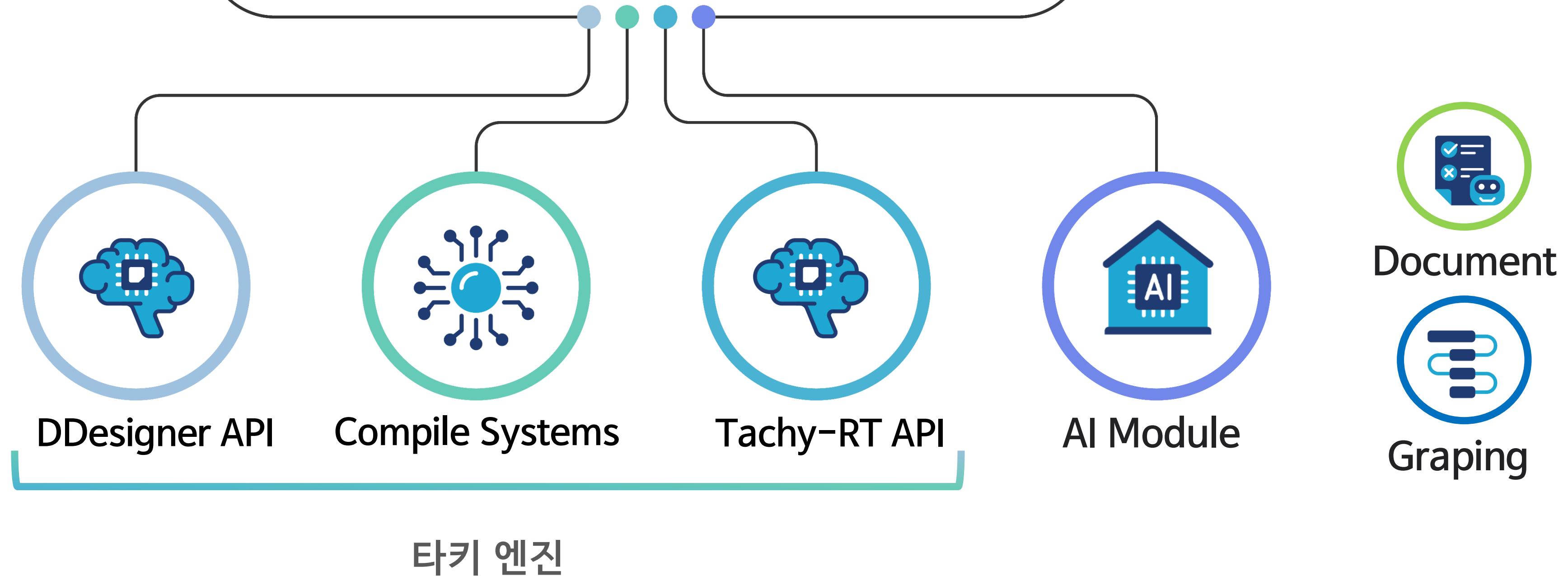




Deeper-I SDK

가이드북

Deeper-I SDK 구성



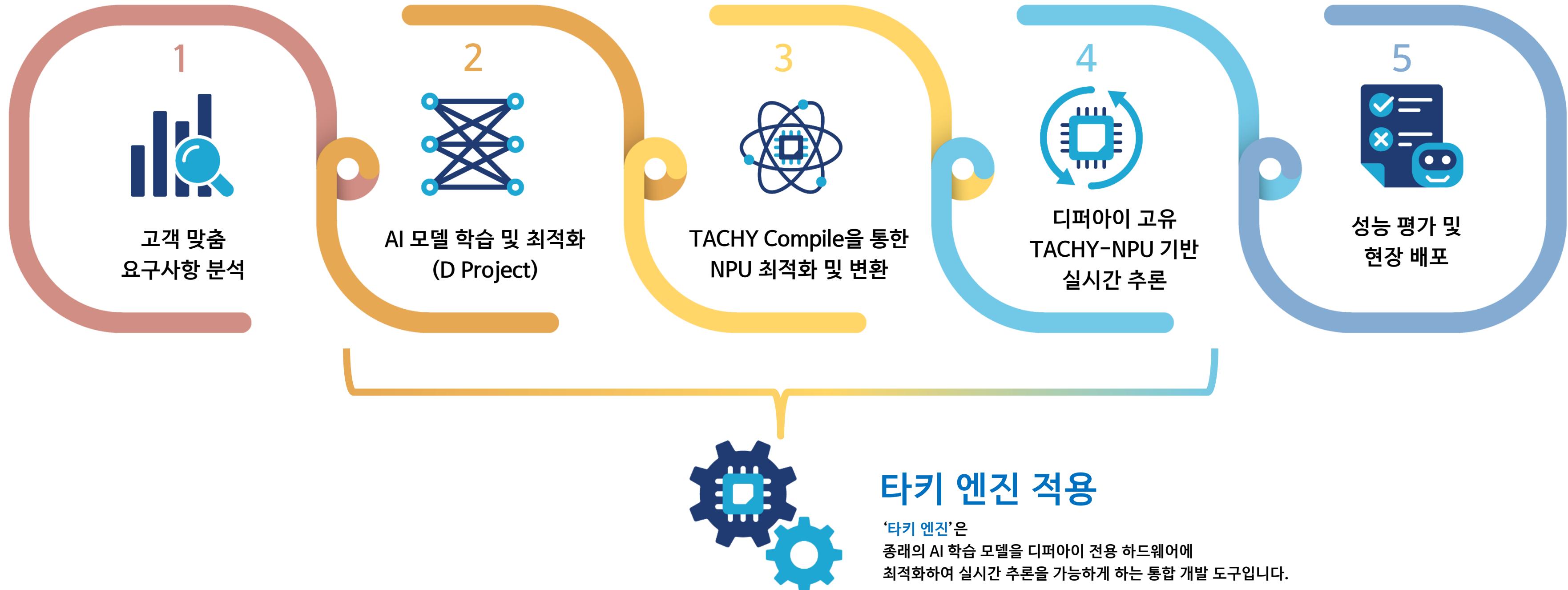
Deeper-I 는 ‘**인공지능 모델 최적화 – 컴파일러 – 디퍼아이 고유 NPU 환경 기반의 실시간 추론**’까지 AI 솔루션 개발 전 과정을 독자적인 기술로 구현하고 있으며, 스마트 모빌리티, 스마트 비전, 스마트 팩토리, 스포츠 등, 다양한 일상과 산업 현장에 적용해 독보적인 가치를 만들어내고 있습니다.



Deeper-I AI 솔루션 개발 프로세스

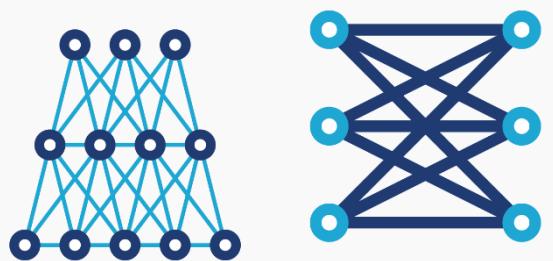
완전한 End-to-End 개발 프로세스

모델 최적화(Training) - 컴파일(Compile) - 추론(Inference)까지 End-to-End 개발 프로세스를 지원하며,
Deeper-I 고유 양자화 구조적 경량화 기술(XWN 기반)을 통해 고성능, 저전력, 고속처리 및 다중 보드 환경을 지원합니다.



Deeper-I 타키엔진의 주요 기능 및 흐름

1. Training (모델 최적화)



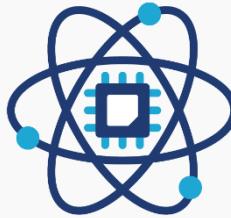
PyTorch(model.pth/model.pt) 또는 TensorFlow(model.h5,saved_model.pb)로 학습된 모델을 DDesigner API를 통해 디퍼아이 하드웨어 전용 파라미터로 변환하는 과정을 거치게 됩니다.

[XWN 구조적 경량화 적용]



서버

2. Compile (모델 컴파일)



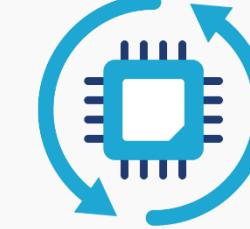
Tachy Compile는 Training 단계에서 회신받은 최적화된 모델 포맷을 ONNX로 변환하고 디퍼아이 전용 확장 파일인 .tachyrt로 변환하는 역할을 수행합니다.

[ONNX → 디퍼아이 전용 포맷 .tachyrt로
컴파일
→ tachyrt 파일 생성 및 하드웨어 적용]



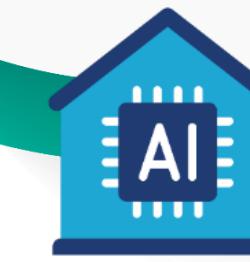
서버

3. Inference (실행 및 추론)



Tachy Compile 단계에서 최종 포맷 변환된 .tachyrt (디퍼아이 전용 확장 파일)을 받게 되면 User는 Tachy-RT API를 통해 ./run명령으로 디퍼아이 하드웨어를 원격 제어와 추론을 동시에 실행하게 됩니다.

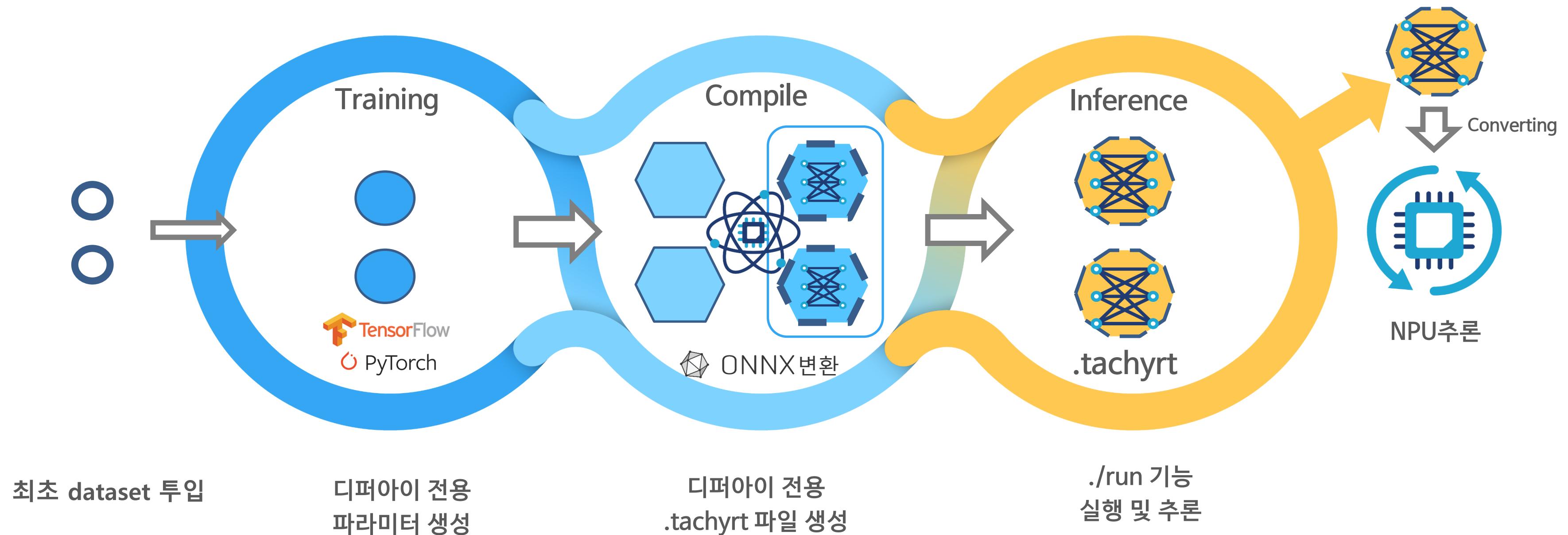
[RPC 기반으로 외부에서
원격 제어 및 추론 수행]



H/W
(NPU)

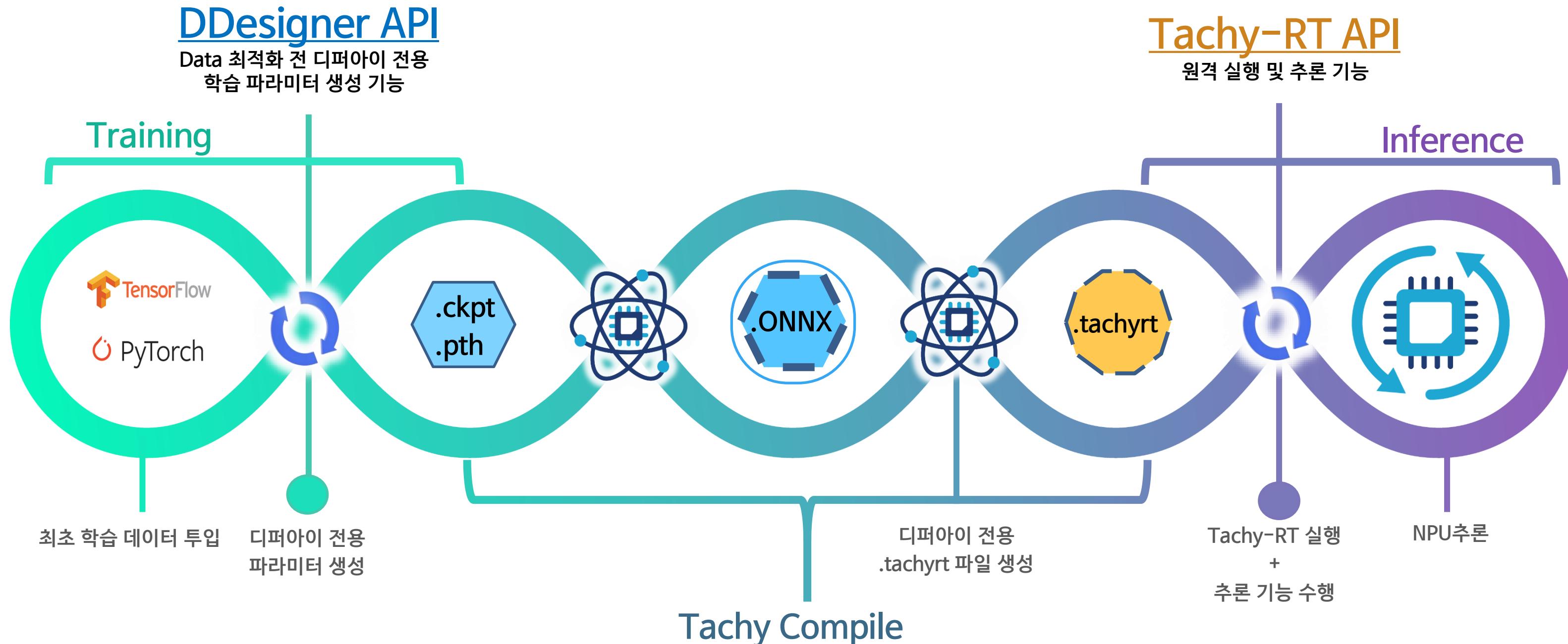
Deeper-I 전용 Data 최적화 변환 흐름

타키엔진을 통한 Data 최적화 변환 흐름은 아래와 같이 이루어집니다.



Data 최적화와 추론 기능을 지원하는 API 적용 흐름

DDesigner API와 Tachy-RT API의 실행 적용은 아래와 같은 흐름으로 data 변환 및 적용 과정이 이루어집니다.



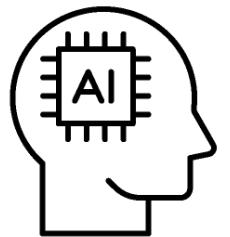
[DDesigner API 가이드북 PDF 링크](#)

[Tachy-Compile 가이드북 PDF 링크](#)

[Tachy-RT API 가이드북 PDF 링크](#)

Deeper-I NPU Board 제품

Deeper-I의 SDK 가이드가 적용되는 제품은 아래와 같습니다.



NPU: BS402

[Tachy-Shield](#)
[브로셔 링크](#)



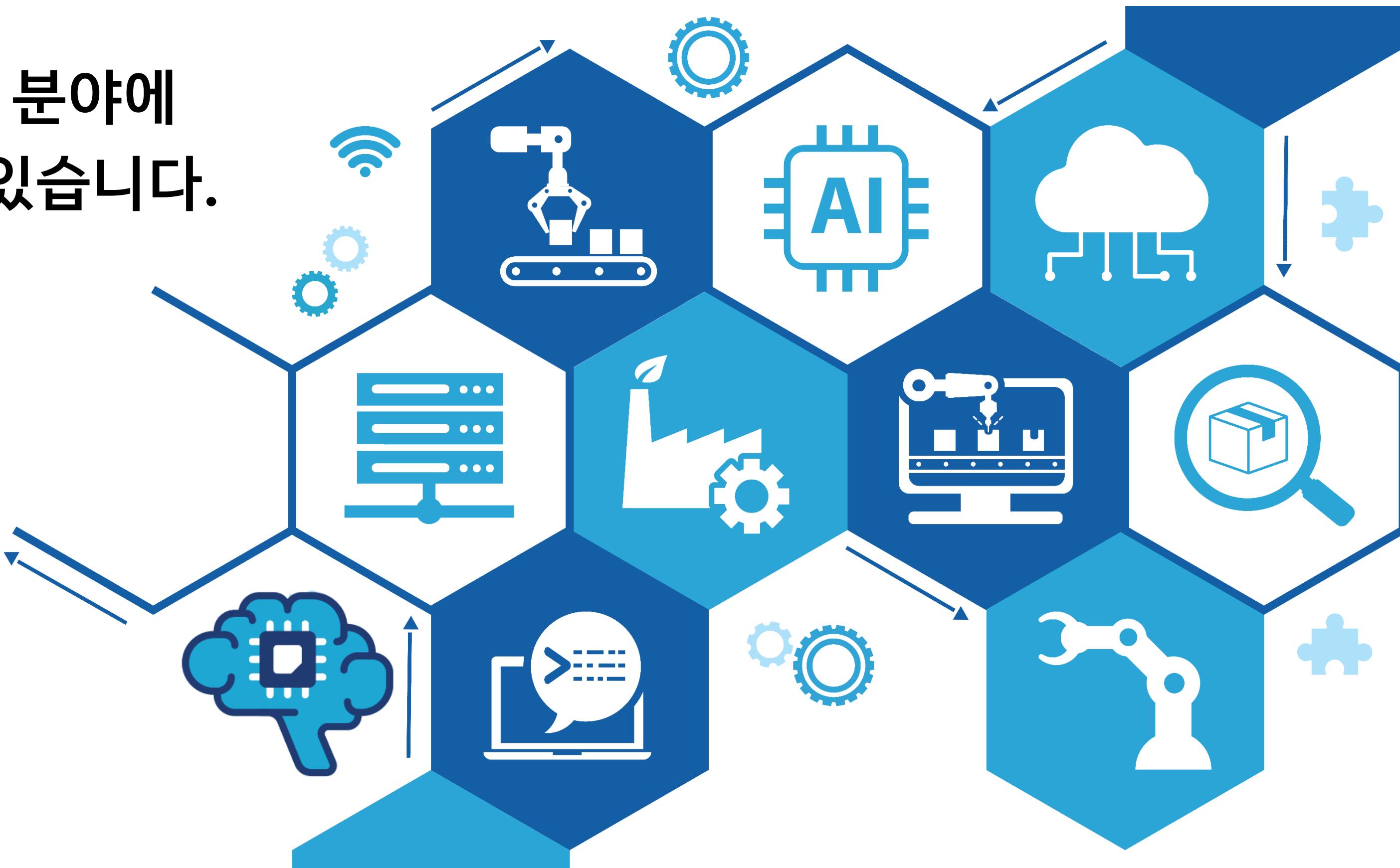
[TAU_Sodimm](#)
[브로셔 링크](#)



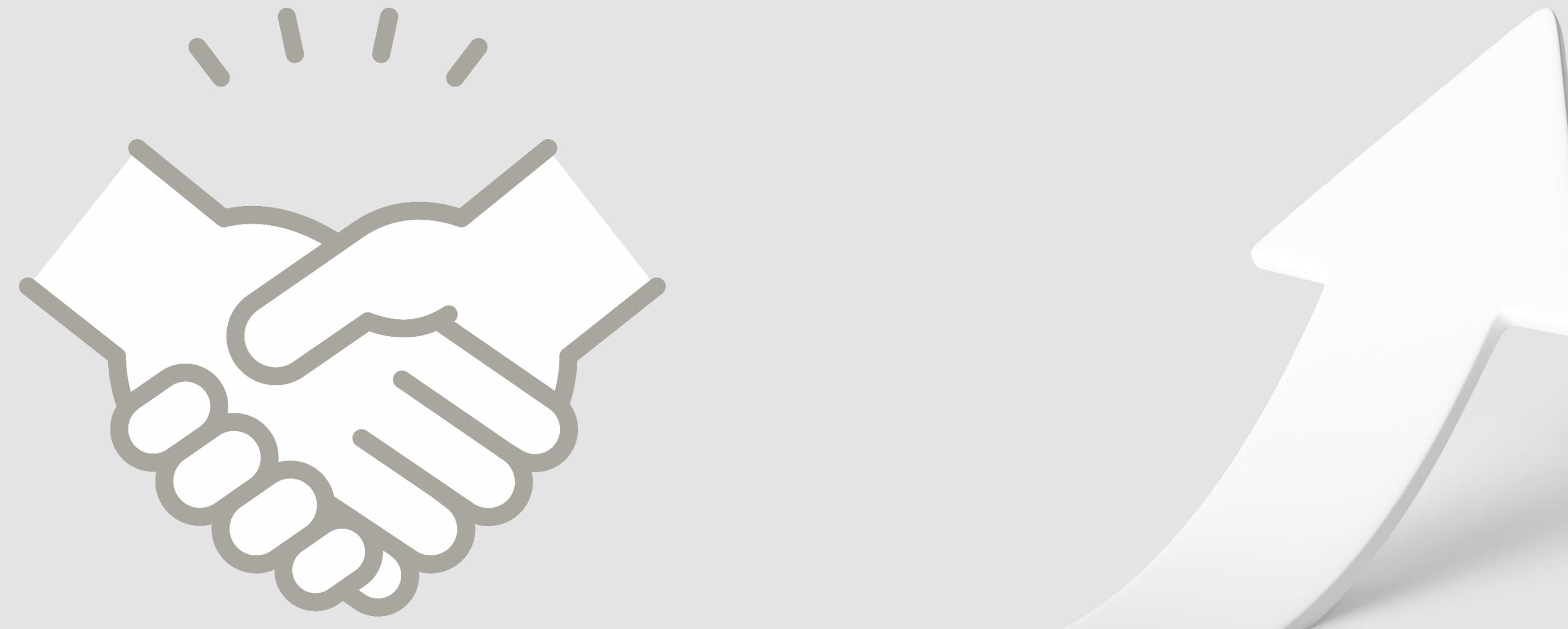
[TAU_M.2](#)
[브로셔 링크](#)

디퍼아이의 타키엔진은

현재 여러 산업 분야에
실적용시키고 있습니다.



Deeper-I 만의 독자적 타키엔진은
여러 기업들과의 협업을 통해 그 기술력이 이미 증명되었습니다.



귀사의 성공과 Deeper-I 는 함께합니다 .

**Innovation in edge technology
creates a smarter tomorrow.**

고객 지원 (Customer Support)
• 문의처: partner@deeper-i.ai
• 웹사이트: <https://www.deeper-i.ai>