

THE COOKIE-LESS CHALLENGE : HOW AI HELPS MARKETERS ADAPT AND WIN



Issue 39
December 2025

ROS + PLC THE BLUEPRINT FOR SEAMLESS **INDUSTRIAL INTEGRATION**

ANTROMEDA META'S NEW GALAXY FOR SMARTER ADVERTISING

Contributors



Antromeda

Meta's New Galaxy for Smarter Advertising

Harish Krishnan

Digital Marketing Analyst, Trivandrum

Page No

02



The Cookie-Less Challenge

How AI Helps Marketers Adapt and Win

Ankita Sharma

Digital Marketing Executive, Pune

Page No

07



ROS + PLC

The Blueprint for Seamless Industrial Integration

Paul Manuel

Project Engineer, Mysore

Page No

12



The Thinking Machine

Neural Networks at Work

Jeyasooriya.J

IT Analyst, Madurai

Page No

17



One Language, Every Layer

The JavaScript Advantage

Arya Saju

IT ENGINEER, Kochi

Page No

24



Lost on the Web

The Surprising Life of 404 Pages

Abisekha Melfin

IT Engineer, Triunelveli

Page No

29

Contributors



Automation to Collaboration The Real Meaning of Industry 5.0

Navya Rajan

Junior Project Engineer, Kollam

Page No

35



Beyond ChatGPT The Evolution from Transformers to RAG

Muhammed Farhan

IT Engineer, Attingal

Page No

40



OPC UA Powering the Connected Factory Revolution

Malleshwaran

Teach lead, Coimbatore

Page No

47



Beyond Text The New Era of Image, Voice and Video Intelligence

Chandana P

Technical Lead, Mysore

Page No

53



TinyML The Science of Machine Learning on Microcontrollers

Karthikeyan

Embedded Engineer, Anna Nagar

Page No

59

About Us

Our journey began in 2008 with the establishment of our first office in Kochi, where our operations team initiated industrial automation projects. Just a year later, we launched our first training center in Calicut. With an unwavering commitment to quality, we quickly gained the trust of students not only across India but also from countries in Africa.

Over time, our presence expanded into Nigeria, Qatar, the UAE, Kenya, and the Kingdom of Saudi Arabia. By 2025 IPCS global Operates 33+ Centres worldwide, earning a reputation as one of the most trusted and respected providers of core technical training—offering programs designed to be truly future-ready.

Each of our programs is carefully crafted to align with global industry trends, employment opportunities, and evolving market needs. Our current offerings include:

Industrial Automation

Building Automation
Technology

Digital Marketing

Python & Data Science

Embedded & IoT

Artificial Intelligence

Software Testing

Key highlights of our training include:

- 100% live and interactive sessions
- Government and internationally recognized certifications
- Comprehensive placement support

Looking ahead, we are on track to expand our network to 50 centers by 2025, reflecting our vision for growth and commitment to excellence. We welcome passionate entrepreneurs and visionary investors to join us—whether as franchisees under our proven model or as strategic partners driving our global expansion. Together, we can build opportunities, shape careers, and create lasting impact in communities around the world.

At IPCS, our mission is to equip students with the skills of tomorrow by staying aligned with emerging technologies, while upholding the highest ethical standards. We cultivate a culture of teamwork, professionalism, and mutual respect, ensuring student success and client satisfaction across all domains.

In today's digital age, technology is the backbone of growth and innovation. Embracing this reality, we continue to deliver excellence across the globe.

To further our vision, Team IPCS proudly presents Iziar—a magazine dedicated to exploring technological insights, industry trends, startups, and digital culture. Iziar aims to make technology accessible, engaging, and inspiring, keeping readers informed about the innovations shaping our future.

Technology is like air—indispensable to life. Step into the world of Iziar and experience the future.

Visit us at www.ipcsglobal.com

“TIME AND TECHNOLOGY WAIT FOR NONE”

Antromeda

Meta's New Galaxy for Smarter Advertising



Harish Krishnan
Digital Marketing Analyst
Trivandrum

I am Harish Krishnan from Pathanamthitta, working as a Digital Marketing Trainer & Analyst, as well as an Internet Marketing Analyst. I specialize in developing and optimizing effective online marketing strategies that drive traffic, boost sales, and enhance overall brand visibility. My role involves a strong focus on data-driven decision-making. I regularly monitor and analyze digital marketing campaigns to evaluate performance, identify growth opportunities, and ensure continuous improvement. By staying updated with the latest trends and tools in the industry, I aim to deliver impactful strategies that support business goals and strengthen digital presence.

Just as the Andromeda Galaxy (M31)—a massive spiral galaxy located 2.5 million light-years away—is on course to merge with the Milky Way, Meta's advertising ecosystem is undergoing its own monumental transformation. Meta's newest upgrade, named Antromeda, marks one of the biggest shifts in digital advertising since the introduction of machine-learning-powered targeting.

Before understanding why Meta chose the name Antromeda, it's important to look back at its previous architecture.



From Galactica to Antromeda: Why the New Name?

Meta's earlier ad delivery system, Galactica, collected user signals across Facebook, Instagram, WhatsApp, and other Meta apps. This included:

- ♦ Interests
- ♦ Demographics
- ♦ Behaviors
- ♦ Actions
- ♦ Custom audience signals

Galactica was the core engine that powered how ads were retrieved and matched to audiences.

However, in 2025, Meta rebuilt this architecture from the ground up using advanced AI retrieval models. Engineers described this upgrade as an entirely new galaxy of ad delivery—smarter, larger, more connected, and far more adaptive than its predecessor.

That's why Meta named the new system Antromeda—symbolizing a massive leap forward, similar to discovering a new galaxy.

What's Changing for Advertisers?

For years, advertisers followed the standard approach:

- ♦ 1 campaign
- ♦ 2–3 ad sets with different interests/behaviors
- ♦ 1–3 creatives in each ad set

But over the past few weeks, many advertisers started noticing something unusual:

- ♦ Higher CPR (Cost Per Result)
- ♦ Slower optimization
- ♦ Fewer conversions
- ♦ Unstable campaign performance

This is not accidental—Meta has been gradually shifting its system since 2024, and Antromeda is now fully taking over.

The New Reality: Creative Diversity > Detailed Targeting

Meta has officially confirmed the direction:

Creative diversity is the new key to performance.

This means your ads must now include variety—different angles, lengths, tones, and formats.

Examples include:

- ♦ Short videos
- ♦ Long videos
- ♦ Static posters
- ♦ Carousels
- ♦ Testimonial ads
- ♦ Problem–solution videos
- ♦ Before–after demonstrations
- ♦ Persona–based creatives

Why?

Because detailed targeting is no longer the primary driver.

With Antromeda, creative targeting becomes more important than interest or demographic targeting. The AI will analyze your creatives, understand their context, and automatically find the best audiences—much like content recommendation engines.

Better creatives = better audience discovery = better results.

How to Succeed With the New Meta Strategy

1. Use Broad Targeting

Let Meta's AI find your ideal customers instead of forcing the audience.

2. Add Maximum Creative Variations

At the ad level, include multiple forms of content focusing on:

- ♦ Pain points
- ♦ Product problems
- ♦ Exclusive offers
- ♦ Unique angles
- ♦ Testimonials
- ♦ Different audience personas
- ♦ Before–after transformations

The more creative diversity you provide, the faster Meta identifies high-performing combinations.

3. Focus on High-Quality Storytelling

Engaging, authentic content will outperform hyper-targeted ads.



Major Additional Updates in Meta Ads (2025)

Meta's Antromeda update is not just about creative diversity—several new features have been rolled out globally and in India.

1. Campaign Consolidation

Advertisers can now combine Awareness, Consideration, and Conversion stages inside a single campaign.

This is especially useful in funnel marketing.

Example:

For a digital marketing course ad, both cold (don't know you) and warm (already aware) audiences can be targeted in one consolidated funnel.

This increases efficiency and helps Meta optimize the entire customer journey.

2. Income-Based Targeting Now Available in India

Previously available only in the U.S., this feature now allows Indian advertisers to target users based on income brackets.

This helps brands:

- ◆ Improve product-audience fit
- ◆ Reduce wasted ad spend
- ◆ Reach more qualified leads

3. Collaboration Ads with Swiggy, Zomato & Blinkit

Meta has partnered with major delivery platforms to integrate real-time purchase behavior into ad optimization.

This means:

- ◆ Better credibility
- ◆ Higher-quality leads
- ◆ More accurate targeting
- ◆ Increased conversions

For brands in food, retail, grocery, and D2C sectors, this is a game changer.

4. Add/Remove Ad Sets Without Restarting Learning Phase

This long-awaited update ensures that the learning phase is not reset when:

- ◆ Adding new ad sets
- ◆ Removing underperforming ones
- ◆ Making structural edits

This preserves optimization data and ensures stable campaign performance.

5. WhatsApp Status Ads

A major expansion of Meta's ad ecosystem:

- ◆ Ads will appear inside WhatsApp Status
- ◆ Huge potential in India due to massive WhatsApp usage
- ◆ More personalized, high-visibility format

Great for lead generation, branding, and promotions.

6. Instagram Skippable & Non-Skippable Ads

Instagram is introducing a dual-format ad model:

- ◆ Skippable ads for user-friendly engagement
- ◆ Non-skippable ads to guarantee message delivery

This brings Instagram closer to YouTube-style advertising and gives brands more control over visibility.



The Real Competition Begins Now

Just like Andromeda represents massive cosmic change, Meta's Antromeda represents a new age in digital advertising—one powered by AI, creativity, and broad audience intelligence.

The message is clear:

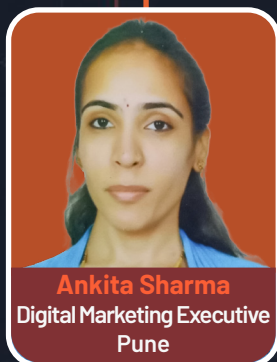
The era of detailed targeting is fading.

The era of creative-led AI targeting has begun.

Those who adapt—by embracing creative diversity, focusing on high-quality content, and leveraging Meta's new tools—will lead the next wave of digital marketing success.

The Cookie-Less Challenge

How AI Helps Marketers Adapt and Win



I am Ankita Sharma, a Digital Marketing Executive at IPCS Global, Pune. I specialize in developing data-driven marketing strategies, managing end-to-end campaigns, and strengthening digital brand presence. With over three years of hands-on experience in SEO, social media management, content marketing, and multi-channel campaign execution, I have consistently delivered measurable results across diverse marketing initiatives. Driven by curiosity and continuous growth, I am committed to staying updated with emerging marketing technologies and tools. My goal is to optimize campaign performance, enhance digital outreach, and contribute meaningfully to my team's overall success.

The digital marketing landscape is undergoing its most significant transformation in over a decade. With the phase-out of third-party cookies, strengthened global privacy laws, and rising consumer awareness about data rights, traditional targeting models are becoming obsolete. According to a Deloitte survey, 78% of marketers believe the death of third-party cookies will disrupt their advertising strategies, while 67% expect increased reliance on AI-driven insights.

The challenge is clear: How do marketers deliver personalisation without compromising privacy? The answer lies at the intersection of first-party data, Artificial Intelligence (AI), and ethical marketing frameworks. Together, they are redefining the future of digital engagement.

The Cookie Collapse: Why It Matters More Than Ever

For nearly 20 years, third-party cookies formed the backbone of digital advertising—fueling retargeting, behavioral profiling, and cross-site tracking. But with Google Chrome, Safari, and Firefox removing support for them, brands are entering uncharted territory.

Why the shift?

- **Privacy regulations** : Laws such as GDPR in Europe, CCPA in California, and India's DPDP Act strengthen individual rights over personal data.
- **Consumer demand**: A Cisco study found that 81% of consumers want more control over how their data is used.
- **Platform policies**: Tech giants now prioritize privacy and transparency, limiting access to granular user tracking.

In this environment, businesses can no longer rely on invisible tracking. Instead, they must pivot toward trusted, consensual, and transparent data practices—with AI playing a crucial enabling role.

First-Party Data: The New Competitive Advantage

As cookies disappear, first-party data becomes the new digital currency. This includes all information a company collects directly from customers through interactions on owned platforms.

Examples of first-party data include:

- ▶ Website engagement metrics
- ▶ Mobile app activity
- ▶ CRM records and purchase history
- ▶ Loyalty program participation
- ▶ Email opt-ins and survey responses

Unlike third-party data, first-party data is:

- ▶ Permission-based
- ▶ More accurate
- ▶ Highly relevant
- ▶ Ethically sourced

Companies that leverage it well gain an edge. McKinsey research shows that brands excelling at first-party data activation achieve 1.5x revenue uplift compared to traditional models.



How AI helps unlock its value

Raw data alone is not enough. AI systems, especially Customer Data Platforms (CDPs), can:

- Clean and unify siloed datasets
- Identify hidden behavioural patterns
- Predict customer needs
- Trigger personalised content in real time

Tools like Salesforce Einstein, HubSpot AI, and Zoho CRM Plus transform first-party data into precise insights—helping marketers build more meaningful, privacy-compliant engagement strategies.

Predictive AI: Understanding Users Without Tracking Them

Modern AI models can “learn” from historical actions and infer future behaviour without relying on third-party cookies. This approach—predictive analytics—is becoming central to marketing.

Key functions enabled by predictive AI:

- ▶ Conversion forecasting
- ▶ Customer lifetime value estimation
- ▶ Churn prediction
- ▶ Product recommendation modelling
- ▶ Optimal timing and channel identification

For example, platforms like Google Cloud AI, Adobe Sensei, and Pecan AI analyse patterns across millions of interactions to predict:

- ▶ Who is most likely to make a purchase
- ▶ Which users may disengage
- ▶ What content format performs best under specific conditions

This ensures accuracy in targeting without violating user privacy.

Contextual Targeting 2.0: Smarter, Safer Advertising

Contextual advertising—once considered old-fashioned—has made a remarkable comeback, evolving into a sophisticated, AI-powered strategy.

How Contextual Targeting 2.0 works

AI systems now analyse:

- ▶ Page topics
- ▶ Sentiment and emotional tone
- ▶ Keyword clusters
- ▶ Visual elements in multimedia content

This allows marketers to place ads that align with the context rather than the identity of a user.

Example:

Instead of targeting “women aged 25–35 who like fashion,”

AI places your ad on content about winter clothing trends.

The benefits are significant:

- ▶ Zero privacy risks
- ▶ Higher user relevancy
- ▶ Stronger brand safety

- ▶ Compliance with global regulations

According to The Drum, contextually relevant ads improve brand recall by up to 40% compared to cookie-based ads.

Generative AI: Personalisation Without User Profiles

Generative AI—powered by large language models (LLMs)—has revolutionised content creation. Instead of building intrusive user profiles, marketers can now personalise messaging using only session-based or contextual data.

Capabilities of generative AI include:

- Creating targeted ad copy in seconds
- Generating multiple creative variations for A/B testing
- Adjusting content in real time based on user behaviour
- Producing hyper-personalised emails or landing pages without storing personal identifiers

Platforms like ChatGPT, Jasper, and Copy.ai enable scalable personalization without breaching privacy norms.

This marks a major shift: personalisation now depends on intelligence, not intrusion.

Consent, Ethics, and Transparency: The New Brand Currency

In today's privacy-first environment, trust is as important as technology.

AI can help organizations maintain compliance by:

- Automating consent management
- Monitoring data usage according to user preferences
- Detecting potential bias in marketing models
- Ensuring transparent communication about data practices

Tools such as Fairly AI perform ethics audits, bias checks, and compliance assessments on machine learning models.

Best practices for brands:

- Provide clear opt-in/opt-out interfaces
- Explain why data is collected and how it benefits users
- Avoid dark patterns and manipulative UX
- Regularly review automated decision-making systems

According to Edelman's Trust Barometer, 88% of consumers say trust is a key factor in deciding which brands they engage with, making ethical AI a strategic priority, not just a legal requirement.

Conclusion: A Smarter, More Ethical Era of Marketing

The cookie-less world is not the end of personalized marketing—it is the beginning of a smarter, more responsible era. AI is emerging as the backbone of this transformation, enabling brands to:

- Understand audiences without intrusive tracking
- Deliver meaningful personalization using first-party and contextual data
- Build systems rooted in transparency, fairness, and consent
- Create human-centered experiences that respect privacy

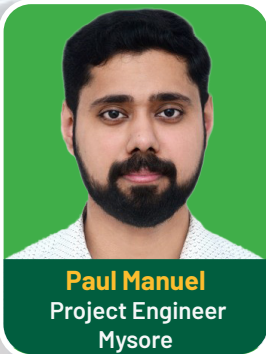
Marketers who embrace this shift will not only stay ahead of regulatory changes but also earn deeper customer trust and long-term loyalty.

In this new landscape, success will no longer rely on tracking people across the internet.

It will come from understanding them—ethically, intelligently, and responsibly.

ROS + PLC

The Blueprint for Seamless Industrial Integration



Paul Manuel
Project Engineer
Mysore

I am an enthusiastic advocate of industrial automation and robotics, dedicated to harnessing intelligent systems to achieve precision and foster innovation. My professional repertoire encompasses the programming, calibration, and troubleshooting of PLCs, the deployment and integration of SCADA systems and possess advanced designing and implementing of human-machine interfaces (HMIs). Furthermore, I am skilled in the conceptualization and assembly of control panels and possess profound knowledge of the Robot Operating System (ROS), including navigation algorithms and mobile robot localization. My focus encompasses diverse facets of industrial environments wherein Programmable Logic Controllers (PLCs) are seamlessly integrated with real-time operational scenarios, complemented by the strategic deployment of SCADA systems. Through the meticulous selection and application of precise communication protocols and modalities, I endeavor to transmute complex industrial challenges into efficacious, real-time solutions.

In today's smart factories, the lines between robotics, automation, and information systems are rapidly disappearing. Robots are becoming more intelligent and autonomous, while industrial control systems continue to deliver the reliability and safety that manufacturing demands. This raises a critical question for modern engineers: How do we make these two worlds—ROS and PLCs—work together seamlessly?

This article explores how the Robot Operating System (ROS) can be integrated with Programmable Logic Controllers (PLCs), the technologies that enable the connection, and practical strategies for implementing the integration in real industrial environments.

The Power of Two Worlds

ROS: Intelligence and Flexibility

ROS (Robot Operating System) is an open-source robotics framework widely used for advanced automation. It provides powerful capabilities such as:

- Path planning and navigation
- Computer vision and AI-driven decision-making
- Sensor fusion and real-time data processing
- Simulation and digital twin development

ROS serves as the robot's "brain," managing high-level decision-making and coordination across sensors and actuators.

PLC: Reliability and Real-Time Control

PLCs remain the backbone of industrial automation, designed for deterministic, real-time control and continuous operation in harsh industrial conditions. They excel at:

- Managing field I/O
- Running ladder logic, FBD, or structured text

- Enforcing safety standards
- Ensuring reliable machine operation with minimal downtime

ROS brings intelligence and adaptability, while the PLC ensures safety and reliability. Together, they form a powerful foundation for modern smart automation.

Why Integrate ROS and PLCs?

The benefits of combining ROS and PLC systems are significant:

- Smart manufacturing: Robots interact smoothly with conveyors, machines, and sensors controlled by PLCs.
- Real-time data exchange: AI-driven algorithms can use PLC data for smarter decisions.
- Collaborative robotics: PLCs handle safety, ROS manages adaptive behavior.
- Industry 4.0 readiness: Both systems integrate seamlessly with cloud analytics, MES, and digital twins.

Integration doesn't replace the PLC—it elevates it.



Communication Pathways Between ROS and PLC

A communication protocol is essential for linking ROS and a PLC. Several proven options exist, each suited to different performance and compatibility needs.

1. OPC UA : The Industrial Standard

OPC UA is the most robust and secure method of integrating ROS with PLCs.

- The PLC functions as an OPC UA server, exposing tags.
- ROS runs a client node that reads/writes these tags.

Advantages

- Vendor-neutral and platform-independent
- Secure communication with encryption and authentication
- Supported by most modern PLCs (Siemens, Beckhoff, Allen-Bradley)

ROS Tools

- `ros_opcua` (ROS1)
- `opcua-ros2` (ROS2)

Use Case:

A ROS 2-based robot checks the PLC's "conveyor ready" tag before initiating a pick-and-place cycle.

2. Modbus TCP/IP: Simple and Effective

Modbus remains a widely supported, lightweight protocol.

- PLC: Modbus server
- ROS node: Modbus client

Advantages

- Open and license-free
- Easy to configure and debug
- Supports Ethernet and serial

Using libraries like `pymodbus`, ROS can read coils, registers, and command PLC outputs.

3. MQTT: Ideal for IIoT and Cloud Systems

MQTT is the backbone of IoT communication, using a publish/subscribe model.

- Both ROS and PLC connect to an MQTT broker.
- Messages are exchanged through topics (e.g., `/status/robot_arm`).

Advantages:

- Highly scalable for distributed systems
- Works smoothly with ROS 2
- Perfect for cloud or IIoT gateways

4. Vendor-Specific Protocols: Profinet, EtherNet/IP, ADS

PLCs also offer high-performance proprietary communication:

- Beckhoff TwinCAT ADS → `ros_twincat`
- Siemens Profinet → via OPC UA or middleware
- Allen-Bradley EtherNet/IP → via third-party tools or Kepware

These often offer better speed or determinism but may require licenses or additional software.

A Practical Architecture

A typical ROS-PLC architecture involves:

- ROS handling advanced logic like AI, vision, or motion planning
- PLC handling deterministic I/O control, actuators, safety, and machine states

Example : A robot receives a "part present" signal from the PLC, uses ROS for vision alignment, and then signals the PLC to activate the conveyor.

Example : ROS-PLC Integration Using Modbus TCP

Below is a simplified example of a ROS node interacting with a PLC through Modbus TCP. This setup allows a robot to detect PLC input changes or send output commands

```
from pymodbus.client.sync import ModbusTcpClient
client = ModbusTcpClient('192.168.0.10', port=502)
while True:
    plc_input = client.read_coils(0, 1).bits[0]
    if plc_input:
        print("PLC button pressed – robot action triggered")
        client.write_coil(1, True)
    else:
        client.write_coil(1, False)
```

Design Considerations for a Robust Integration

To ensure reliability and safety in real-world industrial environments:

1. **Keep all safety-critical logic in the PLC.**
ROS should never manage emergency stops or interlocks.
2. **Use ROS for high-level intelligence only.**
Tasks like path planning or AI-based alignment belong to ROS.
3. **Maintain deterministic behavior.**
Time-sensitive operations stay in the PLC.
4. **Implement strong network security.**
Use encryption and authentication, especially with OPC UA or MQTT.
5. **Enable logging and monitoring.**
ROS is excellent for diagnostics, data collection, and cloud reporting.

Applications Across Industry

Application	ROS Role	PLC Role
Collaborative robot + conveyor	Motion control, AI perception	Conveyor logic, safety interlocks
Vision inspection	Image processing, decision logic	Lighting, ejector control
AGVs/AMRs	Navigation, SLAM	Station control, charging
Packaging lines	Coordination, optimization	Machine sequencing, sensors

The Road Ahead: Toward Smart Factories

The future of automation lies in convergence—robots, PLCs, sensors, and cloud ecosystems communicating effortlessly. ROS delivers intelligence and flexibility, while PLCs provide stability and reliability. Together, they form the backbone of tomorrow's smart, adaptive factories.

As standards like OPC UA, DDS, and 5G advance, ROS-PLC integration will evolve from an optional capability to a standard feature of industrial automation.

Conclusion

Integrating ROS and PLC isn't just a technical interface—it's a bridge between intelligent robotics and industrial-grade control. Whether using OPC UA for a sophisticated robot cell or Modbus for a simple lab setup, this integration unlocks the next generation of connected, intelligent manufacturing.

Robots gain awareness. PLCs remain the guardians of safety. And factories become smarter, faster, and more capable than ever before.

The Thinking Machine

Neural Networks at Work



I'm Jeyasooriya, an IT Analyst at IPCS Global, Madurai, specializing in Full-Stack Development, Data Analysis, and Data Science. I have hands-on experience in Python programming and building data-driven solutions that optimize processes and deliver actionable insights. At IPCS Global, I contribute to critical IT projects while also training and mentoring students in programming, full-stack development, and data analytics, helping them gain practical, industry-ready skills. My passion for technology and education drives my work, and I am committed to leveraging data and software solutions to make meaningful impacts in the field of IT and analytics.

Artificial Intelligence (AI) has become a natural part of our everyday life. From unlocking a smartphone with our face to speaking with assistants like Alexa or Siri, AI quietly works behind the scenes to make technology smarter and more responsive. One of the major reasons behind this intelligence is a system known as Artificial Neural Networks (ANNs).

Inspired by the structure of the human brain, ANNs help computers learn from experience rather than simply following fixed instructions. Just as our brain contains billions of neurons that communicate with each other, a neural network uses artificial neurons that process information, identify patterns, and improve over time. In simple terms, ANNs enable machines to "think," "learn," and make decisions based on data.

How Neural Networks Work

A neural network consists of three key layers:

1. Input Layer

This layer receives the information – such as images, sounds, text, or numerical data.

2. Hidden Layers

These layers perform the actual learning. They detect relationships, patterns, and features within the data. The more hidden layers a network has, the more complex patterns it can learn.

3. Output Layer

This layer produces the final result – a prediction, classification, or decision.

To understand it simply : if you show a neural network thousands of pictures of cats and dogs, it gradually learns what features define each. Eventually, it can identify a new image as a cat or dog with impressive accuracy.

During training, the network compares its predictions with the correct answers. When it is wrong, it adjusts itself – a process called backpropagation. This cycle continues until the network becomes highly accurate.

Common Types of Neural Networks

Different types of neural networks are used for different tasks. Here are the most widely used ones:

1. Feed forward Neural Network (FNN)

The simplest form, where information flows in one direction – from input to output. These networks are used for basic pattern recognition and classification tasks.

2. Convolutional Neural Network (CNN)

CNNs are designed for processing images and videos. They automatically detect patterns such as edges, shapes, and objects. Facial recognition, medical image analysis, and object detection are powered by CNNs.

3. Recurrent Neural Network (RNN)

Used for sequential data where order matters – like speech, music, or text. RNNs have memory and can remember previously processed information. Their advanced versions, LSTM and GRU, are capable of remembering information for longer durations, making them ideal for translation, chatbots, and voice assistants.

4. Radial Basis Function Network (RBFN)

RBFNs are fast and efficient for tasks involving pattern classification, function approximation, and time-sensitive predictions.

5. Modular Neural Network (MNN)

MNNs break down a large problem into smaller modules. Each module solves a part of the task, and their outputs are combined. This structure improves speed, reliability, and scalability – commonly used in robotics and complex decision-making systems.

6. Generative Adversarial Network (GAN)

One of the most exciting neural network types. A GAN has two parts:

- ♦ **Generator** – creates data that looks real.
- ♦ **Discriminator** – checks whether the data is real or fake.

They compete with each other until the generator produces highly realistic content. GANs are used for creating images, enhancing old photos, generating art, and even creating synthetic video.

Where Neural Networks Are Used

Neural networks have become a backbone of modern technology. Some major applications include:

- ♦ **Face and Image Recognition** – smartphones, security systems, social media.
- ♦ **Speech Recognition** – voice assistants, transcription tools.
- ♦ **Healthcare** – detecting diseases from X-rays, CT scans, and MRIs.
- ♦ **Finance** – fraud detection, stock price prediction.
- ♦ **Self-Driving Cars** – identifying roads, traffic signs, and obstacles.
- ♦ **Language Translation** – apps like Google Translate.
- ♦ **Customer Service** – chatbots and virtual assistants.

Advantages and Limitations

Advantages

- ♦ Learn directly from data without needing detailed programming.
- ♦ Work extremely well with complex data types like images, audio, and text.
- ♦ Can automate decision-making after training.
- ♦ Perform repetitive or large-scale tasks faster and more accurately than humans.

Limitations

- ♦ Require large amounts of high-quality data.
- ♦ Training them involves significant computing power and time.
- ♦ Their decision-making process can be hard to interpret (“black box” issue).
- ♦ May give incorrect results if trained with biased or incomplete data.

Conclusion

Artificial Neural Networks are the reason AI feels intelligent, adaptable, and human-like. They have transformed how computers see, hear, understand, and react to the world around us. Although they demand substantial data and computational resources, their ability to learn and improve makes them incredibly powerful.

As technology continues to advance, neural networks will become even smarter, more efficient, and more deeply integrated into our everyday lives. Simply put, ANNs are teaching machines to learn from experience – making our world faster, safer, and more connected than ever before.

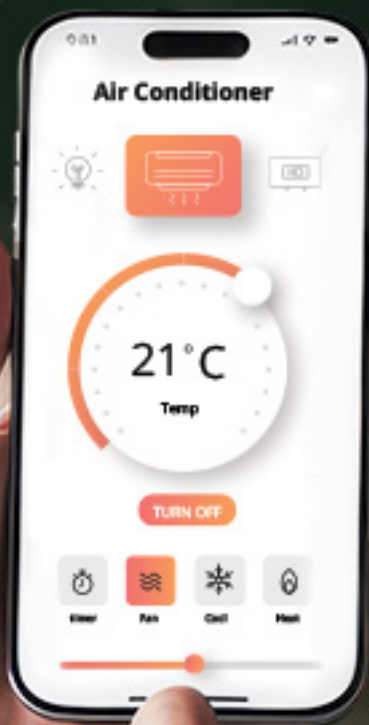
Shape Your Tomorrow

With the

Power of

Smart Home Automation

Diploma in **Building Automation Technology** (One Year)



- 📍 Assured placements
- 📜 International Certification
- 💰 Stipend-Based Internship



Transforming the skill landscape

Cooperation partner
of TUV SÜD



+91 9846 770771



www.ipcsglobal.com

Our Success



15000+

Skilled Graduates

12000+

Global Placements

1200+

Industry Projects

50+

**Training Centers
Worldwide**

120+

**Corporate
Partnerships**

500000

Students Enrolled

20+

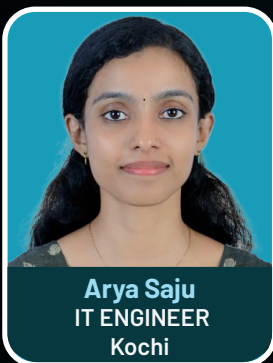
**Nationalities
Represented**

16000+

**Trained
Professionals**

One Language, Every Layer

The JavaScript Advantage



Arya Saju
IT ENGINEER
Kochi

Experienced Python Full Stack Trainer with a strong background in software development and technical training. Skilled in building and deploying scalable web applications using Python, Django, React.js, and Node.js, with expertise in REST APIs, MySQL, Git, Postman, and VS Code. Proficient in delivering comprehensive full stack development courses and mentoring students on real-world projects. Recognized for the ability to simplify complex technical concepts and foster practical, industry-focused learning experiences.

In the fast-evolving world of software engineering, choosing the right technology stack can make or break a project. As a Senior IT Engineer at IPCS Global Kochi, I've seen first-hand how JavaScript has risen to the top as the most versatile and impactful language for full-stack development. What began as a browser scripting tool has transformed into a powerhouse capable of driving everything from intuitive user interfaces to high-performance server systems. For today's developers, JavaScript isn't just an option—it's a complete solution for building modern, scalable applications.

One Language, Endless Possibilities

Imagine developing an entire application—from the interface users see to the logic running on the server—using just one language. That's the power and simplicity JavaScript brings. Instead of juggling multiple languages like Python for the backend and JavaScript for the frontend, developers can work end-to-end using React or Vue.js on the client side and Node.js or Express.js on the server side. This unified approach minimizes context switching, speeds up development, and makes debugging significantly easier.

JavaScript also excels at sharing logic between layers. Data models, validation rules, and utility functions can be reused across the stack, thanks to the natural compatibility between frontend frameworks and Node.js. And since JavaScript handles JSON—today's universal data format—seamlessly, communication between APIs and interfaces is fast, clean, and predictable.

```
init();
```



```
app.start();  
console.log('App initialized for ${user.role}');
```

A Massive Ecosystem of Tools and Frameworks

JavaScript's ecosystem is one of the largest in the software world. With millions of packages on npm, developers get access to tools for everything—UI animations, database connections, serverless functions, testing, deployment, and more.

Whether it's React, Angular, or Vue.js for building rich, interactive frontends, or Express.js, NestJS, and Fastify for backend APIs and microservices, JavaScript offers battle-tested frameworks for every requirement. Libraries like Mongoose simplify database operations, while tools like AWS SDK make cloud integration effortless. Using npm or Yarn for both sides of the stack keeps projects organized and consistent.

Built for Speed and Real-Time Performance

JavaScript thrives in real-time environments. Its asynchronous, non-blocking nature makes it ideal for applications that require instant updates and high concurrency. On the frontend, users enjoy fluid interactions, instant form feedback, and dynamic page updates. On the backend, Node.js handles thousands of simultaneous connections without breaking a sweat, making it perfect for chat platforms, streaming services, and collaboration tools.

A real-time chat app built with Socket.IO on the backend and React on the frontend is a perfect example of how JavaScript delivers responsiveness and speed in one seamless experience.

Beyond the Browser: Mobile and Desktop, Too

JavaScript's reach goes far beyond websites. With frameworks like React Native, developers can target iOS and Android using the same language. Desktop applications become possible with Electron, enabling cross-platform apps for Windows, macOS, and Linux.

This means a single JavaScript codebase can run a full web platform, its backend, and its mobile or desktop version—massively reducing development time and ensuring consistency across devices. An e-commerce website, for example, can share logic with its mobile app without starting everything from scratch.

JSON: The Quiet Hero Behind Smooth Data Flow

JavaScript's natural partnership with JSON is a major advantage in today's API-driven world. Whether it's a Node.js microservice sending data to a Vue.js interface or multiple services communicating within a distributed system, JSON keeps everything lightweight and easy to handle. This makes JavaScript an excellent fit for modern architectures like microservices, serverless applications, and REST or GraphQL APIs.



HTML

Backed by Big Players and an Ever-Growing Community

Tech giants such as Netflix, PayPal, Uber, and Airbnb rely heavily on JavaScript for large-scale applications. Its dominance is reflected in developer surveys, where it consistently ranks as the most widely used language. With countless tutorials, open-source projects, and global communities contributing daily, support is always available. New ideas, updates, and tools keep the ecosystem evolving faster than nearly any other language.

Tools That Power Productivity

Modern JavaScript development is backed by a suite of tools that streamline workflows. Webpack and Vite make builds fast and efficient. TypeScript adds type safety, reducing bugs and improving maintainability. ESLint and Prettier ensure clean, consistent code across teams. Unit testing becomes simpler with Jest, which works equally well on frontend and backend codebases.

These tools eliminate repetitive tasks and let developers focus on what matters most—building features and delivering value.

The MERN Stack: JavaScript at Its Best

MongoDB, Express.js, React, and Node.js—together known as the MERN stack—demonstrate the true power of JavaScript as a unified full-stack solution. MongoDB's JSON-like documents integrate naturally with JavaScript. Express.js simplifies API creation. React builds user-friendly interfaces. Node.js keeps everything running efficiently behind the scenes. With MERN, one skilled developer can build a complete application from database to UI, making it a favorite for startups and enterprises alike.



Not Perfect, but Powerful Enough

JavaScript does have limitations. It's not ideal for heavy computational tasks or domains like deep learning, where languages like Python or C++ perform better. Its dynamic typing can lead to errors if not managed carefully—but TypeScript solves most of those issues. For low-level system programming, JavaScript isn't the first choice. Still, for most modern applications, its strengths far outweigh its drawbacks.

Why JavaScript Truly Stands Out

For full-stack developers, JavaScript checks every box: user interfaces, business logic, databases, APIs, mobile apps, desktop apps, and cloud deployments. It scales from quick prototypes to enterprise-grade systems. It aligns perfectly with today's trends—real-time apps, PWAs, microservices, and cloud-native development. Mastering one language instead of juggling many improves efficiency and helps developers stay sharp and productive.

By focusing on JavaScript and its ecosystem, any developer can confidently build complete, high-performing applications without spreading themselves thin.

Final Thoughts

JavaScript's versatility, ecosystem strength, and ability to power both frontend and backend development make it the ultimate choice for modern full-stack engineering. At IPCS Global Kochi, I've seen how it empowers developers to build fast, scalable, and innovative applications with ease. For full-stack developers, JavaScript isn't just enough—it's a strategic advantage that unlocks creativity and productivity across every layer of development. Whether you're starting your journey or leading a major project, JavaScript remains one of the smartest tools you can choose.

```
218 }
219
220 background: -webkit-linear-gradient(#eaf2f6 0, #e0ebf1 25%);
221 background: -moz-linear-gradient(#eaf2f6 0, #e0ebf1 25%);
222 background: -ms-linear-gradient(#eaf2f6 0, #e0ebf1 25%);
223 background: linear-gradient(#eaf2f6 0, #e0ebf1 25%);
224 }
225 to {
226   -webkit-box-shadow: inset 0 1px 2px rgba(20, 60, 75, .7);
227   -moz-box-shadow: inset 0 1px 2px rgba(20, 60, 75, .7);
228   box-shadow: inset 0 1px 2px rgba(20, 60, 75, .7);
229 }
}

215 {}
216 @keyframes newlinetest
217 {
218   from {
219     background: -webkit-linear-g
220     background: -moz-linear-gra
221     background: -ms-linear-gradi
222     background: linear-gradient(a
223   }
224   to {
225     -webkit-box-shadow: inset 0 :
226     -moz-box-shadow: inset 0 1px
227     box-shadow: inset 0 1px 2px
228   }
229 }
```

Lost on the Web

The Surprising Life of 404 Pages



I am an IT Engineer and AR/VR Developer with a strong passion for creating innovative digital experiences through technology and design. Skilled in building interactive AR/VR applications and developing web solutions using Unity, Django, React, and Java, I bring a blend of creativity and technical expertise to every project. My core strengths include AR/VR development, web and mobile application development, UI/UX and interaction design, project planning, and creative graphic design. I am always eager to explore new technologies, transform ideas into impactful digital solutions, and stay ahead in the rapidly evolving world of innovation.

Whether you're browsing a blog, shopping online, or exploring a corporate website, you've likely encountered the infamous Error 404: Page Not Found message. While this HTTP error feels like a dead end for users, it has evolved far beyond a simple warning. For modern designers and developers, the 404 page has become a technical safeguard, a branding opportunity, and – surprisingly – a tool for improving user experience and site performance.

What Is a 404 Error, Technically?

A 404 error is an HTTP status code indicating that the server is reachable but cannot locate the requested resource.

This typically occurs due to:

- Deleted or moved pages
- Incorrect URLs
- Broken or outdated internal links
- Typing mistakes
- Misconfigured redirects

According to web analytics reports, 404 errors make up 8–12% of all HTTP errors globally, affecting millions of browsing sessions every day. For large websites with thousands of pages, 404s are almost unavoidable.



Beyond a Warning: The Modern 404 Page as a Micro-Experience

In the 1990s, 404 pages were plain white screens with basic text generated by a server. But as web design matured, businesses realized that a user encountering an error is experiencing a moment of high friction—and that moment can either frustrate or engage them.

Today, over 73% of top-ranked websites use customized 404 pages that include:

- Visual design
- Brand elements
- Navigation options
- Search functionality
- Humor or storytelling
- Interactive elements (mini-games, animations, Easter eggs)

These improvements are not just aesthetic—they're strategic.

Examples That Made 404 Pages Famous

Some brands have mastered the art of transforming errors into memorable moments:

GitHub

Its 404 page features the Octocat drifting in space. The design reinforces the theme of “lost” while staying on-brand and visually delightful. GitHub’s 404 page is so iconic that users often visit it intentionally.

Lego

Lego showcases characters in a “broken” scene, reflecting the idea of missing pieces. It uses minimal text and strong visuals—perfect for a brand built on imagination.

Netflix

Netflix uses witty text paired with simple illustrations. Their tone remains friendly and human, acknowledging the issue without overwhelming the user.

These pages succeed because they merge creativity with function, making errors feel intentional rather than disruptive.

The UX Psychology Behind a Creative 404 Page

A well-designed 404 page isn’t just visually appealing—it taps into user psychology. Research from the Nielsen Norman Group and UX Collective highlights several reasons why creative 404 pages improve user experience:

1. Humor Reduces Friction

Humor lowers stress and transforms frustration into amusement.

A study on UX emotional design showed that users were 22% more likely to continue browsing when error messages included visual or humorous elements.

2. Brand Memory Increases

A unique 404 page acts as an additional brand touchpoint.

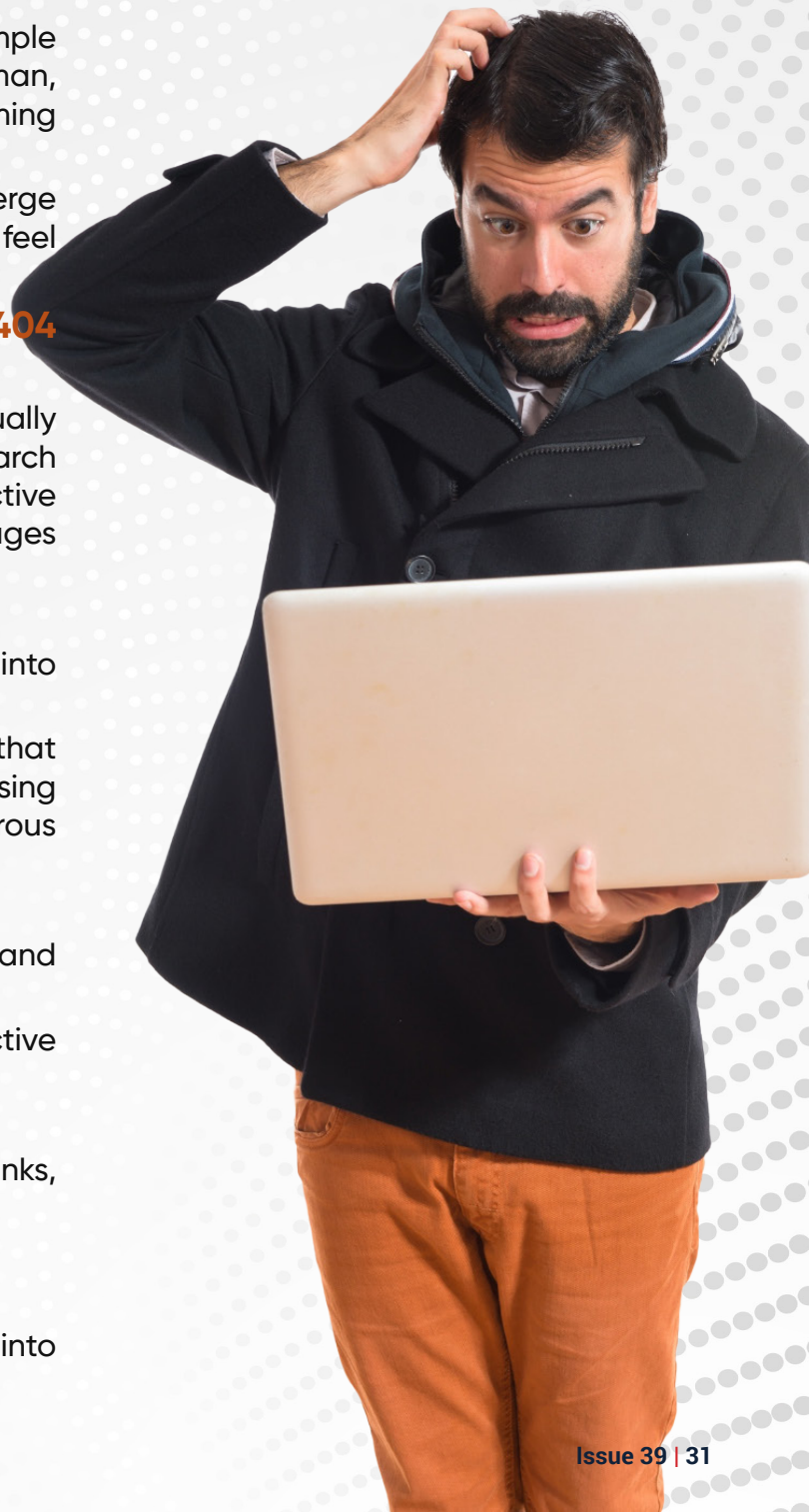
According to Google’s UX principles, distinctive visuals can improve brand recall by up to 30%.

3. Engagement Goes Up

Websites that offer search bars, navigation links, or suggested content in their 404 pages see:

- +15% increase in session continuation
- Reduction in bounce rate by 12–20%

A 404 page can literally redirect users back into the site’s value instead of pushing them away.



404 Pages as Business Tools: Not Just UX, But Also SEO

From a technical perspective, 404 pages impact more than user experience—they influence SEO and site health.

SEO Benefits of Smart 404 Pages

- Google crawlers use 404 pages to understand which links are broken.
- A custom 404 prevents indexing of blank or confusing server errors.
- When users navigate back into the site, bounce rates drop—boosting SEO.

Search Engine Journal reports that optimized 404 pages can:

- Reduce bounce rate by up to 18%
- Improve average session duration by 10–13%
- Increase the number of pages viewed per session

This makes a well-designed 404 page essential for overall site performance.

Turning Errors Into Opportunities

A 404 page can also serve strategic purposes:

E-Commerce Platforms

Display:

- Trending products
- Seasonal deals
- Discount codes

Large retailers like Amazon and Shopify-based stores report 3–6% conversion recovery from users who landed on a 404 but continued shopping thanks to helpful navigation or offers.

Travel Websites

Use high-quality destination photos and links to popular packages, converting user frustration into inspiration.

News & Blog Sites

Feature:

- Recent articles
- Popular categories
- Search tools

Users discover new content instead of leaving immediately.



Best Practices for Designers and Developers

To make a 404 page effective, designers and developers need to blend creativity with usability:

1. Maintain Brand Consistency

Color schemes, tone, fonts, and visuals should mirror the main site.

2. Provide Clear Navigation

Include:

- Home button
- Search bar
- Main categories
- Contact options
- Sitemap link

3. Add Interactive or Visual Elements

Mini-games, animations, or illustrations make the experience memorable.

4. Show Empathy

A friendly message like:

"Oops! Looks like you're lost. Let's get you back on track."

can soften irritation.

5. Ensure Mobile Optimization

Over 58% of 404 visits occur on mobile devices, making responsive design critical.

6. Log and Monitor 404 Errors

From a developer standpoint:

- Use analytics tools (Google Analytics, Search Console)
- Monitor 404 frequency
- Fix broken links or set up redirects

Proper monitoring improves site reliability and SEO.

More Than an Error : The New Role of 404 Pages in UX Design

The rise of creative 404 pages reveals a deeper shift in digital design. Modern UX emphasizes:

- Empathy
- Micro-interactions
- Emotional design
- User journey continuity

When companies pay attention to even the error states of their website, users feel valued—and that strengthens trust and loyalty.

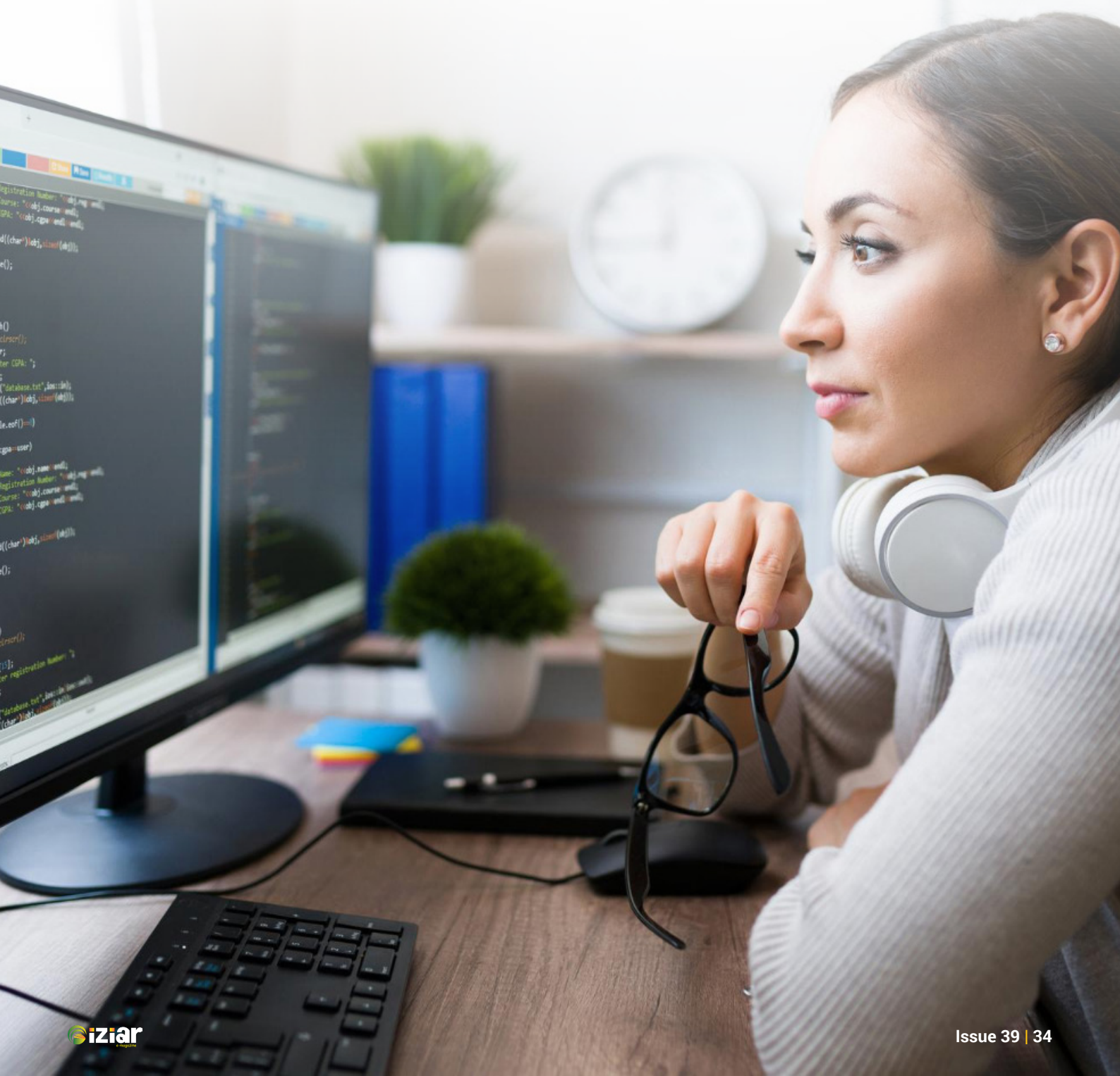
Conclusion: The Hidden Art of Error Pages

The next time a 404 message appears, take a moment to appreciate what's happening behind the scenes. What used to be a lifeless system warning has evolved into a blend of technical precision, creative design, UX psychology, and business strategy.

A great 404 page isn't just a notification—it's an opportunity:

- To engage
- To guide
- To entertain
- To express brand personality
- To retain users
- To improve SEO

In the world of digital experiences, even a dead end can point the way forward. The humble 404 page proves that every moment—no matter how unexpected—can be crafted into something delightful, memorable, and meaningful.



Automation to Collaboration

The Real Meaning of Industry 5.0




I am an enthusiastic Junior Project Engineer with a strong academic background in Electronics and Instrumentation Engineering, further strengthened by a PG Diploma in Industrial Automation. I'm driven by curiosity, innovation, and a constant desire to learn. I actively work on developing smart, reliable, and efficient engineering solutions, and I enjoy using modern automation technologies to improve system performance, optimize processes, and contribute to industrial innovation. I see myself as part of the next generation of forward-thinking professionals shaping the future of automation.

For more than a decade, Industry 4.0 has been the dominant force driving industrial transformation around the globe. It introduced a new era of cyber-physical systems, advanced robotics, IoT sensors, cloud platforms, and predictive analytics that brought factories closer to full autonomy than ever before. Manufacturing lines became highly optimized digital ecosystems capable of self-monitoring and self-correction, while enterprise systems could access real-time production data from anywhere in the world. Yet, as organizations approached the upper limits of automation, a fundamental realization emerged: technological sophistication alone does not guarantee resilience, creativity, or long-term sustainability. Pure automation, while efficient, cannot substitute the adaptability, ethical judgment, and creative problem-solving that humans naturally possess. This realization opened the door to Industry 5.0 – an evolution that places humans back at the centre of industrial innovation, not as operators overshadowed by machines, but as empowered collaborators working alongside intelligent systems.

Industry 4.0 transformed factories into interconnected networks where every machine, sensor, and process contributed to a continuous data loop. With the rise of industrial IoT platforms, SCADA-to-cloud integration, and AI-driven analytics, organizations gained unprecedented visibility across the value chain. Predictive maintenance reduced downtime by up to 30%, digital twins enabled virtual testing before physical deployment, and autonomous robots handled high-precision, repetitive tasks with unparalleled consistency. These breakthroughs redefined productivity and set new global benchmarks in manufacturing excellence. But even as automation improved operational efficiency, it introduced challenges that technology alone could not solve. Systems optimized strictly for speed and cost often lacked the flexibility required during unexpected global disruptions. Machine-controlled environments struggled to adapt to rapidly changing market demands, especially when personalization, sustainability, and innovation became core expectations. These limitations revealed that Industry 4.0 had reached a plateau where further progress required more than better algorithms or faster robots—it required the reintegration of human intelligence.

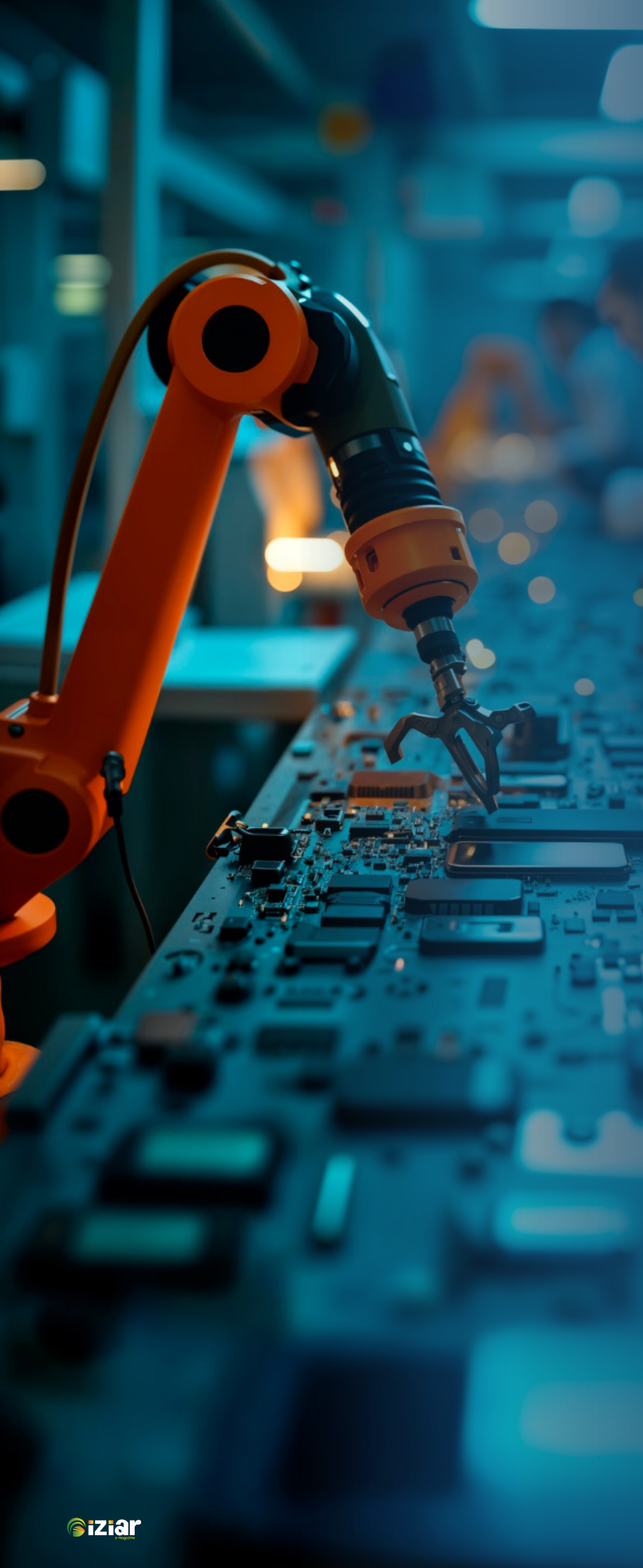
Industry 5.0 reimagines the factory not as an autonomous machine hub but as a collaborative ecosystem where humans and machines work symbiotically. Instead of designing factories where people fit into rigid automated workflows, Industry 5.0 creates technologically augmented environments that adapt to humans. Collaborative robots (cobots) replace isolated industrial robots, sharing workspaces and responding safely to human presence through advanced sensing and AI vision. Cognitive AI systems act not as decision-makers but as advisors, interpreting vast datasets and presenting insights for human judgment. Production lines evolve from mass manufacturing to flexible, demand-driven micro-factories capable of delivering customization at near-industrial scale. These changes reflect a larger philosophical shift: technology becomes a tool that amplifies human capability rather than replacing it.





A defining characteristic of Industry 5.0 is its embrace of mass personalization. While Industry 4.0 optimized mass production, modern consumers increasingly demand products tailored to individual needs—from personalized medical implants produced via additive manufacturing to customized consumer electronics designed with AI-aided creativity tools. Industry 5.0 makes such personalization economically viable by combining human creativity with machine efficiency. Designers, engineers, and operators can co-create solutions with AI-driven generative design platforms, while modular robotic workcells adjust automatically to unique product configurations. This shift from uniformity to individuality marks one of the most significant commercial advantages of Industry 5.0.

The rise of Industry 5.0 is also fuelled by global challenges that exposed the limitations of machine-only approaches. The COVID-19 pandemic demonstrated how highly automated supply chains could fail when unforeseen disruptions occur. Fully automated processes lacked the improvisational intelligence humans naturally exhibit in crises. Likewise, climate commitments, carbon regulations, and resource constraints require decision-making that considers ethics, sustainability, and long-term impact—areas where human judgment is irreplaceable. Industry 5.0 introduces the concept of human-in-the-loop resilience, where automated systems provide speed and efficiency, while humans ensure adaptability and ethical responsibility.



Human-machine collaboration is also solving workforce challenges. The manufacturing sector faces a global shortage of skilled labour, with an estimated gap of 8 million jobs by 2030. Younger generations often avoid factory work that is monotonous, hazardous, or physically demanding. Industry 5.0 transforms these environments into high-tech, creative, and safer spaces. Cobots handle heavy lifting and precision tasks, while augmented reality interfaces guide workers through complex procedures with real-time overlays. Even inexperienced technicians can perform expert-level tasks using AI-assisted work instructions, bridging the skill gap and democratizing advanced manufacturing knowledge.

Sustainability forms another foundational pillar of Industry 5.0. With energy costs rising and environmental regulations tightening, manufacturing cannot rely solely on automation-driven output. Industry 5.0 uses AI-driven energy optimization, smart resource allocation, and waste-reducing precision robotics to create environmentally conscious production systems. Digital twins allow manufacturers to simulate eco-friendly processes, measure carbon footprints, and optimize material usage before implementing changes physically. Instead of treating sustainability as an afterthought, Industry 5.0 integrates it as a measurable performance metric on par with efficiency and profitability.

Ultimately, the future of manufacturing is not framed as a contest between human and machine intelligence but as a partnership that elevates both. Machines bring endurance, precision, and computational power. Humans bring empathy, imagination, ethical reasoning, and the ability to navigate ambiguity. Together, they form an industrial model that is more innovative, more resilient, and more meaningful than either could achieve independently. The factory of the future is not a silent, fully automated environment—it is a dynamic, intelligent workspace where humans orchestrate complex operations with machines as their most capable partners.

Industry 5.0 represents the next chapter of industrial evolution, one that recognizes that progress is not measured only in output but in creativity, sustainability, and human empowerment. It ensures that technology continues to advance without losing sight of the people it is meant to serve. As the world faces rising global complexities, the industries that thrive will be those that blend digital intelligence with human insight, forging a future where innovation is not only efficient, but profoundly human-centered.



Beyond ChatGPT

The Evolution from Transformers to RAG



Muhammed Farhan
IT Engineer
Attingal

I'm Muhammed Farhan PC, an IT Engineer driven by a deep passion for Python programming and Data Science. With hands-on experience in building intelligent solutions and uncovering meaningful insights from data, I focus on using technology to solve practical, real-world challenges. My work blends logical thinking with strong analytical skills, enabling me to design data pipelines, automate processes, and develop predictive models that support smarter, data-driven decisions. From handling complex datasets to visualizing trends and deploying machine learning models, I'm always exploring the latest tools and innovations within the data science ecosystem. I thrive in dynamic environments where curiosity fuels learning and innovation. Constantly evolving my expertise in Python, machine learning, and analytics, I aim to transform data into impactful, actionable solutions.

Artificial Intelligence (AI) has transformed the way we communicate, work, and interact with technology. What began with simple assistants like Siri and Alexa has now evolved into conversational systems capable of reasoning, generating ideas, and engaging in human-like dialogue. But what exactly powers this new generation of intelligent chatbots? The answer lies in two groundbreaking innovations: Transformers and RAG (Retrieval-Augmented Generation). This article explores how Transformers reshaped the foundations of modern AI—and how RAG is elevating chatbots even further, making them not just fluent but reliably informed.

The Birth of the Transformer

Before 2017, most language models relied on RNNs and LSTMs, which processed text sequentially. These models often struggled with long sentences and lost track of earlier context.

Everything changed when Vaswani et al. introduced the Transformer in the landmark paper *"Attention is All You Need."*

Instead of reading text word by word, Transformers use self-attention, a mechanism that lets them analyze all words at once and understand their relationships.

For example, in the sentence:

"The bird flew over the river because it was wide."

the Transformer correctly identifies that "it" refers to the river, not the bird.

This architecture became the backbone of models like BERT, GPT, and T5, which now power most of today's intelligent language applications.



ChatGPT and the Power of Generative AI

Built on the Transformer architecture, OpenAI's GPT models learned from massive amounts of text—books, articles, websites, and more. When you interact with ChatGPT, it doesn't perform real-time internet searches. Instead, it predicts the most logical sequence of words based on patterns learned during training.

This is why ChatGPT can:

- ♦ Write poems
- ♦ Answer complex questions
- ♦ Summarize documents
- ♦ Generate code

However, since it relies solely on its training data, it sometimes produces incorrect answers — a problem known as hallucination. And this is where RAG steps in to enhance accuracy.

Enter RAG – Retrieval-Augmented Generation

Imagine a chatbot that can search relevant documents before answering you—combining the creativity of Transformers with the factual grounding of a search engine. That is exactly what RAG delivers.

A RAG system has two main components:

1. **Retriever:** Finds relevant information from external sources (e.g., PDFs, knowledge bases, websites).
2. **Generator:** Uses a Transformer model to create a natural, well-informed response based on the retrieved data.

Instead of guessing, RAG-powered systems read first and answer next.



Why RAG Matters

Feature	Transformer	RAG
Knowledge Source	Internal training data	External documents & databases
Accuracy	Can hallucinate	More grounded and factual
Updating Info	Requires retraining	Easy—just update knowledge base
Best Use Cases	Text generation, summarization	Question answering, chatbots, knowledge retrieval

RAG bridges the gap between static knowledge (what a model was trained on) and dynamic knowledge (real-time or updated information). This makes it ideal for:

- ♦ Customer support systems
- ♦ Healthcare assistance tools
- ♦ Educational tutors
- ♦ Enterprise-grade knowledge bots

Real-World Adoption

Major tech companies are already leveraging RAG:

- ▶ Meta introduced the first RAG model in 2020.
- ▶ OpenAI integrates RAG capabilities into ChatGPT using user-uploaded documents.
- ▶ Google Gemini and Microsoft Copilot also employ retrieval-based architectures for real-time information access.

These advancements are making AI conversations more accurate, reliable, and contextually rich.

Challenges Ahead

Despite its advantages, RAG still faces hurdles:

- ▶ **Retrieval quality** : Poor search results lead to poor answers.
- ▶ **Response speed** : Searching large databases can slow the model.
- ▶ **Data privacy** : Sensitive documents require secure handling.

Researchers are now exploring hybrid systems that combine RAG with reasoning engines and reinforcement learning to further improve reliability.

The Future: Knowledge-Aware AI

The next generation of chatbots will not just sound intelligent—they will think intelligently. With RAG, AI systems will gain the ability to verify facts, reference external knowledge, and explain their reasoning, much like humans do.

We are entering an era where AI continuously learns from the world, not just from its training data.



Conclusion

Transformers taught machines how to understand language.

RAG teaches them how to use information wisely.

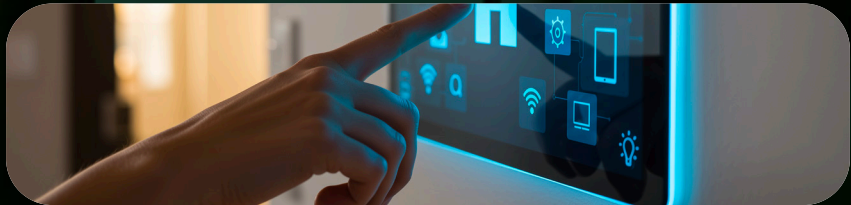
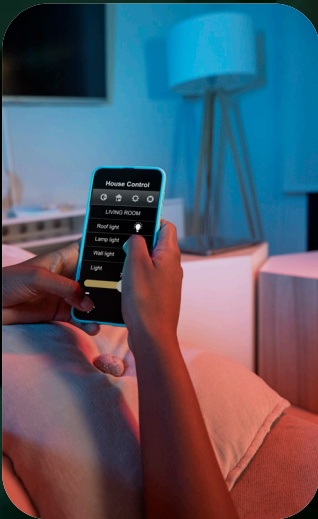
From simple assistants to deeply knowledgeable companions, chatbots have evolved dramatically—and RAG marks the next major leap toward building AI systems that are not only powerful but truly trustworthy.



Diploma In Building Automation Technology

.....

The One-Year Smart Building Automation Technology course by IPCS Global offers 410 hours of intensive, hands-on training. Covering electronics, CCTV, fire alarms, access control, IoT, renewable energy, and BMS systems, it equips students with real-world skills in automation, safety, and surveillance. Designed for technical learners, it ensures 100% practical exposure, internship, and placement support for a successful career.



IPC'S GLOBAL,
THE WORLD'S TRUSTED
INDUSTRY - BASED
TRAINING INSTITUTE!

A Thousand Plus Success Stories

The 2024-25 Batch Witnessed Over 1000+ Placements
Stories Written with Skill and Ambition...



OPC UA

Powering the Connected Factory Revolution



Mallethwaran
Teach lead
Coimbatore

I am a dedicated, results-driven Automation Engineer with strong experience in PLC programming, SCADA development, HMI design, and VFD integration across a wide range of platforms, including Siemens, Allen-Bradley, Schneider, Delta, ABB, Mitsubishi, and Omron. I also have solid expertise in Building Management Systems (BMS), CCTV, fire alarm, and access control systems, ensuring smooth integration and reliable operation.

My hands-on background in control panel design, troubleshooting, and industrial communication protocols enables me to deliver safe, efficient, and well-optimized automation solutions. I take pride in combining strong technical knowledge with analytical problem-solving and collaborative teamwork to enhance plant performance.

Passionate about continuous improvement, I stay updated with the latest automation technologies to deliver innovative, high-quality solutions that boost industrial efficiency and productivity.

In the era of Industry 4.0, the success of a smart factory depends on secure, standardized, and high-fidelity communication between machines, edge devices, control systems, and enterprise-level software. Traditional industrial protocols—Modbus, Profibus, DeviceNet, and vendor-specific Ethernet variants—have served the industry for decades but suffer from closed architectures, limited security, and poor cross-vendor interoperability.

To overcome these limitations, the Open Platform Communications Unified Architecture (OPC UA) emerged as a platform-agnostic industrial interoperability standard engineered to unify IT and OT communications, enable complex data modeling, and support secure, scalable, service-oriented architectures (SOA). OPC UA is not simply another protocol; it is a complete industrial communication framework.

What is OPC UA? – A Technical Definition

OPC UA is the next-generation evolution of the classic OPC standard created by the OPC Foundation. Unlike classic OPC—built on Microsoft COM/DCOM—OPC UA is based on cross-platform binaries, standardized communication stacks, and service-oriented architectures.

OPC UA is implemented in three major profiles:

1. 1. UA Binary over UA TCP (High-performance)

- ♦ Optimized for deterministic, low-latency industrial communication.
- ♦ Widely used in PLCs, embedded systems, and SCADA.

2. UA over HTTPS / WebSockets (Cross-platform IT integration)

- ♦ Ideal for cloud, web-based clients, and IoT devices.

3. OPC UA Pub/Sub (UDP Multicast, MQTT)

- ♦ Enables real-time, event-driven, many-to-many communication.
- ♦ Foundation for Time Sensitive Networking (TSN) integration.

The architecture supports:

- ♦ PLCs
- ♦ Sensors
- ♦ SCADA
- ♦ Historian servers
- ♦ MES/ERP platforms
- ♦ SCADA/HMI
- ♦ Cloud analytics, edge gateways

All using a common, secure, object-oriented data exchange model.

Why OPC UA Was Needed – Technical Limitations of Legacy Protocols

Traditional fieldbus and Ethernet protocols were designed for device-level communication but lacked features required for Industry 4.0:

1. Vendor Lock-In

Each manufacturer used its own proprietary ecosystem:

- ♦ Siemens → PROFINET
- ♦ Rockwell → EtherNet/IP
- ♦ Schneider → Modbus TCP

Cross-vendor integration required protocol converters and engineering time.

2. No Data Semantics

- ♦ Legacy protocols exchange raw bytes without structure.
- ♦ OPC UA supports rich, semantic, object-modeled data.

3. Poor Cybersecurity

Classic protocols lack:

- ♦ Encryption
- ♦ Certificate-based authentication
- ♦ Secure session establishment
- ♦ Role-based access control

4. Limited Scalability

Adding devices required manual configuration and individual drivers.

OPC UA supports auto-discovery, namespace browsing, and unified addressing.



Key Technical Features of OPC UA

1. Platform Independence

OPC UA uses:

- ♦ UA Binary for high-speed embedded systems
- ♦ JSON, XML for enterprise systems
- ♦ TCP, UDP, HTTP, MQTT transports

It runs on:

- ♦ Linux, Windows
- ♦ Android
- ♦ Embedded RTOS
- ♦ Microcontroller-class hardware (even ARM Cortex-M)

2. Unified Information Model (UAModel)

OPC UA defines a Node-based, object-oriented address space containing:

- ♦ Objects
- ♦ Variables
- ♦ Methods
- ♦ DataTypes
- ♦ ReferenceTypes
- ♦ Views

Nodes include metadata, semantics, and relationships, enabling:

- ♦ Device parameter descriptions
- ♦ Calibration metadata
- ♦ Configuration structures
- ♦ Hierarchical modeling
- ♦ Custom domain models (via OPC UA Companion Specifications)

Example: A temperature sensor exposes:

- ♦ Value
- ♦ EngineeringUnits (°C, K)
- ♦ CalibrationDate
- ♦ Accuracy
- ♦ Manufacturer info

This semantic richness makes OPC UA ideal for analytics, MES, and digital twin applications.

3. Built-in Security Architecture

OPC UA follows a defense-in-depth model with:

- ♦ TLS/SSL secure channels
- ♦ X.509 certificate-based authentication

- ♦ Asymmetric key exchange
- ♦ Symmetric encryption (AES-128/256)
- ♦ Message signing & integrity checks
- ♦ User role-based authorization

It complies with modern standards such as: ISA95, ISA99/IEC 62443, NIST, and industrial cybersecurity best practices.

4. Scalability Across Levels

OPC UA supports:

Level	Examples	OPC UA Role
Field Devices	Sensors, actuators	UA Embedded Server
Control Layer	PLCs, PACs	Full server + client functionality
SCADA/MES	Historian, alarms	Namespace modeling, subscriptions
Cloud/IT	Analytics, dashboards	HTTPS, MQTT Pub/Sub

5. Companion Specifications (Industry-Specific Models)

Over 70+ standardized information models, including:

- ♦ PackML
- ♦ Euromap (plastics machines)
- ♦ VDMA (machine builders)
- ♦ IEC 61850 (power systems)
- ♦ Auto industry, building automation, pharmaceutical systems

This guarantees semantic interoperability across vendors.

6. Cloud and IIoT Integration

OPC UA supports:

- ♦ Native MQTT integration
- ♦ JSON encoding for IoT platforms
- ♦ Direct communication with Azure IoT Hub, AWS IoT Core, Google IoT Core
- ♦ OPC UA over WebSockets for browser-based HMIs

How OPC UA Works – Technical Architecture

OPC UA uses a client-server pattern for structured data access and a Pub/Sub pattern for scalable, high-performance communication.

OPC UA Client–Server Workflow

1. **Discovery** : Clients locate servers using mDNS or LDS (Local Discovery Server).
2. **Secure Channel** : Certificates exchanged and a secure session established.
3. **Address Space Browsing** : Clients explore nodes using standardized browse services.
4. **Subscriptions** : Clients subscribe to monitored items for real-time updates.
5. **Method Invocation** : Clients can call functions implemented inside devices.
6. **Writes/Commands** : Secure control commands or configuration data sent.

Benefits of OPC UA for Industry 4.0

1. True Cross-Vendor Interoperability

A Siemens S7-1500 PLC can send data to a Rockwell SCADA or a Schneider EMS without any protocol converters.

2. Unified System Architecture

Reduces need for vendor-specific drivers and custom middleware.

3. Future-Proof Communication

Supports:

- ◆ Pub/Sub
- ◆ TSN
- ◆ Cloud-native interfaces
- ◆ Edge computing
- ◆ Digital twins

4. Enterprise-Level Data Visibility

Rich data models support:

- ◆ AI analytics
- ◆ Predictive maintenance
- ◆ Real-time optimization
- ◆ Big-data platforms

5. Best-In-Class Cybersecurity

Built to meet global regulatory and industrial cyber standards.

Applications of OPC UA – Technical Use Cases

Smart Manufacturing

- ◆ Cross-vendor PLC networking
- ◆ centralized SCADA
- ◆ unified data logging
- ◆ machine-to-cloud connectivity

Energy Sector

- ◆ Substation automation
- ◆ renewable energy monitoring (wind, solar, hydro)
- ◆ integration with IEC 61850 models

Pharmaceuticals

- ◆ GxP-compliant secure data logging
- ◆ traceability using standardized models

Automotive

- ◆ High-speed robotic communication
- ◆ vision inspection integration
- ◆ digital twins and MES connectivity

Building Automation

- ◆ HVAC, metering, lighting systems with unified namespace

Challenges in Deploying OPC UA

1. Certificate Management

PKI setup and certificate rotation can be complex.

2. Engineering Skill Gap

OPC UA requires understanding of:

- ♦ Data modeling
- ♦ Namespace design
- ♦ Cybersecurity concepts (PKI, TLS)
- ♦ Pub/Sub and multicast networking

3. Legacy System Integration

Older PLCs may only support Modbus or proprietary protocols, requiring:

- ♦ OPC UA gateways
- ♦ protocol conversion appliances
- ♦ edge computing devices

Despite these challenges, its benefits make OPC UA the core technology for future-ready automation systems.

Conclusion

OPC UA is not a minor upgrade but a complete re-architecture of how industrial systems communicate. Through:

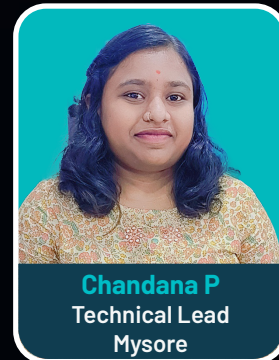
- ♦ platform independence
- ♦ secure communication models
- ♦ rich information modeling
- ♦ IT/OT unification
- ♦ cloud and IIoT readiness
- ♦ real-time Pub/Sub + TSN

OPC UA becomes the central pillar of Industry 4.0 and the foundation for fully connected, intelligent factories.

As industries advance toward autonomous operations, predictive analytics, and digital twins, OPC UA will remain the de facto communication and data modeling standard for industrial connectivity—from the factory floor to the enterprise cloud.

Beyond Text

The New Era of Image, Voice and Video Intelligence



I work with Python, Machine Learning, and SQL to develop data-driven solutions that address real-world challenges. Over the years, I've built a strong foundation in Machine Learning and Deep Learning, continuously expanding my expertise through hands-on projects and ongoing learning. Recently, I've been exploring emerging advancements in AI and staying updated with the technologies shaping the future of data science. I am also enhancing my ability to communicate insights effectively by learning Power BI and Tableau, enabling me to create clear, insightful visualizations and dashboards that support informed decision-making. My approach is practical, growth-oriented, and focused on delivering meaningful, data-backed solutions that create tangible impact.

Artificial Intelligence has come a long way—from rule-based systems designed for narrow tasks to advanced neural architectures capable of perceiving the world in ways that resemble human cognition. Early AI models were unimodal: an image model could analysis only images, a language model only text, and a speech model only audio.

Today, that limitation is fading rapidly. The rise of *multimodal AI* represents one of the most transformative shifts in machine intelligence.

These systems can process and integrate multiple data types—text, images, audio, video, and even sensor inputs—creating a more holistic and context-rich understanding of information.

Imagine speaking a description of a scene and an AI instantly generating a realistic image, analyzing the objects, interpreting your tone, and answering follow-up questions. This seamless cross-modal reasoning is the promise of multimodal AI.

The Concept of Multimodal Learning

Multimodal learning is inspired by human cognition. We interpret the world using several senses simultaneously—vision, hearing, language, and spatial awareness. Multimodal AI replicates this by combining information from different sources to capture nuance, emotion, and context that a single modality cannot express.

For example, the sentence “He smiled while saying sorry” means one thing, but adding facial expression or vocal tone may completely alter the intent. Multimodal AI aims to recognize that deeper context, making systems more perceptive, intuitive, and reliable.

This cross-modal capability enhances applications from intelligent assistants to medical diagnostics, enabling smarter, more human-like interaction.

How Multimodal Models Are Built

A multimodal AI system typically relies on three core components:

1. Feature Extraction

Each data type is first converted into numerical embeddings:

- ◆ Text → Transformers like BERT, GPT
- ◆ Images → CNNs or Vision Transformers (ViT)
- ◆ Audio → wav2vec, Whisper, or spectrogram-based models

These encoders capture semantic, spatial, or acoustic features.

2. Data Fusion

The extracted features are combined into a shared embedding space where cross-modal relationships are learned—for example, connecting the word “cat,” an image of a cat, and the audio of a meow.

3. Representation Learning

Through large-scale training, the model learns to reason across modalities, such as associating gestures with emotion or linking video frames to narrative descriptions.

This architecture enables AI to integrate information the way humans do—holistically rather than in isolation.





Leading Multimodal Models and Frameworks

Several groundbreaking models have accelerated progress:

- ♦ **CLIP (OpenAI):** Learns strong image–text associations from large-scale datasets.
- ♦ **Flamingo (DeepMind):** Performs few-shot multimodal reasoning with minimal examples.
- ♦ **GPT-4o (OpenAI):** A real-time multimodal model capable of analyzing text, images, audio, and video.
- ♦ **Google Gemini:** Integrates multiple sensory inputs for simultaneous multimodal analysis.
- ♦ **LLaVA, Kosmos-2:** Open-source projects democratizing multimodal AI for research and development.

These systems are pushing AI closer to genuine contextual intelligence.

How Multimodal AI Works in Practice

A multimodal pipeline generally includes:

1. **Data Collection** – Image–caption datasets, video–transcript pairs, audio–text pairs
2. **Preprocessing** – Cleaning, aligning, and standardizing different data types
3. **Encoding** – Converting each modality into vector embeddings
4. **Fusion** – Combining embeddings into a unified space
5. **Prediction/Generation** – Producing outputs across modalities

For instance, an AI assistant can analyse a graph, understand a spoken query about it, and respond with a verbal explanation—much like a human instructor.

Real-World Applications of Multimodal AI

Healthcare

- ♦ Combines medical images, doctor’s notes, and patient speech for more accurate diagnosis.

Education

- ♦ Intelligent tutors analyse facial expressions, spoken questions, and writing patterns to offer personalized learning support.

Retail

- ♦ Visual search, gesture recognition, and voice-based shopping enhance customer experience.

Security

- ♦ Audio-visual surveillance systems detect anomalies through a combination of sound and movement analysis.

Entertainment

- ♦ Automated dubbing, video editing, and interactive storytelling powered by multimodal inputs.

Challenges in Building Multimodal Systems

Despite their promise, several challenges remain:

- ♦ **Data Alignment:** Difficulty in synchronizing audio, video, and text.
- ♦ **Data Requirements:** Multimodal datasets are expensive to collect and store.
- ♦ **High Computational Costs:** Require advanced GPUs and large-scale training.
- ♦ **Ethical Concerns:** Potential biases, privacy risks, and misuse.
- ♦ **Interpretability:** Hard to explain decisions made across multiple modalities.

Addressing these issues is essential for safe and wide-scale adoption.

Explainability and Interpretability

Multimodal systems amplify the “black box” challenge. When decisions are influenced by combinations of text, visuals, and sound, tracing the reasoning becomes complex.

To mitigate this, researchers use:

- ♦ Attention heatmaps
- ♦ Modality influence graphs
- ♦ Visualization tools for cross-modal interactions

These methods improve trust, especially in high-stakes fields like healthcare and finance.

Impact on Analytics

Analytics is expanding beyond numbers and text to incorporate new modalities:

- ♦ **Marketing:** Analyzing facial expressions + voice tone + feedback text.
- ♦ **Customer Service:** Sentiment analysis combining speech and text.
- ♦ **Sports:** Merging video action data with commentary and statistics.

The result is richer, more actionable insights.

The Indian Context

India's multilingual and visually diverse environment makes it ideal for multimodal AI adoption. Applications include:

- ♦ **Education:** Adaptive learning tools in regional languages.
- ♦ **Agriculture:** AI interpreting farmers' videos and voice queries.
- ♦ **Customer Support:** Multilingual voice + text + image-driven support.

Government initiatives like Digital India and India AI Mission are accelerating indigenous innovation.

Future of Human–AI Collaboration

As multimodal AI becomes more intuitive, human–AI interaction will shift from keyboards to natural communication—voice, gestures, and visuals.

Professionals will need skills in:

- ♦ Designing multimodal workflows
- ♦ Interpreting multimodal outputs
- ♦ Working alongside AI-driven tools

This will reshape workplaces, training programs, and digital collaboration.

Multimodal AI in Education and Training

The classrooms of the future will leverage:

- ♦ Speech recognition for interactive Q&A
- ♦ Video analytics for engagement tracking
- ♦ AI-based summarizers for quick assessment
- ♦ Immersive AR/VR experiments

Educators and trainers will blend traditional methods with intelligent multimodal tools.



Ethics, Privacy & Policy

Handling audio and video data raises serious ethical considerations. Governance frameworks must emphasize:

- ▶ Transparent data practices
- ▶ Consent-driven collection
- ▶ Explainable AI decision-making
- ▶ Fairness and inclusivity in model design

Responsible AI is essential for sustainable development.

Case Studies

- ▶ CLIP helps visually impaired users by describing scenes.
- ▶ Flamingo supports radiology by merging text reports with imaging data.
- ▶ Google Gemini performs live video-audio translation.
- ▶ Meta's Ego4D builds first-person video datasets for next-gen assistants.

These examples illustrate rapid progress toward holistic AI understanding.

Emerging Trends

- ▶ Zero-shot multimodal reasoning
- ▶ Video-language models
- ▶ Real-time multimodal translation
- ▶ AR/VR integration
- ▶ Wearable AI with continuous sensory input

These innovations will define the next generation of AI systems.

Tools & Libraries for Developers

Key frameworks include:

- ▶ Hugging Face Transformers
- ▶ OpenCLIP
- ▶ TensorFlow Multimodal
- ▶ PyTorch Lightning
- ▶ LangChain with Vision or Speech models

These tools make multimodal AI accessible to researchers, students, and industry.

Business Impact

Multimodal AI is reshaping industries:

- ◆ **Retail** : Visual search + sentiment analytics
- ◆ **Banking** : Tone detection + behavioral analysis for risk modeling
- ◆ **Media** : Automated content summarisation, captioning, and tagging
- ◆ **Telecom** : Voice + image inputs for customer troubleshooting

Businesses adopting these tools gain efficiency, accuracy, and stronger customer engagement.

The Road Ahead

The future of AI will be multimodal, adaptive, and increasingly human-aware. Models will understand individual communication styles, emotional cues, and contextual signals with greater precision.

As hardware becomes faster and models more efficient, multimodal AI will integrate seamlessly into daily life—from classrooms to clinics, farms to workplaces.

Conclusion

Multimodal AI marks a new chapter in artificial intelligence—one where machines can perceive the world through multiple senses, enriching their understanding and expanding their capabilities.

This evolution is not just technological but human-centric. By bridging sight, sound, and language, multimodal AI brings us closer to AI systems that understand context, emotion, and intent just as we do.

For India's technology community, the rise of multimodal AI presents an extraordinary opportunity to lead innovation and shape the future of human-AI interaction.

TinyML

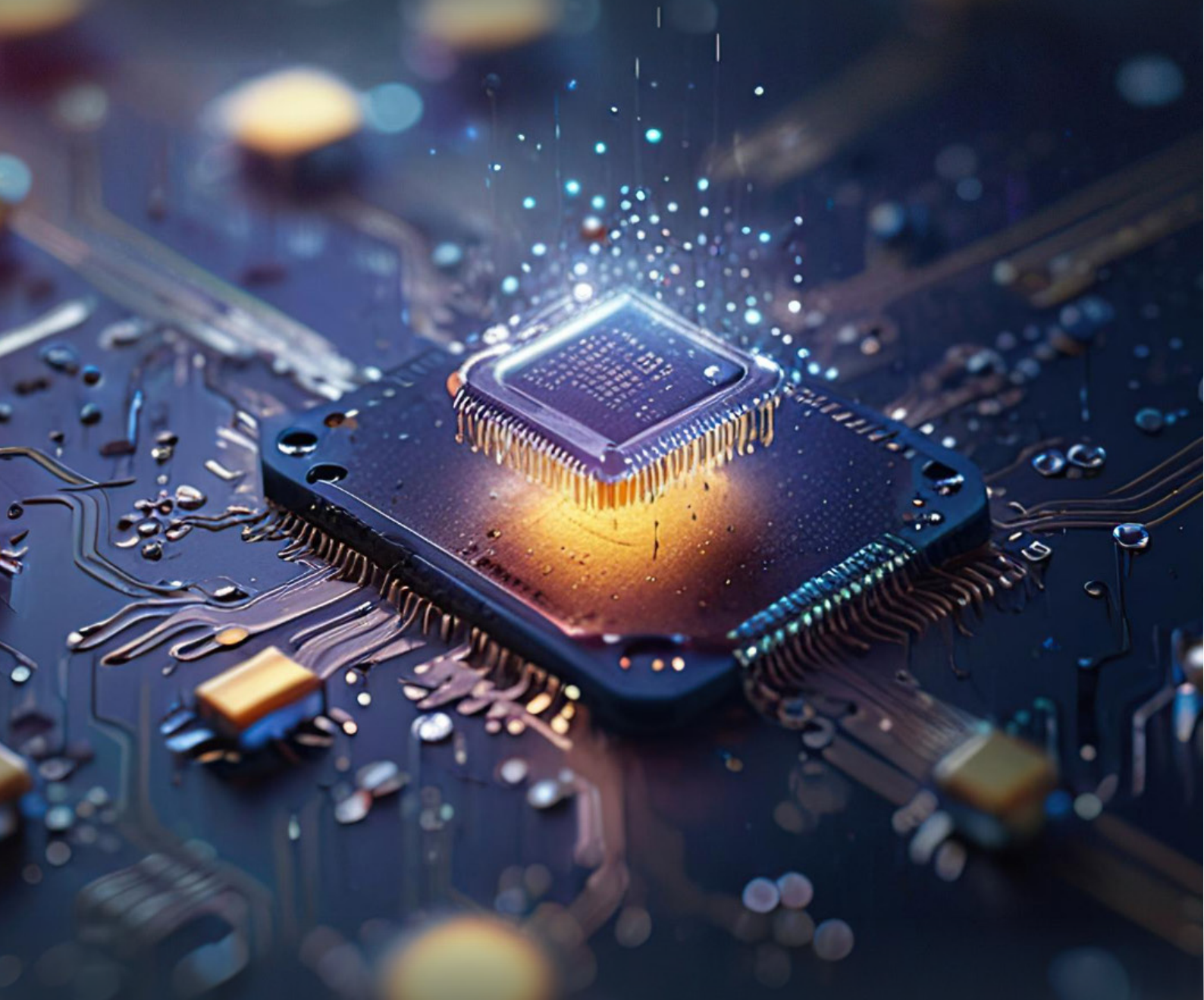
The Science of Machine Learning on Microcontrollers



Results-oriented With more than 5 years of expertise, the embedded engineer has designed, developed, and optimized hardware and firmware solutions for embedded systems. competent in RTOS, C/C++, and microcontroller-based development, with practical knowledge of circuit design, performance optimization, and debugging. competent in using communication protocols including CAN, SPI, I2C, and UART. devoted to developing scalable, dependable, and effective embedded solutions for a variety of uses, including industrial automation, automotive, and the Internet of Things.

TinyML represents a major shift in how machine learning models are designed, optimized, and deployed on resource-constrained hardware. Unlike traditional ML systems that rely on cloud servers or GPU-accelerated machines, TinyML operates directly on microcontrollers—devices that typically feature clock speeds below 200 MHz, RAM measured in kilobytes, and strict power budgets often limited to a few milliwatts. To make ML feasible in such environments, TinyML demands a combination of algorithmic compression, hardware-aware model design, and highly efficient runtime execution.

At the core of TinyML lies model optimization. Conventional neural networks are far too large for microcontroller memory, so several compression strategies are employed. Quantization reduces model weights and activations from 32-bit floating-point to 8-bit integers, or in some cases even down to binary or ternary values. This drastically decreases both memory footprint and computational load while enabling inference using integer arithmetic instead of floating-point operations. Pruning eliminates weights or even entire neurons that contribute minimally to the output, reducing the number of computations required per inference. Advanced approaches like knowledge distillation allow a large “teacher” model to train a smaller “student” model, transferring essential predictive behavior while keeping the model compact enough for microcontroller deployment.



Hardware architecture also plays a defining role in TinyML. Microcontrollers such as ARM Cortex-M0, M3, and M4 provide a combination of low power consumption and deterministic, real-time execution. The ARM Cortex-M4 and M7 series include DSP (Digital Signal Processing) extensions that accelerate operations like convolution, dot products, and matrix multiplications—core components of neural network inference. Many MCUs also include SIMD (Single Instruction, Multiple Data) instructions that allow parallel mathematical operations on multiple data elements with a single instruction cycle. Some modern microcontrollers integrate specialized accelerators or NPUs (Neural Processing Units), designed to execute quantized neural networks efficiently within milliwatt power budgets.

The TinyML software ecosystem is equally important. Frameworks such as TensorFlow Lite for Microcontrollers (TFLM) provide inference engines optimized for small memory regions, often requiring less than 30–50 KB of RAM. These runtimes avoid dynamic memory allocation, operate within statically allocated buffers, and perform kernel-level optimizations tailored for specific MCU architectures. The model conversion workflow involves training a conventional ML model, running post-training quantization, and compiling it into a C/C++ array that can be flashed directly into microcontroller memory. Libraries like CMSIS-NN further enhance performance by offering hand-optimized kernels for ARM Cortex-M processors, reducing cycle counts for convolution, depthwise convolution, fully connected layers, and activation functions.

On-device data processing is another fundamental component of TinyML. Because microcontrollers interface directly with sensors—accelerometers, microphones, temperature probes, and more—they must implement efficient signal preprocessing pipelines. Techniques such as FFTs, Mel-frequency cepstral coefficients (MFCCs), windowing, filtering, and feature normalization are performed within tight timing constraints. These operations must maintain low latency to support real-time inference while minimizing energy consumption. Developers often balance the complexity of preprocessing with the size of the neural network, ensuring that the overall workload fits within the MCU's computational and memory limits.



Power management strategies further distinguish TinyML from conventional ML workflows. MCUs often operate in deep-sleep modes and wake periodically for sampling, preprocessing, and inference. Because inference typically accounts for a major portion of energy usage, the number of computations per inference becomes a critical design constraint. Models are tailored to run within precise energy and timing budgets, ensuring long-term autonomous operation on coin-cell batteries or energy-harvesting sources.

TinyML continues to evolve through research on ultra-lightweight architectures such as MobileNet-Micro, Edge-LSTM variants, and binarized neural networks. Novel hardware platforms integrate event-driven or neuromorphic computing principles, allowing inference where computations occur only when new input data arrives, further reducing power consumption. As TinyML pushes machine learning into increasingly smaller and more power-efficient systems, it expands the boundary of what embedded intelligence can achieve, enabling continuous on-device processing while maintaining strict computational and energy constraints.



THANKS

Expert Panels

Ajith Surendran | Rakesh K C | Jomesh Jose | Vishnu

Magazine Editor

D A Anand

Content Editing

Harish Krishnan | Ankita Sharma | Paul Manuel | Jeyasooriya. J
Arya Saju | Abisekha Melfin | Navya Rajan | Muhammed Farhan
Malleshwaran | Chandana P | Karthikeyan

Design

Merin Sujith M R

Editing

D A Anand