preferable encryption approach and provides the required keys to Irida Labs to decrypt the data in order to be sent to the "Raw data" pool.

<u>Who has access to PerCV.ai external data?</u>

Data gathered with our partners and customers are processed only in Irida Labs premises and only by Irida Labs staff. Access is provided only to Irida Labs team members that are involved in the specific project – Vision AI solution, and their path within the company (i.e. path for storing and processing, personnel access, results) is always monitored. Part of the work is to prepare the PerCV.ai data for ML training and evaluation purposes, which also involves data annotation and checking for specific objects or events in an image or video. Although annotation or labelling is a mind-numbing work, it is very crucial not only for the data protection pipeline but also for following ML and computer vision tasks. Therefore, in Irida Labs we take it very seriously and it is carried out by Irida Labs staff with proper employment contracts and confidentiality terms in order to avoid outsourcing that will increase risks with quality of ML data and data security.

<u>How PerCV.ai external data are used for the specific customers and other customers?</u>

Let's start with the easier last part. PerCV.ai external data along with the data generated by the Data Engine are used only for the specific PerCV.ai customer and NOT for other customers.

The PerCV.ai data are used for annotation and labelling and subsequently for ML and deep learning purposes. That also involves preparing datasets for training, validating and evaluating the ML models but also the overall system approach. These data are used to increase the Day-0 initial performance of the vision AI solution in order to meet the requirements described in Step 1 of the PerCV.ai workflow. The evaluation stage may be repeated depending on the different real-world conditions that we want to cover but is always carried out with the participation and approval of the customer.

<u>What happens to PerCV.ai external data after we reach Step 5?</u>

Irida Labs data policy is NOT to store any data that are not needed for specific purposes; therefore, when we reach the Operational stage, our primary choice would be to delete or return to our customer any PerCV.ai external data acquired and used during the Evaluation stage. Anything else will be part of a mutually agreed Data Processing Agreement for the Operational Phase.