



OAAX: One Standard for AI Vision on Any Compute Platform

Robin van Emden

Senior Director of Data Science

Network Optix



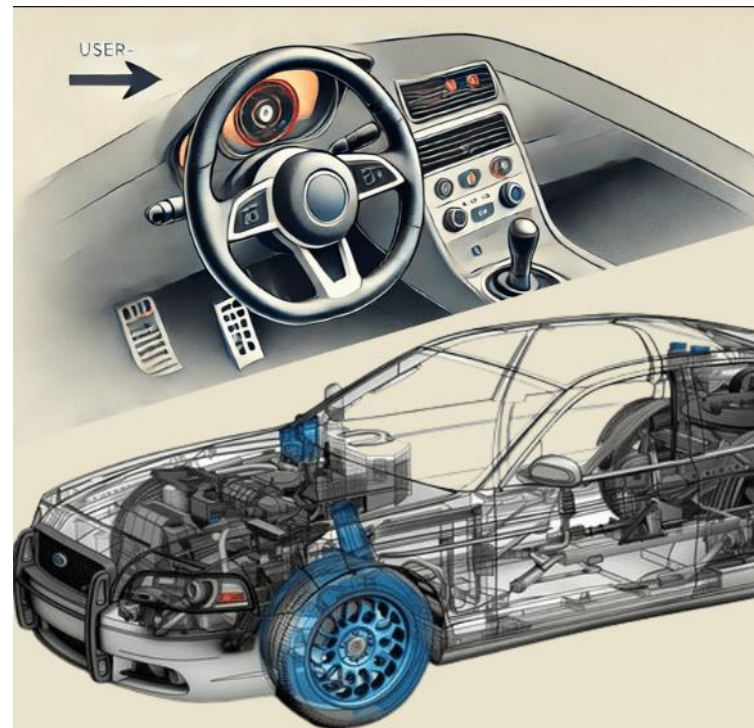
Open
AI Accelerator
eXchange

Outline

- The problem OAAX addresses
- Proposed solution
- OAAX's impact on chipmakers and users
- Ecosystem and collaboration
- OAAX in practice
- Roadmap and future development
- Get involved: Linux Foundation project & community resources

The value of standardized interfaces

- The standard design of cars:
 - has consistent features like pedals, gears, dashboard, trunk...
 - boosts usability and mass adoption
 - eliminates the need for training or license for each car brand
- Likewise, OAAX:
 - has a unified interface
 - boosts mass adoption of AI accelerators
 - allows ease of integration of any OAAX-compliant AI accelerator



The problem that gave rise to OAAX

Before

Deploying on **CPU/GPU**

```
$ pip install torch
```

```
import torch
model = torch.load(...)
model = model.cuda()
# model = model.cpu()
```

Now

Deploying on **XPU X**

1. Learn about **X** API & docs



2. Setup development env



3. Experiment with examples



4. Adapt example code



5. Integrate custom code



6. Compile code for target

Next

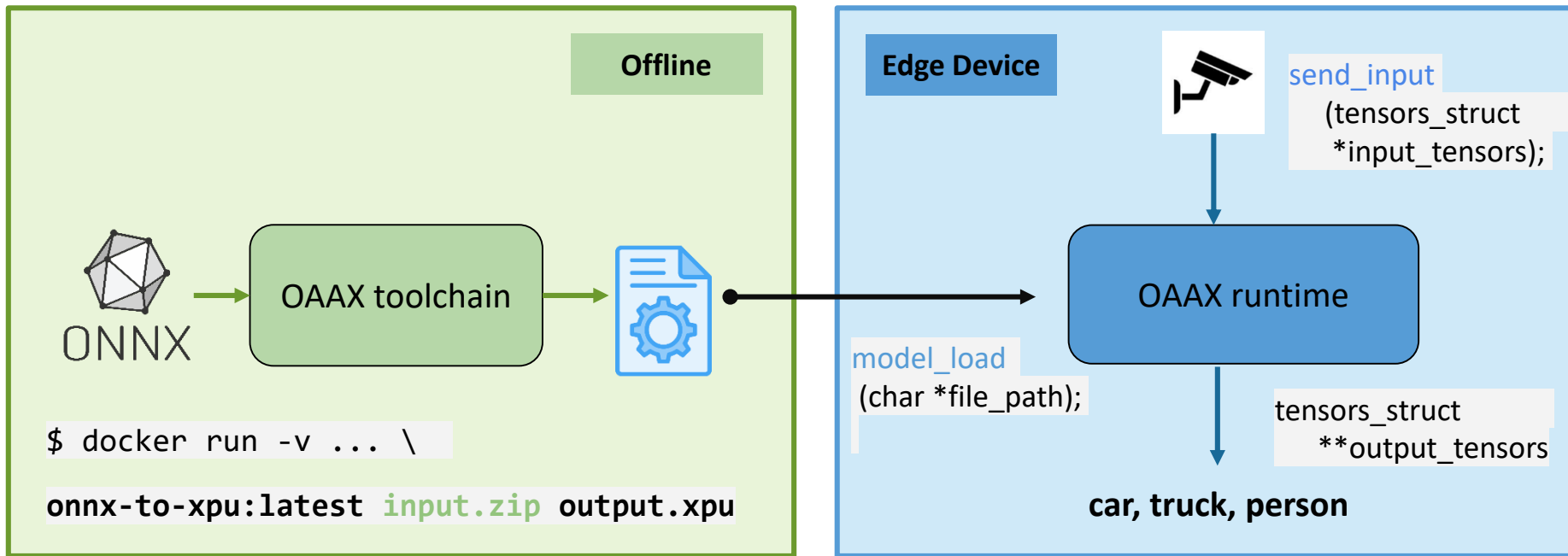
Deploying with **OAAX**

Download runtime library

```
$ docker run
  onnx-to-xpu:latest
  input.zip model.xpu
```

```
runtime.load("model.xpu");
runtime.run(tensors_in);
```

Solution using OAAX



OAAX proposes a standardized interface to **compile a model** and a process to **run it**.

OAAX for users

- Makes it easier to deploy AI models
- Reduces time, effort and money for deploying a trained model
- Allows them to focus on training models
- Enables smooth transition from one AI Accelerator to another
- Reduces time to market for companies

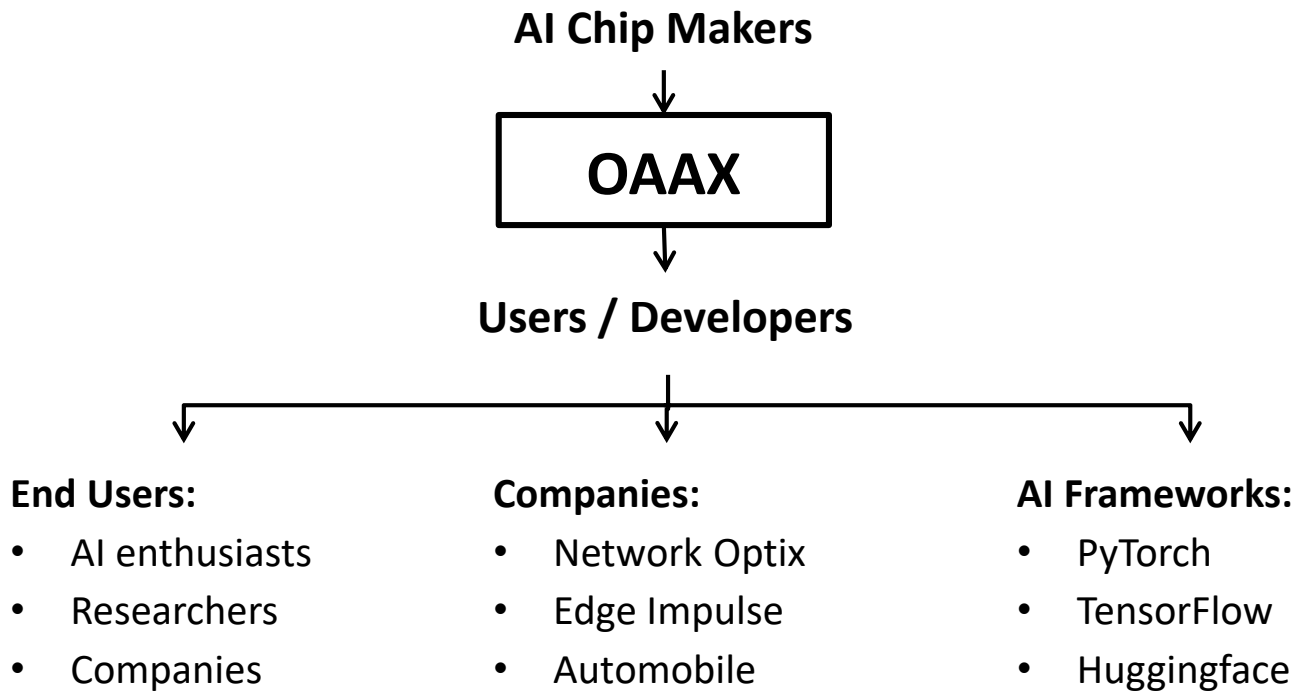
OAAX for AI chip makers

- Simplifies the process of showcasing their hardware, saving time and effort
- Enables new XPUs to compete on equal footing with existing ones
- Allows them to access broad segment of users
- Enables them to access the ecosystem of examples and tools already developed

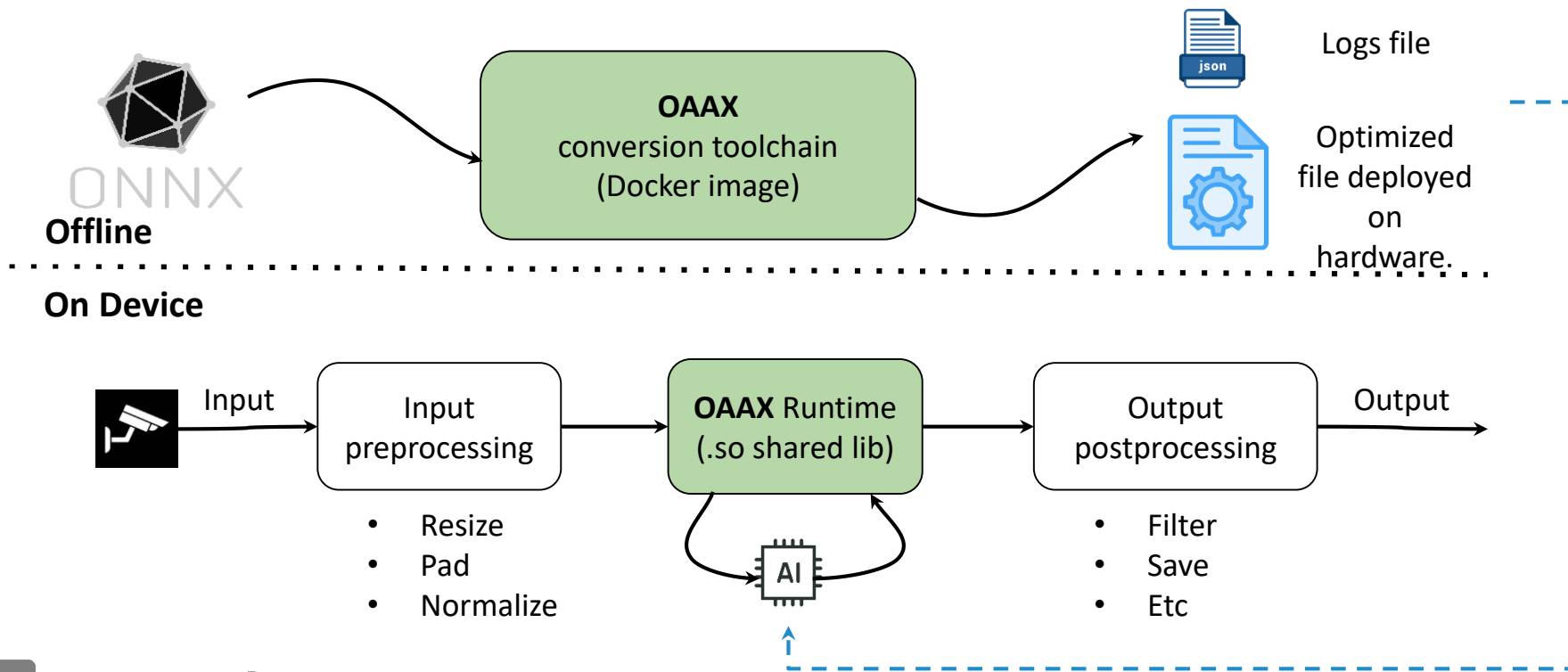
Current state of OAAX

- **OAAX simplifies AI deployment:** It enables effortless deployment of AI models across various hardware platforms.
- **Proven in real-world use:** As the first adopter, Network Optix uses OAAX to deploy AI models from the Nx Cloud to Nx supported servers and devices at the edge.
- **Open and collaborative:** OAAX is a project under the Linux Foundation.
- **Broad hardware support:** OAAX provides toolchains and runtimes at different levels of maturity for a growing list of AI accelerators, including:





OAAX in practice – general overview



OAAX from a user perspective

**User trains and exports
model to ONNX**

```
input_tensor = torch.rand((1,3,256,256));  
model = MyModel();  
torch.onnx.export(model, input_tensor,  
    "model.onnx", input_names=["input"])
```

**Compile ONNX to the
XPU format**

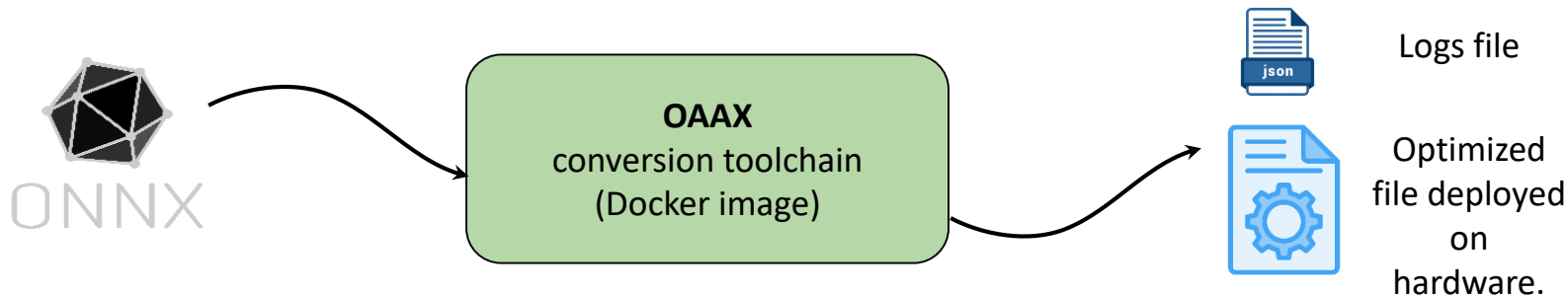
```
$ docker pull  
onnx-to-xpu:latest
```

```
$ docker run  
onnx-to-xpu:latest  
input.onnx model.xpu
```

```
int main() {  
    runtime = load_runtime("./xpu.so");  
    runtime.runtime_initialization();  
    runtime.runtime_load("./model.xpu");  
  
    while (true) {  
        runtime.send_input(input);  
        tensors_struct output;  
        runtime.receive_output(output);  
    }  
  
    runtime.runtime_destruction();  
    return 0;  
}
```

**Run the model on the
XPU**

Conversion toolchain interface

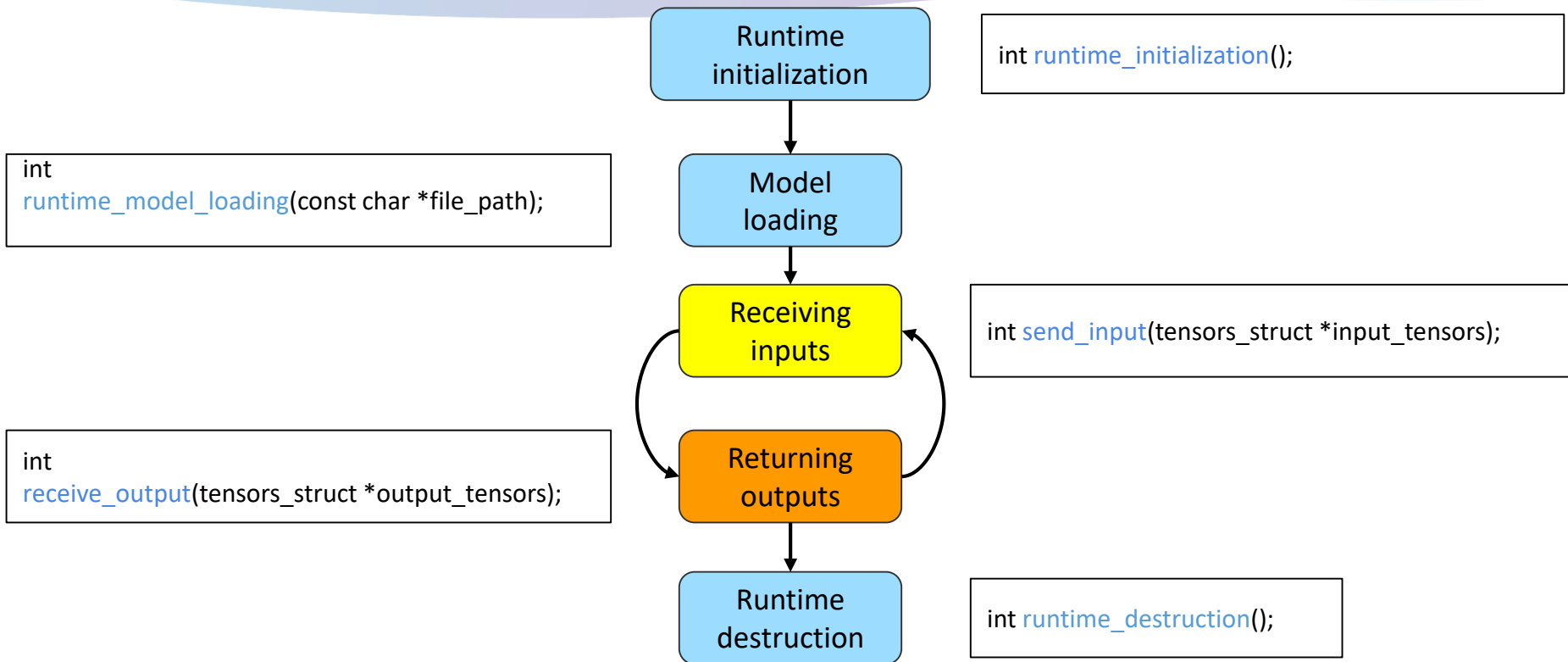


A process to go from a generic model specification (ONNX), to a specification that can be executed by the target specific runtime.

- Implemented as a standalone **Docker image**: easy to distribute and to maintain, while being totally flexible regarding the implementation of the conversion process.
- Usage:

```
> docker run -v ./app onnx-to-xpu:latest /app/input.onnx /app/output.xpu
```

Runtime interface - process



Runtime interface - definition

```
int runtime_initialization();

int runtime_initialization_with_args(int    length,
                                     char**  keys,
                                     void**  values);

int runtime_model_loading(const char*  path);

int send_input(tensors_struct*  input_tensors);

int receive_output(tensors_struct**  tensors_out);

int runtime_destruction();
```

```
typedef struct tensors_struct {
    size_t      num_tensors;
    char**      names;
    tensor_data_type* data_types;
    size_t*     ranks;
    size_t**    shapes;
    void**      data;
} tensors_struct;
```

- The interface is:**
- Minimalist.
 - Abstract.
 - Works for any kind of AI model: vision, LLM, multimodal.

- Support more AI chips: SiMa, Axelera, RockChip, etc.
- Improve the standard to accommodate for needs of both users & AI chip makers.
- Implement tools to make it easier for users to the toolchains & runtimes at scale.
 - Auto-detection of supported AI accelerators on a machine.
 - Download the correct runtime and toolchain programmatically.
 - Benchmark and profiling.

✦ Support the Future of Open AI Acceleration!

OAAX is now a **Sandbox Project** under the **Linux Foundation AI & Data**.

We're building an open, collaborative ecosystem for AI acceleration —
and we invite developers, companies, and innovators like you to join us!

- 🔧 Contribute your expertise
- 🔍 Collaborate on groundbreaking solutions
- 🚀 Help shape the future of AI infrastructure
- 👉 **Support OAAX — Let's build the future, together!**



OAAX Linux Foundation Project – Reach out

- **Sponsor:** **Network Optix** <https://networkoptix.com>
- **Tech lead:** **Ayoub Assis** aassis@networkoptix.com
- **OAAX Team:** **Robin van Emden** rvanemden@networkoptix.com
 Josef Joubert jjoubert@networkoptix.com
 Maurits Kaptein mkaptein@networkoptix.com

OAAX Linux Foundation Project – Resources

Links to OAAX related resources:

GitHub repository:

<https://github.com/OAAX-standard>

Project home:

<https://www.oaax.org/>

Linux Foundation AI & Data project site:

<https://lfaidata.foundation/projects/oaax>

2025 Embedded Vision Summit

We warmly invite you to visit us at the
Network Optix booth, #302.

We'd love to connect and show you
what we're working on!



OAAX: One Standard for AI Vision on Any Compute Platform

Robin van Emden