



Implementing and Scaling AI with Utilities

Marc Spieler
Sr. Managing Director

The AI Revolution is Here and Accelerating

Perception AI

Speech Recognition
Deep RECSYS
Medical Imaging



Generative AI

Digital Marketing
Content Creation



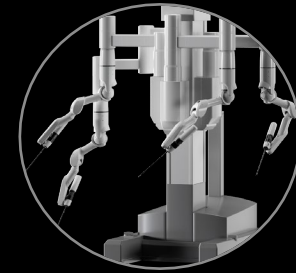
Agentic AI

Coding Assistant
Customer Service

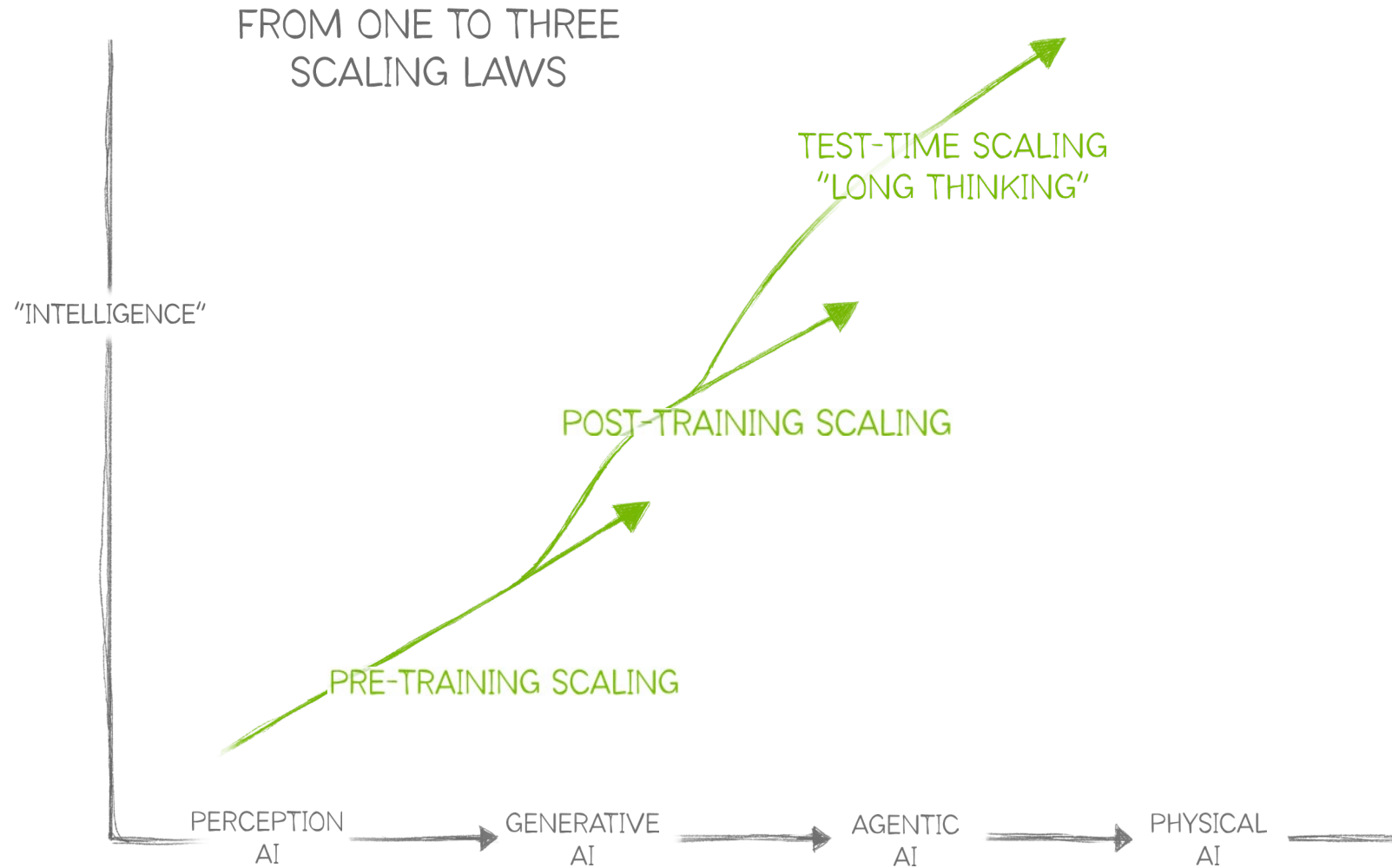


Embodied AI

Self-Driving Cars
General Robotics

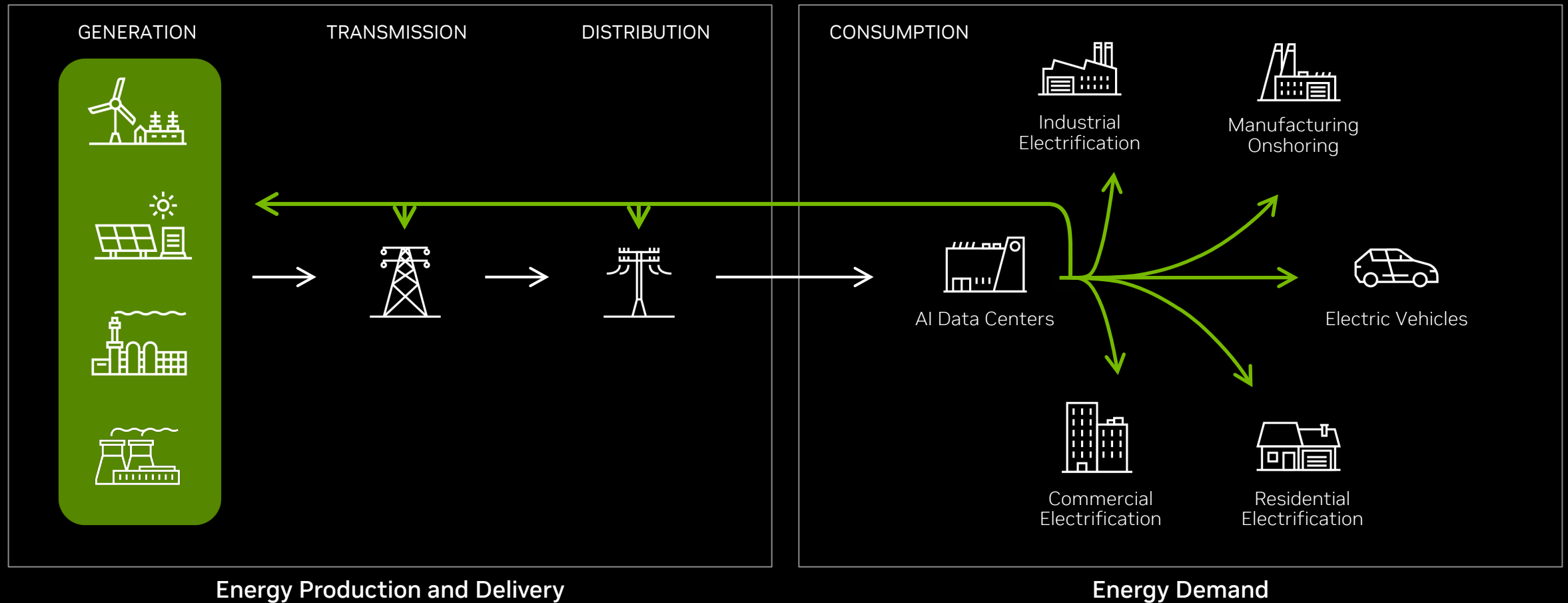


AI Scaling Laws Drive Exponential Demand for Compute



AI Is Key to Solving Longstanding and Emerging Energy Challenges

Optimize Energy Production, Delivery, and Consumption to Meeting Growing Demand



SCE PROJECT ORCA BACKGROUND

Generative AI Assistant for the Command Center and Field called **Orca**. Initially, this Gen AI RAG Knowledge Assistant will benefit the Grid Telecom Command Center. Ultimately, **Orca** will **be the AI Assistant used across the business**.

BACK- GROUND

- Digital Grid Services (DGS) vision for Orca
- NVIDIA & WWT partner to develop use case for SCE

Business Challenges

- Orders of magnitude more telecom devices to manage & maintain - 30% increase per year projected
- Many new device types added to all legacy devices - currently 60 vendors
- Can't increase staff
- Staff turnover / reduce training time
- Need to reduce resolution time - data network more critical to more dynamic electric grid

GOALS

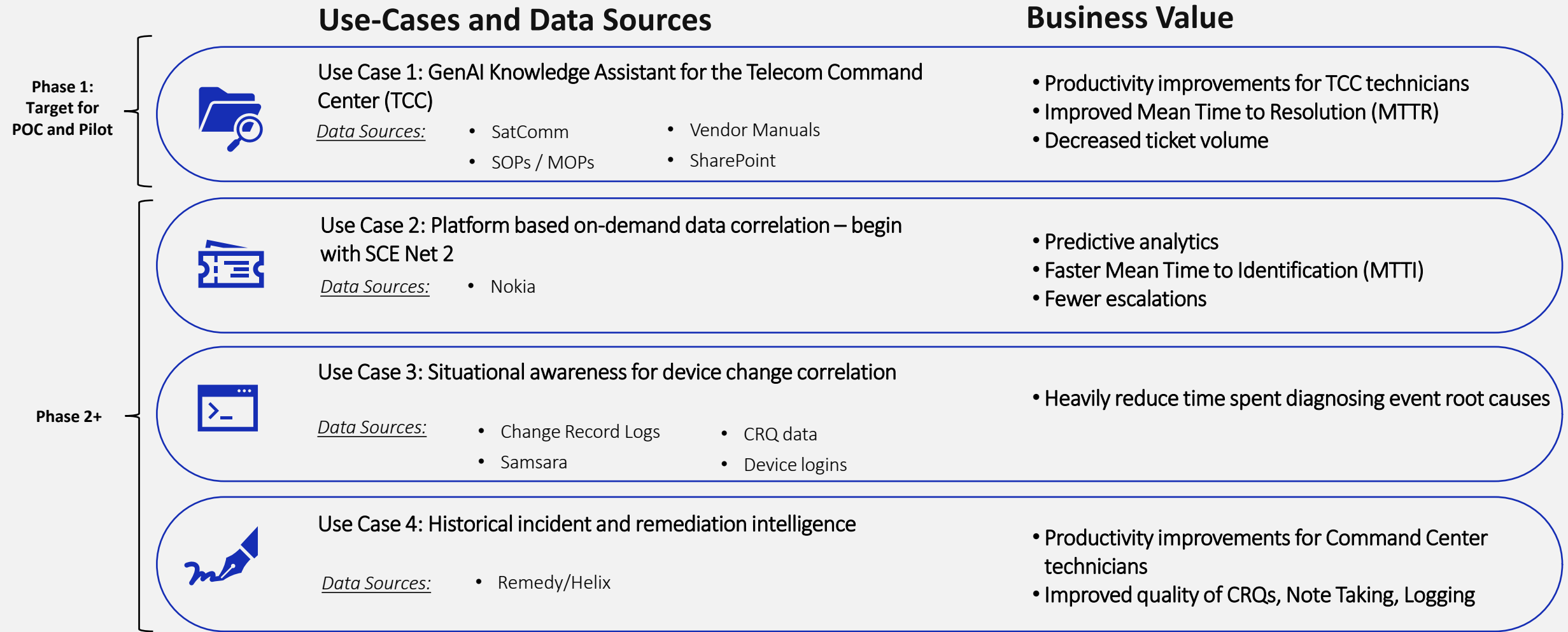
Expected Benefits

- Improve command center efficiency, gain deeper insights to proactively avoid incidents

Key Metrics

- Improve training & productivity for Command Center Agents
- MTTR/MTTI reduction
- Decrease in ticket volume
- Fewer Escalations
- Improve quality of CRQs, Note Taking, Logging

A well-defined set of use cases will inform Orca design and roadmap



SCE PROJECT ORCA NEXT STEPS

*The goal of ORCA is to build out **a base set of high-performance architecture** to run **AI Workloads for the Digital Grid Services Organization**. Next, we are working to scale this architecture to **meet the needs of SCE's AI Strategy for the Grid**.*

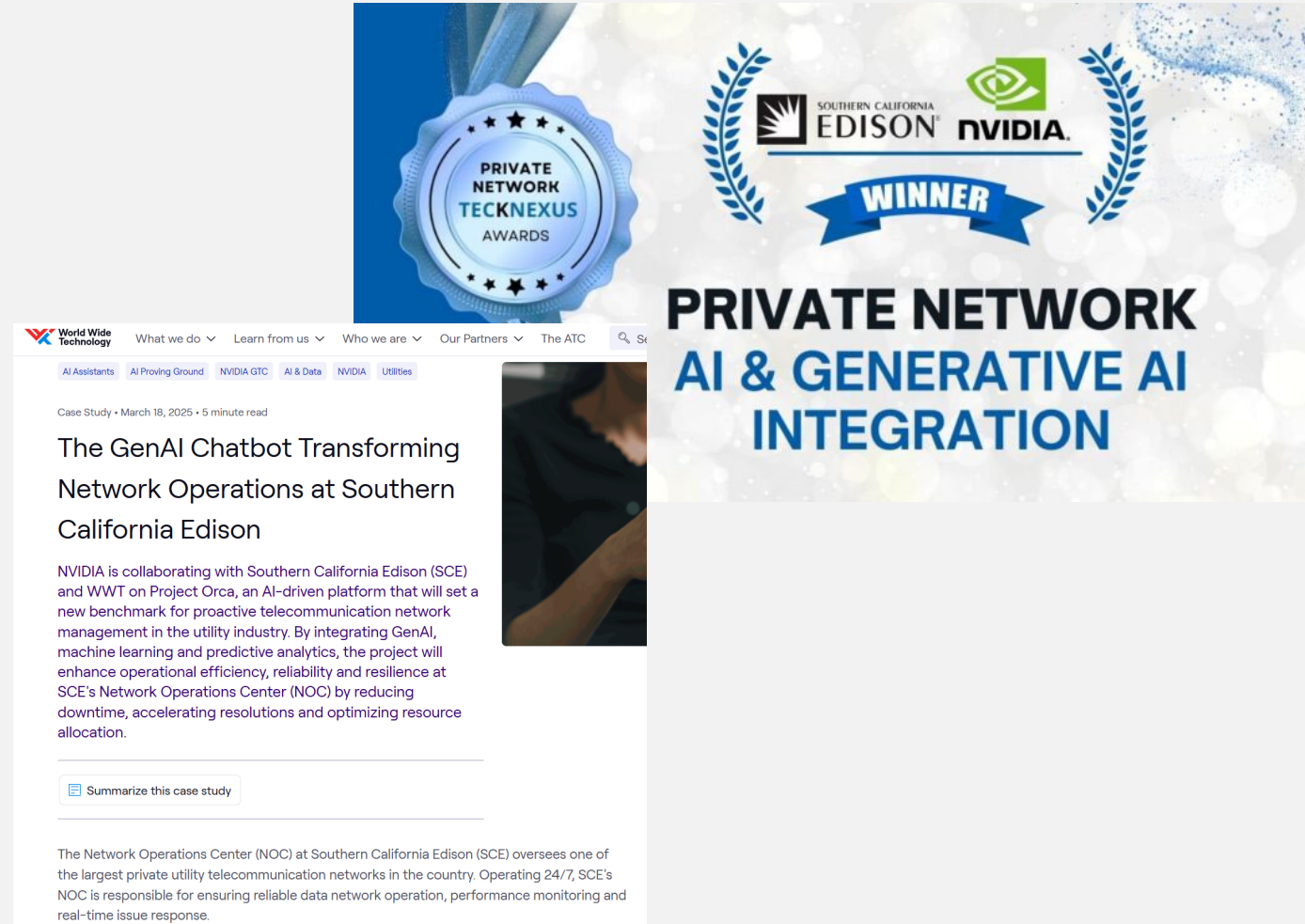
NEXT STEPS

Deliver

- 3 X HPE DL380 w/2 NVIDIA L40S GPUs
- Net App A90 – Hi Performance Storage
- High Performance Switching (Cisco)
- NVIDIA AI Enterprise + NIM
- NVIDIA Retrieval QA E5 Embedding v5 NIM & Llama 3.x 8B Instruct NIM

Next Use Cases

1. Platform based on demand data correlation beginning with SCE Net 2 (Nokia)
2. Greater situational awareness for device change correlation



The collage features three main elements: a blue award medal with the text 'PRIVATE NETWORK TECKNEXUS AWARDS', a 'WINNER' banner for Southern California Edison and NVIDIA, and a screenshot of a World Wide Technology case study. The case study is titled 'The GenAI Chatbot Transforming Network Operations at Southern California Edison' and describes a collaboration between NVIDIA, SCE, and WWT on Project Orca, an AI-driven platform for proactive telecommunication network management. It highlights the use of GenAI, machine learning, and predictive analytics to enhance operational efficiency and reduce downtime at SCE's Network Operations Center (NOC).

<https://www.wwt.com/case-study/genai-chatbot-transforming-network-ops-at-sce>





AI ENABLED NUCLEAR POWER PLANT - PG&E, ATOMIC CANYON, & NVIDIA

Use Case

- Generative AI search for nuclear plant operations.
- Fast, accurate retrieval of billions of documents.
- Supports engineering, regulatory, and operational tasks.

Challenges

- Billions of unstructured documents across many systems.
- Slow, difficult search—delaying key decisions and compliance.
- Risk of incomplete or outdated information.

Solution

- Atomic Canyon's AI-powered, nuclear-specific search platform.
- Unified, intelligent search across all data sources.
- Advanced OCR, classification, and nuclear-specialized AI models (FERMI).

NVIDIA Solution Stack

- Hardware: NVIDIA H100 GPU
- Software: NVIDIA Nemo

Outcome

- Search time cut from hours/days to seconds/minutes.
- 98%+ search accuracy reduces errors and rework.
- Faster, more reliable compliance and safety.



PHOTOVOLTAIC POWER PREDICTION — GCL, PKU & NVIDIA

Use Case

- Enhancing photovoltaic (PV) power prediction with high-resolution weather forecasting from NVIDIA Earth-2

Challenges

- PV power generation is volatile due to the influence of weather and environment, which affects the stability of the power grid
- PV power generation enterprises spend a large amount of money on meteorological and power prediction services every year, but they still face heavy fines due to inaccurate power prediction

Solution

- Using Earth-2 and PhysicsNeMo, NVIDIA developed an AI pipeline which can forecast solar irradiation at a much lower cost
- GCL and PKU developed a time-series model which improves the accuracy of PV power prediction with solar irradiation forecast data

NVIDIA Solution Stack

- Hardware: NVIDIA Ampere & Hopper Tensor Core GPUs
- Software: NVIDIA Earth-2, NVIDIA PhysicsNeMo, NVIDIA Omniverse

Outcome

- 10% increase in PV power prediction accuracy can reduce fines and save the industry billions of Yuan annually in China
- 5,000X speed-up and 10,000X increase in energy efficiency for solar irradiation forecast compared to NWP methods, and save PV power generation enterprises' expense in weather service

SMART SUBSTATIONS - LARGE GRID OPERATOR

Use Case

- Substation Autonomous Operation & Maintenance

Challenges

- Traditional manual inspection for Substation with the heavy manpower workload and more mistake

Solution

- Engage and support local ISV partners and start-ups to integrate NVIDIA technology into the smart substation solution
- Using AI/IVA technology to do substation equipment inspection like meter reading, defect detection, switch state, workers' face recognition, asset health, helmets, behaviors, etc.

NVIDIA Solution Stack

- Hardware: NVIDIA T4/A2 Tensor Core GPUs
- Software: NVIDIA Metropolis, TAO, DeepStream, TensorRT, etc

Outcome

- Total 42 algorithms and inference time less than 400ms
- The frequency of substation inspections has been increased by 43%
- AI-enabled inspections detect over 1,000 defects annually, enhancing the reliability and safety of grid equipment

