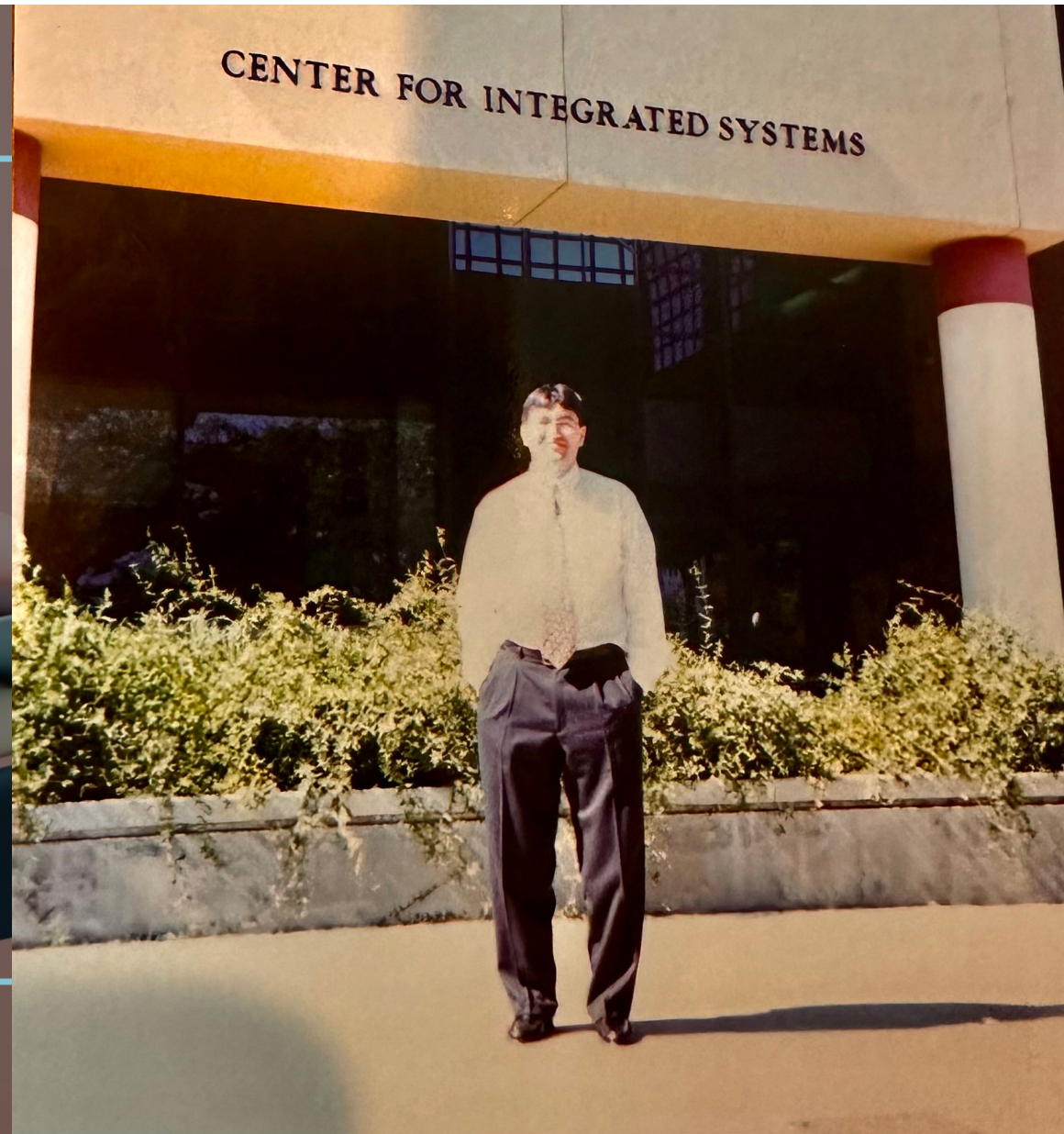


A JOURNEY FROM EDA TO THE TIMES OF "THE MYTHICAL LLM- MONTH"

Claudionor N. Coelho Jr, PhD/MBA
Chief AI Officer
Zscaler





CLAUDIONOR COELHO JR

Chief AI Officer at Zscaler

Invited Professor for ML/DL/NLP at Santa Clara University

Chief AI Officer / SVP of Engineering at Advantest

VP/Fellow of AI at Palo Alto Networks

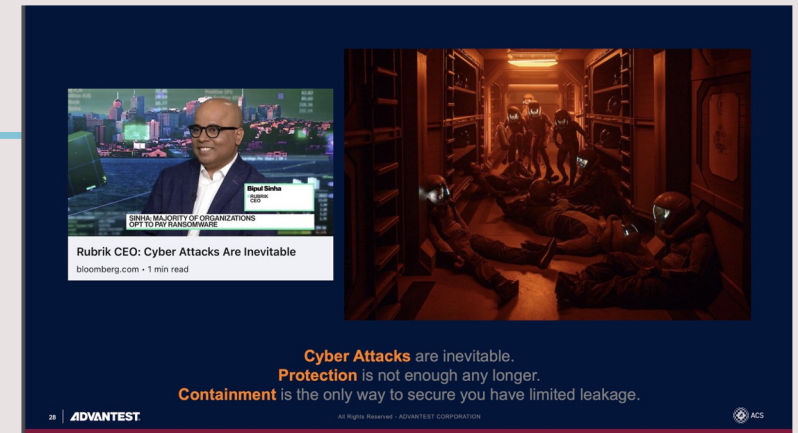
Deep Learning Researcher at Google (creator of Qkeras, used by CERN/STM)

VP at Synopsys

SVP of Engineering at Jasper Design Automation / Cadence Design Systems

Professor DCC/UFMG - Brazil

PhD from Stanford, MBA from IBMEC/Brazil, MS/BS from UFMG/Brazil



THE PHD TIMES STILL RINGS A BELL TODAY

- First met Nanni in SBMICRO 1989 when he was talking about Hercules HLS
- When I joined Nanni's group, he asked me to solve the problem of High-Level Technology Mapping, could not solve this problem at that time because needed a killer application
 - In 2017, when I was the VP of Engineering at NVXL Technology, working on inference engines for DL workloads, I contacted him, and I said I solved this problem elegantly for Deep Learning problems
- Ended up doing PhD on synthesizing dynamic controllers in High-Level Synthesis using constraints
 - Constraining applications and dynamic schedules became important again when creating AI Accelerators
 - Turns out this is extremely important in formal verification

HIGHLIGHTS OF MY CAREER

1990-2000

- Did formal verification of microcode of 386 processor at Integrated Information Technology (IIT) using the same techniques I was using for my PhD
- Took ACM Computer Programming Contest to the Southern Hemisphere of the Planet
- One of the creators of IEEE Open Verification Library as a strategy to promote formal verification
- Sold Verplex to Cadence

2000-2020

- Took JDA when investors were shutting it down and sold it to Cadence when it reach 80% of the market share
- Started looking at ML for EDA
 - First ML for EDA by fine tuning SAT solvers
 - Looked at word embeddings for debugging emulation traces
- Worked on AI acceleration engines at NVXL before tensorflow and pytorch
- Work at Google on model compression and quantization was used at CERN to search for sub-atomic particles (cover page of Nature Machine Intelligence)

2020-now

- Worked on neurosymbolic techniques to debug logs looking for patterns in zero-day cybersecurity attacks
- Build secure private cloud for semiconductor industry to run ML workloads
- Released fourth Copilots
- Started working on how to build almost hallucination-free Autopilots using Multi-Agent systems

STARTED BUILDING COPILOTS WHEN NO ONE WAS TALKING ABOUT IT...

- **2016: Synopsys for ZeBu (under Rohit Vora)**
 - Interpreting user's intents and using tools to scan logs and suggest courses of actions
- **2021-2022: Palo Alto Networks (before ChatGPT, under Karan Gupta)**
 - Solving customer problems by scanning tickets for common answers to problems
- **2023: Advantest**
 - RAG pipelines to answer questions about complex products
- **2024: Zscaler**
 - Multi-agent systems, avoiding hallucinations and the pitfalls of "*The Mythical LLM-Month*"

INFLUENCE FROM NANNI IN MY CAREER IS EVIDENT

- Used Nanni's thesis and book on synthesis several times in my life
 - When creating ActiveProp @ JDA (generation of automatic properties by figuring out what's missing), used Espresso to optimize tables of the complement of coverage vectors
 - In QKeras @ Google for model compression and quantization, one of the modules inside QKeras used Nanni's PhD thesis to optimize symbolic deep-quantized trees to efficiently implement 1-3 bit neural networks
- Be humble
- Solve the right problems at the right time
 - Look at problems where no one is looking
- Innovate and create value
- Strategy with innovation
- During your career, you need to reinvent yourself several times...

Thanks a lot, Nanni, for a lifetime of learning