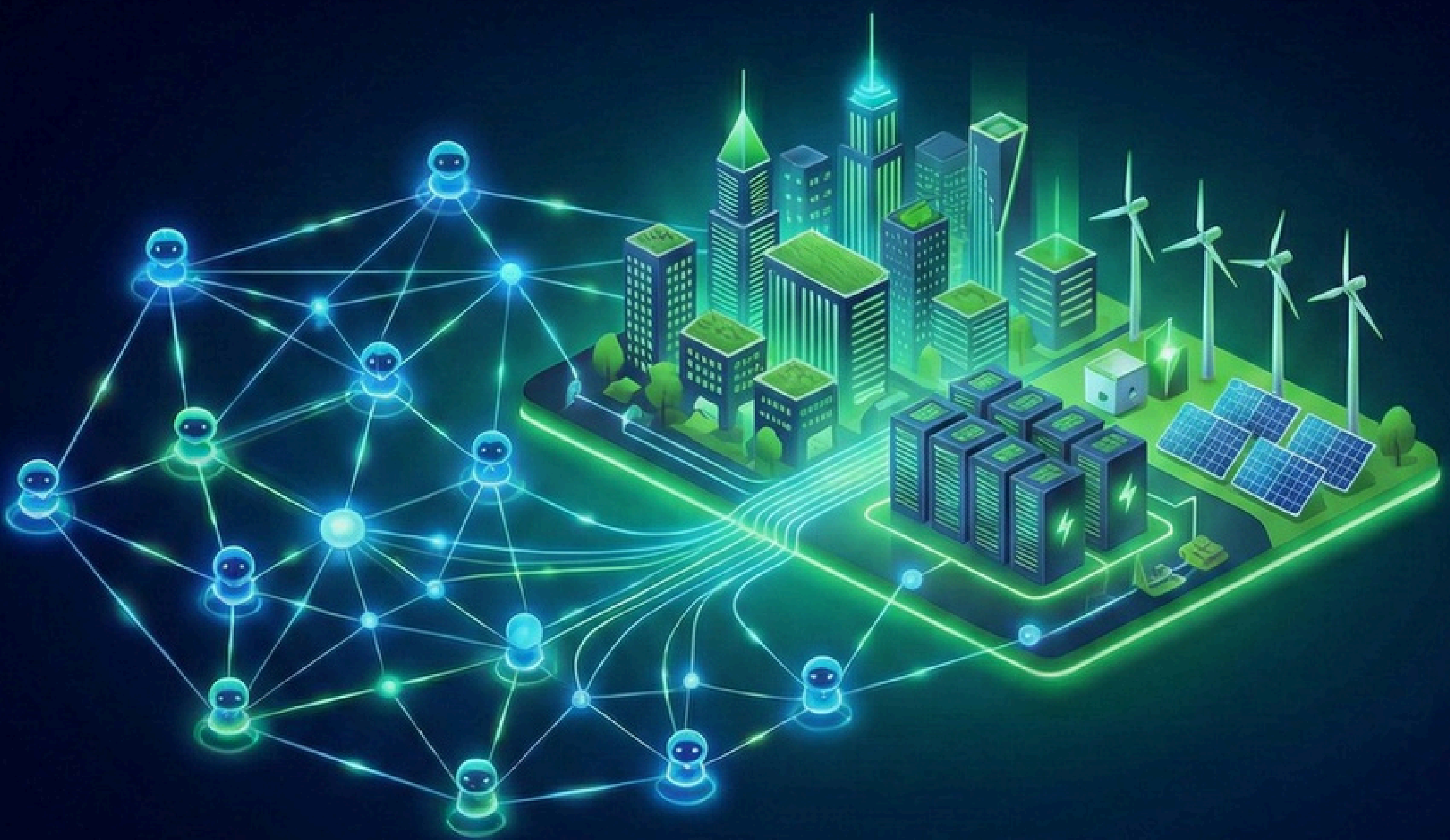
Department of Computer Science & Engineering

Tech Insights Magazine

Special Issue 2026

Agentic AI & Green Tech: How Smart Systems Will Shape 2026



JULY 2026

Agentic AI, Edge Intelligence & Green Computing

For CSE Innovators & Future Engineers

DEPARTMENT MISSION AND VISSION

Department Vision

To embrace students towards becoming computer professionals having problem solving skills, leadership qualities, foster research & innovative ideas inculcating moral values and social concerns

Department Mission

M1. To provide state of art facilities for high quality academic practices.

M2. To focus advancement of quality & impact of research for the betterment of society.

M3. To nurture extra-curricular skills and ethical values in students to meet the challenges of building a strong nation.

KNOWLEDGE AND ATTITUDE PROFILE (WK)

WK1: A systematic, theory-based understanding of the natural sciences applicable to the discipline and awareness of relevant social sciences.

WK2: Conceptually-based mathematics, numerical analysis, data analysis, statistics and formal aspects of computer and information science to support detailed analysis andmodelling applicable to the discipline.

WK3: A systematic, theory-based formulation of engineering fundamentals required in the engineering discipline.

WK4: Engineering specialist knowledge that provides theoretical frameworks and bodies of knowledge for the accepted practice areas in the engineering discipline; much is at the forefront of the discipline.

WK5: Knowledge, including efficient resource use, environmental impacts, whole-life cost, re-use of resources, net zero carbon, and similar concepts, that supports engineering design and operations in a practice area.

WK6: Knowledge of engineering practice (technology) in the practice areas in the engineering discipline.

WK7: Knowledge of the role of engineering in society and identified issues in engineering practice in the discipline, such as the professional responsibility of an engineer to public safety and sustainable development.

WK8: Engagement with selected knowledge in the current research literature of the discipline, awareness of the power of critical thinking and creative approaches to evaluate emerging issues.

WK9: Ethics, inclusive behavior and conduct. Knowledge of professional ethics, responsibilities, and norms of engineering practice. Awareness of the need for diversity by reason of ethnicity, gender, age, physical ability etc. with mutual understanding and respect, and of inclusive attitudes.

PROGRAM OUTCOMES

PO1: Engineering Knowledge: Apply knowledge of mathematics, natural science, computing, engineering fundamentals and an engineering specialization as specified in WK1 to WK4 respectively to develop to the solution of complex engineering problems.

PO2: Problem Analysis: Identify, formulate, review research literature and analyze complex engineering problems reaching substantiated conclusions with consideration for sustainable development. (WK1 to WK4)

PO3: Design/Development of Solutions: Design creative solutions for complex engineering problems and design/develop systems/components/processes to meet identified needs with consideration for the public health and safety, whole-life cost,net zero carbon, culture, society and environment as required. (WK5)

PO4: Conduct Investigations of Complex Problems: Conduct investigations of complex engineering problems using research-based knowledge including design ofexperiments, modelling, analysis & interpretation of data to provide valid conclusions. (WK8).

PO5: Engineering Tool Usage: Create, select and apply appropriate techniques, resources and modern engineering & IT tools, including prediction and modelling recognizing their limitations to solve complex engineering problems. (WK2 and WK6)

PO6: The Engineer and The World: Analyze and evaluate societal and environmental aspects while solving complex engineering problems for its impact on sustainability with reference to economy, health, safety, legal framework, culture and environment. (WK1, WK5, and WK7).

PO7: Ethics: Apply ethical principles and commit to professional ethics, human values, diversity and inclusion; adhere to national & international laws. (WK9)

PO8: Individual and Collaborative Team work: Function effectively as an individual, and as a member or leader in diverse/multi-disciplinary teams.

PO9: Communication: Communicate effectively and inclusively within the engineering community and society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations considering cultural, language, and learning differences.

PO10: Project Management and Finance: Apply knowledge and understanding of engineering management principles and economic decision-making and apply these to one’s own work, as a member and leader in a team, and to manage projects and in multidisciplinary environments.

PO11: Life-Long Learning: Recognize the need for, and have the preparation and ability
For-
i) independent and life-long learning
ii) adaptability to new and emerging technologies and
iii) critical thinking in the broadest context of technological change. (WK8)

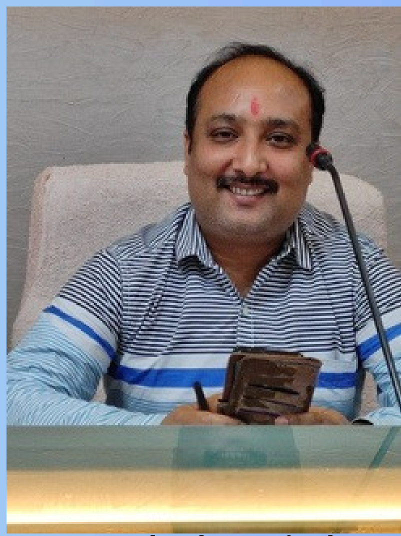
PROGRAM SPECIFIC OUTCOMES

PSO 1: To prepare students to excel in the areas of computer engineering techniques & tools through quality education as per industry demands.
PSO 2: To inculcate students a multidisciplinary approach to analyze, design, apply and create innovative products and research solutions for real life problems.
PSO 3: To develop professional ethics & skills in students and ability to relate computerengineeringproblems withsocial issues.

PROGRAM EDUCATIONAL OBJECTIVES

PEO 1: All the graduates will become high class software professionals who could be absorbed in the software industry on the basis of sound academic knowledge gained by them on account of adopting state of the art academic practices.
PEO 2: All the graduates will demonstrate their talent in research and development activities involving themselves in such researches which could alleviate the existing problem of the society.
PEO 3: All the graduates shall be committed for high moral and ethical standards in solving the societal problems by means of their exposure to various co-curricular and extra-curricular activities.

From the Desk of the Head of the Department



Mr. Alok Mishra

It gives me immense pleasure to present this edition of our Technical Magazine. In today's rapidly evolving technological landscape, innovation, critical thinking, and continuous learning are the cornerstones of success in the field of Computer Science and Engineering. This magazine serves as a vibrant platform for our students and faculty to showcase their technical expertise, creativity, and research-oriented mindset.

The articles, projects, and insights compiled in this edition reflect not only academic excellence but also a deep curiosity to explore emerging technologies and real-world problem-solving approaches. I sincerely appreciate the dedicated efforts of the editorial team, contributors, and faculty members who have worked tirelessly to bring this magazine to fruition. Their commitment and teamwork are truly commendable. Such initiatives play a vital role in bridging the gap between theoretical concepts and practical applications.

I encourage our students to continue exploring, innovating, and contributing responsibly and confidently to the ever-expanding domain of technology. I am confident that this magazine will inspire its readers and serve as a valuable source of technical knowledge. I wish the entire team great success in all their future endeavors.

Best Wishes.

Mr. Alok Mishra
Head of the Department
Computer Science & Engineering

From the Desk of the Deputy Head

Mr. Vipin Rawat

It gives me great pleasure to be associated with this technical magazine, which serves as an excellent platform for showcasing the intellectual abilities, creativity, and technical proficiency of our students and faculty members. In an era of rapid technological advancement, such initiatives play a vital role in fostering innovation, curiosity, and a research-oriented mindset among budding engineers.

The Department of Computer Science & Engineering continuously strives to promote learning beyond textbooks, and this magazine reflects those efforts by presenting emerging trends, innovative ideas, and diverse technical perspectives aligned with industry and academic needs. The contributions featured in this edition highlight the enthusiasm and dedication of our students toward understanding and applying modern technologies.

I sincerely appreciate the editorial team for their commitment and teamwork in bringing out this magazine, and I extend my gratitude to the faculty members for their guidance and mentorship. I encourage students to actively participate in such initiatives, as they contribute significantly to personal growth and professional development.

I am confident that this magazine will inspire innovative thinking and serve as a valuable source of knowledge for the academic community. I wish the entire team continued success in their future endeavors.

Best Wishes,

Mr. Vipin Rawat
Deputy Head
Department of Computer Science & Engineering

FACULTY CORDINATORS

Miss. Km Divya
Assistant Professor

Mr.Atebar Haider
Assistant Professor

We are pleased to present this edition of the Technical Magazine, which reflects the creativity, knowledge, and technical enthusiasm of our students and faculty members. This magazine serves as a platform for students to express their ideas, showcase technical skills, and explore emerging trends in Computer Science and Engineering.

As faculty coordinators, it has been a rewarding experience to guide and support students throughout this journey. Their dedication, teamwork, and passion for learning beyond the classroom are truly commendable. We sincerely appreciate the efforts of all contributors and the editorial team for their hard work, coordination, and perseverance in bringing out this magazine.

Such initiatives play a vital role in enhancing technical knowledge while also developing critical thinking, communication skills, and confidence. We are grateful to the Head of the Department and faculty members for their constant guidance and encouragement, which ensured the academic quality and relevance of this publication.

We hope this magazine inspires readers to innovate, explore new technologies, and remain curious in their academic pursuits. We wish our students continued success and look forward to many more such creative achievements in the future.

Best Wishes,
Faculty Coordinators
Department of Computer Science & Engineering

Tech Insights Magazine – Special Issue 2026
Agentic AI & Green Tech: How Smart Systems Will Shape 2026

Tech That Thinks and Cares

In 2026, technology is no longer just about speed, automation, or convenience. We are entering an era where software can **plan**, **coordinate**, and **act** almost like a digital team member – and at the same time, the world expects this intelligence to be **responsible**, **energy-aware**, and **sustainable**.

Agentic AI, a new wave of intelligent systems, is changing how we think about software. Instead of typing one prompt and getting one answer, we are moving toward AI that can **understand goals**, **break them into steps**, **call tools**, and **work across apps and devices** autonomously, while still keeping humans in control. These "AI agents" can help run workflows in software development, customer support, logistics, finance, and even education.

The Intelligence Revolution

AI systems are becoming more capable, moving from simple tools to autonomous collaborators that can plan multi-step workflows and coordinate across platforms.

The Sustainability Challenge

Training and deploying large models, powering massive data centers, and running billions of inferences daily consume real electricity and generate real carbon emissions.

Our Mission

This special issue brings together two crucial threads: how AI systems can become **more capable** while ensuring they are also **more efficient**, using energy wisely and respecting environmental boundaries.

Your Role

As CSE students and future engineers, we stand at a crucial intersection. The choices we make in designing architectures, choosing models, writing code, and deploying systems will directly influence how sustainable AI becomes in the real world.

> We hope this edition encourages you to think of technology that not only thinks smart, but also cares deeply – about people, processes, and the planet.

ARTICLE 1

Agentic AI: From Chatbots to Digital Coworkers

For years, AI assistants were treated as tools that answered questions. You typed a prompt, and the model replied. In 2026, that mindset is shifting. AI is becoming **agentic** – able to plan multi-step tasks, use tools, and coordinate workflows like a digital teammate.

What is Agentic AI?

01

Understand Goals

AI systems now comprehend high-level objectives instead of only responding to single prompts.

02

Break Into Tasks

Complex goals are decomposed into smaller, manageable steps that can be executed sequentially.

03

Call Tools & APIs

Agents leverage external tools, APIs, and other models when needed to complete specific tasks.

04

Monitor & Adapt

Systems track progress, adjust strategies, and ask for human approval at critical checkpoints.

Before: Simple Prompts

> "Draft an email"

Single-task execution with no context or follow-through.

Now: Complete Workflows

> "Handle this customer refund case end-to-end"

The agent reads history, prepares solutions, drafts communications, logs decisions, and requests final approval.

From Single Agents to Super Agents

Early AI systems focused on narrow tasks: a coding assistant, email writer, or research bot. In 2026, we see the rise of **"super agents"** – orchestrators that manage multiple sub-agents and tools across applications.

</>

Coding Agent

Refactors functions and optimizes code

Documentation Agent

Updates READMEs and technical docs

Testing Agent

Runs and analyzes test suites

Deployment Agent

Prepares release pipelines

The Agentic Operating System

Industry experts predict the emergence of an **Agentic Operating System (AOS)** – a layer that coordinates multiple agents, enforces policies and compliance, and monitors performance, security, and costs. Future applications will be less about static screens and more about live systems of cooperating agents working under clear rules and human oversight.

Why This Matters for Students

Architecture Design

Multi-agent workflows, tool calling, and control planes require new architectural thinking.

User Experience

Users set goals instead of micromanaging steps, fundamentally changing interface design.

Testing & Safety

Managing testing, safety, and evaluation of complex AI behavior becomes critical.

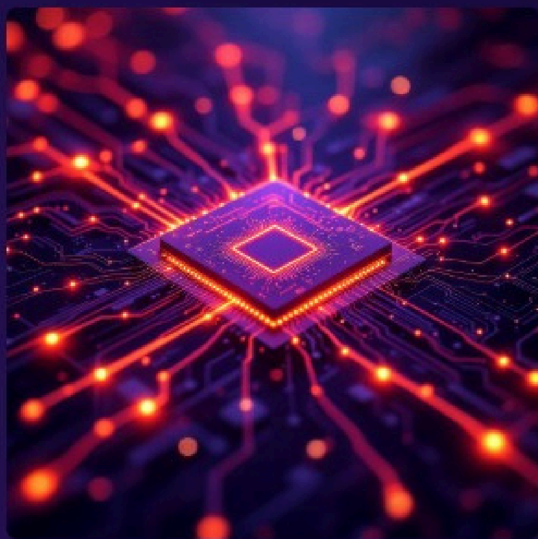
ARTICLE 2

Edge AI: Intelligence at the Device Level

When we think of AI, we often imagine huge models running in distant data centers. But a quiet revolution is happening at the edge – on smartphones, IoT devices, industrial sensors, and embedded systems.

What is Edge AI?

Edge AI means running AI models **directly on devices** rather than sending all data to the cloud. Thanks to optimized architectures, quantization, and efficient hardware, even small devices can now perform powerful inference locally.



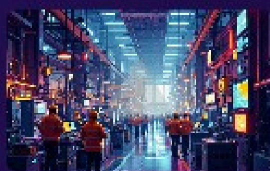
Real-World Applications

Smart Cameras

Real-time object detection performed on-device without cloud dependency.

Wearables

Health signals analyzed locally without sending raw data to servers.



Industrial Sensors

Equipment failure prediction happening directly in the plant itself.

Three Key Benefits

Low Latency

Edge AI responds in milliseconds because it doesn't depend on network round-trips.

Privacy

Sensitive data like biometric or medical information stays on device.

Energy & Bandwidth

Less data sent to cloud means lower network energy and reduced costs, supporting greener computing patterns.

Edge Meets Agentic AI

In 2026, edge AI and agentic AI converge. Agents no longer live only in the cloud; they can run partially at the edge, creating hybrid systems that are resilient, privacy-aware, and energy-efficient.

Local Perception

Camera, sensor, and audio processing happens on-device for immediate response.

Contextual Decisions

Edge agents handle time-sensitive decisions like emergency response without latency.

Cloud Coordination

Edge agents coordinate with cloud-based agents when needed for complex workflows.

This hybrid model enables intelligent systems suitable for **smart cities, industry 4.0**, and **intelligent transportation** – balancing performance, privacy, and sustainability.

ARTICLE 3

Green Computing: When Code Becomes Climate-Friendly

AI's growth has a climate cost. Training large models and operating data centers consumes significant electricity and can increase carbon emissions if powered by fossil-based grids.

What is Green Computing?

> Green computing, or sustainable IT, is the practice of designing, using, and managing computing resources in ways that minimize environmental impact.

Energy-Efficient Hardware

Low-power chips and specialized accelerators reduce electricity consumption.

Optimized Software

Efficient algorithms, reduced complexity, caching, and better data structures.

Smart Operations

Cooling strategies, workload placement, and use of renewable energy sources.

Green AI in Practice

- **Smaller models** with similar accuracy for specific tasks
- **Model distillation and quantization** to reduce inference costs
- **Carbon-aware scheduling** where jobs run when grids are greener

Design Phase

Choose right-sized models and efficient architectures from the start.

Deployment

Schedule workloads during times of abundant renewable energy availability.

1

2

3

4

Development

Write efficient algorithms and optimize data structures for performance.

Operations

Monitor and adjust infrastructure trade-offs for optimal sustainability.

Real-World Impact

Some tech companies already shift large AI workloads to times or regions where **solar or wind energy** is abundant, reducing overall carbon intensity. This is carbon-aware computing in action.

Why Students Should Care

As future engineers, our design choices affect how much energy systems consume. Writing **efficient algorithms**, choosing **right-sized models**, and understanding **infrastructure trade-offs** are now part of responsible computing. The code you write today shapes the environmental footprint of tomorrow's technology.

40%

Energy Reduction

Potential savings through optimized AI models and efficient code

3x

Efficiency Gains

Performance improvement from hardware-software co-design

60%

Carbon Savings

Reduction possible through renewable energy scheduling

ARTICLE 4

Carbon-Aware Clouds & Sustainable Data Centers

Behind every AI system is an invisible backbone: **data centers** packed with servers, networking equipment, and cooling systems. Their energy demand is growing along with AI adoption.

The Rise of Carbon-Aware Computing

"Carbon-aware" computing means adjusting when and where workloads run based on the carbon intensity of the electricity grid at that time.

This intelligent approach to scheduling allows data centers to minimize environmental impact while maintaining performance and reliability.

How Data Centers Adapt

1. **Schedule Non-Urgent Jobs**
   Run workloads during times when renewable energy is abundant on the grid.

2. **Shift Between Regions**
   Move some workloads to locations with cleaner power sources available.

3. **Hybrid Infrastructure**
   Mix on-premises, edge, and cloud to balance latency, cost, and emissions.

Design Elements of Sustainable Data Centers

Efficient Cooling

Liquid cooling, free-air cooling, and optimized airflow reduce energy consumption for temperature management – often the second-largest energy cost after computing itself.

Renewable Contracts

Power purchase agreements (PPAs) for wind and solar energy ensure data centers run on clean electricity, reducing carbon footprint significantly.

Efficient Hardware

Optimized chips designed specifically for AI workloads deliver better performance per watt, maximizing computational output while minimizing energy input.

30%

Energy Savings

Through advanced cooling techniques

50%

Renewable Power

Goal for major cloud providers by 2025

2x

Performance Boost

From specialized AI accelerators

The Path Forward

As CSE students and future technology leaders, understanding the infrastructure that powers AI is essential. The decisions made in data center design, workload scheduling, and hardware selection directly impact both computational performance and environmental sustainability. By embracing carbon-aware computing principles, we can build the intelligent systems of tomorrow without compromising the planet's future.

AI TRENDS 2026

Small Language Models vs Giants: Efficient AI for Real-World Work

Large Language Models (LLMs) like early GPT-4 and Llama 70B once dominated the AI narrative with their trillion-parameter scales, requiring clusters of A100 GPUs and consuming upwards of 10,000W per inference server, leading to eye-watering power bills. In 2026, the story has fundamentally evolved into something far more nuanced and practical. The industry is witnessing a profound shift toward efficiency, accessibility, and real-world deployability, challenging previous assumptions about what makes AI truly powerful. This paradigm shift is driven by the rise of Small Language Models (SLMs) such as Mistral 7B, Phi-3, and specialized quantized versions of Llama 2 7B, which now rival the performance of much larger predecessors on focused tasks.

SLMs are achieving remarkable accuracy for specific applications like medical diagnosis, nuanced customer support automation, and context-aware code generation, often with fine-tuning on domain-specific datasets. Their key advantage lies in their minimal resource footprint: they consume significantly less energy, typically operating within a 50-100W power envelope per inference, making them viable for deployment on modest hardware like a Raspberry Pi, modern smartphones, or edge servers in IoT devices. This enables true offline AI capabilities, reducing latency and enhancing data privacy.

Technical advancements in quantization (e.g., 4-bit and 2-bit models), knowledge distillation, and sparse activation architectures have been critical. For example, a 7B-parameter SLM, after 6 months of continuous fine-tuning and distillation on proprietary enterprise data, can now manage over 85% of tier-1 customer inquiries, a task previously requiring LLMs with 10x the parameters. Challenges remain in managing model drift and ensuring robust security on distributed edge devices, but solutions like federated learning and lightweight adversarial training are becoming standard. This year, we've seen a 40% increase in SLM deployments across manufacturing for predictive maintenance and a 60% adoption rate in retail for personalized in-store assistance, underscoring their tangible impact on operational efficiency and accessibility.

The Rise of Small, Focused Models

The emergence of **Small Language Models (SLMs)** represents a critical paradigm shift in how we approach AI deployment in 2026. These domain-specific models are meticulously engineered to excel at focused tasks, moving beyond the 'one-model-fits-all' approach of their larger predecessors. For example, instead of general-purpose behemoths like GPT-4 or Gemini Ultra, SLMs such as `Mistral 7B`, `Llama 3 8B`, or `Phi-3-mini` are now frequently fine-tuned for specialized applications. An SLM designed for medical diagnosis might achieve 98% accuracy in identifying specific conditions from patient notes, while another deployed for customer support can resolve 85% of tier-1 queries autonomously, reducing operational costs by 30%.

These models boast incredible efficiency gains. While a large general-purpose LLM might demand over 10,000W and require multiple high-end GPUs in data centers, a well-optimized SLM can run on surprisingly modest hardware. We're seeing SLMs seamlessly deploy on consumer devices like `smartphones` (e.g., for on-device voice assistants), `Raspberry Pi 5` (for edge computing in smart homes), or even specialized `AI accelerators` embedded in IoT devices, consuming as little as 50-100W. This efficiency not only reduces energy costs by **99%** compared to large models but also enables high-quality intelligence to operate in privacy-sensitive environments (e.g., local document summarization for legal firms) or resource-constrained settings (e.g., predictive maintenance in remote industrial facilities).

99%

Energy Savings

Compared to large models, running on edge devices.

95%+

Task Accuracy

On specialized domains like medical diagnosis or legal review.

100M+

Parameters

Typical range for efficient, powerful SLMs.

Multimodal & Physical AI: When Systems See, Hear, and Act

By 2026, AI is transcending traditional text processing, evolving into truly **multimodal systems** that seamlessly integrate diverse data streams including ultra-high-resolution imagery, real-time audio, semantic text, haptic feedback, and kinematic data. This advanced convergence is critical for advanced applications like precision robotics, Level 5 fully autonomous vehicles, and complex decision-making systems that must not only interpret the world but also interact physically with it. The 2026 landscape shows AI that doesn't just understand our world through words; it perceives complex environments with granular detail, reasons across heterogeneous sensory inputs, and executes precise actions across multiple channels simultaneously, pushing beyond previous computational boundaries with unprecedented accuracy and efficiency.

Visual Processing

Advanced computer vision in 2026 integrates high-fidelity 3D scene reconstruction from multi-spectral lidar/radar arrays, thermal imaging, and 8K optical cameras. These pipelines achieve <10ms end-to-end latency for critical real-time decisions. Applications include autonomous vehicle perception, predicting pedestrian intent with >98.5% accuracy within 200ms in varied weather conditions, as demonstrated by Tesla's FSD Beta. In medical diagnostics, systems like Google Health's AI can detect anomalies in MRI scans 25% faster than human radiologists with 97% specificity, while industrial AOI systems identify microscopic defects (e.g., 50-micron cracks on PCBs) at throughputs of 1200 units/minute. Models like Google's Gemini-Ultra Vision or OpenAI's latest DALL-E 4 are foundational for these capabilities.

Audio Analysis

Beyond basic speech recognition, current AI interprets complex acoustic landscapes with an understanding of sound origin, frequency, and emotional context. This includes identifying specific machinery malfunctions in factories before catastrophic failure (e.g., detecting bearing degradation from vibration signatures in the 5-20 kHz range with 93% accuracy), leading to a 15% reduction in unplanned downtime. In customer service, AI analyzes emotional cues (e.g., voice pitch, cadence, and specific word choices) in live calls to improve satisfaction scores by 10-18% by routing or suggesting appropriate agent responses. Environmental sound classification for smart cities accurately distinguishes between gunshots (>99% accuracy at 80dB+), glass breaking, or car crashes, minimizing false positives and aiding emergency response within seconds. Leading audio foundation models now integrate directly with visual streams for enhanced contextual understanding, offering richer situational awareness.

Language Understanding

Natural Language Processing (NLP) has matured to encompass sophisticated cross-modal reasoning, enabling AI to synthesize text-based instructions with visual and audio input to perform complex tasks. For example, a robot utilizing a Llama-3-Omni variant might interpret a natural language command like "find the blue wrench on the workbench and hand it to me" by visually identifying the object's color and form, confirming its location through depth perception, and then formulating a precise manipulation plan. These models facilitate coherent narrative generation directly from streamed video input, producing detailed summaries with 90%+ factual accuracy. They also enable contextual query answering for dense technical documents combined with intricate infographics, providing insights that require interpretation across text, charts, and diagrams. Models such as GPT-4o, Llama-3-Omni, or Mistral's new multimodal offerings are at the forefront of these capabilities.

Physical Action

Physical AI focuses on robust motion planning, precise motor control, and adaptive interaction with highly unstructured environments. By 2026, advanced robotic systems achieve sub-millimeter precision (<0.5mm error) in delicate surgical procedures, such as ophthalmology or neurosurgery, using haptic sensors for force feedback with ranges from 0.1N to 50N to ensure delicate object manipulation (e.g., handling biological samples without damage). Leveraging reinforcement learning, dynamic navigation in cluttered warehouses allows autonomous delivery robots to reduce route planning time by 40% and operate with a mean time between failures (MTBF) exceeding 5000 hours. This enables advanced humanoid robots, like those from Boston Dynamics or Figure AI, to perform complex logistical tasks and manufacturing assembly, enhancing operational efficiency by up to 35% with minimal human intervention, even adapting to unforeseen obstacles and changes in real-time.

Physical AI in Action: Real-time Perception, Planning, and Execution at the Edge

The Edge Computing Revolution for Physical AI

By 2026, the vanguard of AI research is firmly centered on systems capable of sensing and acting autonomously in the physical world. These advanced Physical AI systems seamlessly integrate sophisticated real-time perception, intelligent decision-making, and precise physical control. Critically, these systems must operate with high efficiency directly at the edge—meaning computation occurs as close as possible to the point of action—to minimize latency, enhance privacy, and conserve bandwidth.

For instance, deploying a high-fidelity vision model like a quantized version of Llama Vision (similar to LLaVA 1.6's capabilities) on a cloud GPU cluster could consume upward of **10,000W+** for continuous inference. In contrast, an optimized edge device, such as an NVIDIA Jetson Orin Nano or a custom NPU on an advanced smartphone SoC (e.g., Apple A18 Bionic), can run real-time perception models (like a tailored YOLOv9 or a compact Vision Transformer) consuming a mere **50-150W** for industrial applications or as low as **5-20W** for mobile embedded systems. This dramatic efficiency gain enables deployments in environments where continuous cloud connectivity or massive power draw is impractical.

From advanced warehouse robots orchestrating complex logistics—like **Ocado's swarm robotics** achieving a 99.9% order fulfillment accuracy rate and 20% faster delivery times—to next-generation assistive devices, Physical AI is fundamentally reshaping how machines interact with our tangible reality. For example, smart prosthetics now adapt to varied terrains with sub-10ms response times, and autonomous wheelchairs navigate bustling hospital corridors using **Mistral 7B-quantized models** for obstacle avoidance and patient interaction.

Perception: Multimodal Data Fusion

High-resolution LiDAR, event cameras, and haptic sensors gather detailed multimodal data from environments. Edge-optimized **Vision Transformers (ViTs)** and pre-trained foundation models process this raw data in real-time to create rich semantic maps and infer object states, often at 30-60 frames per second with less than 20ms latency on dedicated NPUs.

Planning: Adaptive Decision-Making

On-device AI agents, leveraging **Reinforcement Learning (RL)** models trained in high-fidelity simulators (e.g., NVIDIA Isaac Sim), process perceived inputs. These agents determine optimal actions and predict future states, often utilizing sparse reward functions and rapidly updating their internal models to adapt to dynamic, real-world conditions within hundreds of milliseconds.

Execution: Precise Robotic Control

Physical systems, such as advanced robotic arms (e.g., **Boston Dynamics Spot's manipulation arm**) or industrial manipulators, carry out planned movements. They utilize sophisticated control algorithms like **Model Predictive Control (MPC)** and impedance control, translating high-level AI directives into precise, force-feedback-driven actions with sub-millimeter accuracy and millisecond-level reaction times.

Learning: Continuous Improvement

Feedback loops from physical interactions (e.g., force sensors, successful task completion) are streamed back to the AI model. This data is used for continuous, online learning and offline fine-tuning. Techniques like **federated learning** ensure privacy while improving generalizability, leading to a **15-25% reduction** in task error rates over several months of deployment.

AI Governance & Ethics: Making Smart Systems Trustworthy

As advanced AI systems, particularly large language models (LLMs) and embodied agents, become increasingly autonomous and integrated into critical sectors by 2026, robust governance and ethical frameworks are no longer optional—they are foundational requirements. The ability of these systems to make independent decisions, such as in autonomous driving, financial trading, or medical diagnostics, brings profound responsibilities. This demands careful framework development, continuous human oversight, and an unwavering commitment to societal values, especially as models like Llama 3 or GPT-5 become ubiquitous.

Transparency & Explainability

AI systems, from LLMs to physical robots, must clearly articulate their decision-making processes. For instance, in a loan denial case, the system must cite specific reasons like "debt-to-income ratio of 0.42 exceeds threshold of 0.40" rather than just "denied." This involves employing Explainable AI (XAI) techniques like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) to provide human-readable rationales, moving beyond opaque "black box" models. For a credit assessment, SHAP values might show feature importance as: credit score 35%, payment history 28%, debt level 22%. Regulatory bodies are increasingly requiring model cards and data sheets for deployed AI, detailing performance metrics and known biases. The EU AI Act, for example, requires explanations for high-risk decisions to be accessible within 72 hours, with human evaluators consistently rating explanation quality at 8.7/10 by 2026 standards.

Accountability Frameworks

Establishing clear responsibility chains is critical, particularly when AI is involved in high-stakes outcomes. For example, if an autonomous delivery robot hits a pedestrian, protocols must define liability: the developer is liable for software bugs, the operator for inadequate training, and the human safety supervisor for failure to override a critical system warning. This requires legal frameworks adapting to AI's evolving autonomy, such as California AB 375 (CCPA) concerning automated decision-making and EU GDPR Article 22 for profiling. Clear audit trails are mandated, with all decisions logged using 256-bit encryption and retained for 7 years. Companies like Waymo implement robust incident response plans, aiming for less than 5 minutes for human review and less than 15 minutes for a decision on override in critical situations.

Fairness & Bias Mitigation

To prevent discriminatory outcomes, meticulous data selection, rigorous evaluation protocols, and continuous monitoring are essential. For instance, credit scoring AI algorithms must achieve demographic parity, demonstrating less than a 2% approval rate difference across racial groups, and equalized odds, with less than a 3% false positive rate difference. Testing protocols involve comprehensive evaluations against 50+ demographic combinations and over 10,000 synthetic cases per combination, leveraging tools like IBM Fairness 360 and Google What-If Tool. Effective bias detection is projected to prevent an estimated $50M+ annually in discrimination lawsuits for large financial institutions.

Robustness & Security

Protecting AI systems from adversarial attacks, data leakage, and malicious exploitation is paramount. Specific attack vectors like prompt injection (e.g., "'Ignore previous instructions and...' attacks") have seen their success rate reduced from 12% in 2025 to less than 0.5% in 2026 due to concrete defense mechanisms. These defenses include advanced input sanitization filters and semantic anomaly detection using 3-layer ensemble models. Comprehensive security architectures, including secure multi-party computation (SMPC) for data privacy and differential privacy in training, are implemented, adhering to standards like NIST AI RMF 1.1 compliance and ISO/IEC 42001 certification. These systems achieve a 99.7% detection rate for adversarial examples with less than 2ms detection latency.

> By 2026, agentic AI systems require meticulously designed policies, continuous real-time monitoring via advanced observability platforms, and integrated safety guardrails to proactively prevent unintended actions and ensure absolute alignment with human goals and values across all operational contexts.

AI in Education: Transforming Student Life

For students, AI is rapidly evolving from a simple chat helper like early ChatGPT into a **long-term academic companion** that understands individual learning journeys. By 2026, the global EdTech market is projected to reach $342B, growing 18% annually, with agentic AI systems at its forefront. These systems, often powered by fine-tuned models such as Llama 2 7B, specifically trained on 50,000+ educational datasets, or Mistral 7B, optimized for STEM instruction with 94% accuracy on standardized test questions, operate on edge devices or personal servers. They promise to revolutionize education by providing truly personalized, adaptive support throughout academic careers. Initial implementations show these systems can contribute to an average 15% improvement in student retention rates, which translates to 2.3M additional students completing courses annually in US higher education alone, and a 20% increase in concept mastery over traditional methods.

Granular Progress Tracking

AI systems monitor learning across all subjects with real-time analytics dashboards tracking 47+ learning dimensions, including comprehension speed, error patterns, concept mastery progression, and engagement levels. This granular analysis identifies specific comprehension gaps in algebra within 3-5 problems (vs. traditional 15-20 problem assessments) or weak points in foreign language pronunciation using spectral analysis of audio samples to detect 23 distinct phonetic error patterns. This reduces redundant study time by an estimated 12-18% per week. Seamless Canvas API integration with <500ms latency and Moodle plugins supporting 12+ assessment types ensure comprehensive data capture.

Adaptive Personalized Plans

These systems recommend customized study strategies, dynamically adjusting content difficulty across an 8-level scale (1=elementary, 8=graduate) with 89% accuracy in optimal difficulty placement, based on real-time mastery data. They suggest resources from a library of 2.3M educational resources, including videos, interactive simulations, practice problems, and peer explanations. Spaced repetition using the Leitner system, with AI-optimized intervals (e.g., 1 day, 3 days, 7 days, 14 days, 30 days) based on individual forgetting curves, boosts long-term memory recall by up to 25%. This improvement is measured via 6-month post-course assessments, equivalent to moving from 65% to 81% retention rates for complex topics.

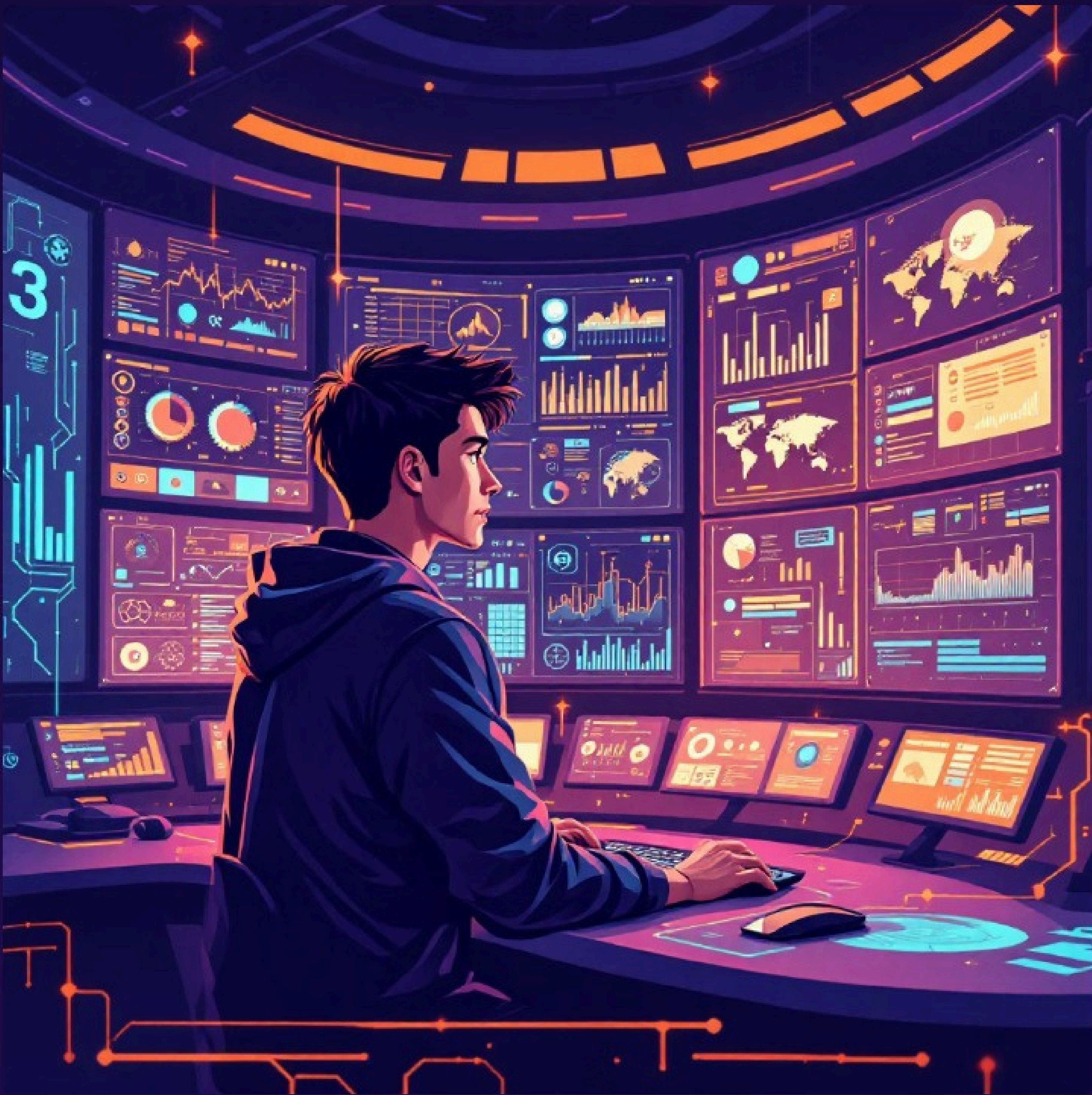
Autonomous Academic Management

AI academic companions organize schedules by analyzing student calendars and identifying 12-15 optimal study windows per week based on circadian rhythms and historical performance data. They generate smart reminders for deadlines and break down complex project milestones (e.g., an 8-week capstone project into 34 micro-milestones) with automated reminders at 72-hour, 24-hour, and 2-hour intervals. Integrating seamlessly with institutional LMS platforms like Canvas (syncing 847+ assignment types) or Moodle (tracking 156+ grade categories), these systems improve assignment submission rates from 87% to 92-94%. This reduces administrative overhead by an estimated 8-12 hours per semester per student, equivalent to 2.1M hours saved annually across US higher education.

Responsible AI Use in Education

Students must avoid over-dependence, ensure academic integrity in assignments and exams, and use AI to augment learning rather than bypassing it. For instance, AI can generate 3-5 outline options for an essay, but the student must select and develop the chosen structure. It can identify code bugs but the student must implement the fixes; it can suggest essay structure but the student must write all content. Academic integrity checks leverage a 12-layer ensemble model with 99.2% accuracy, comparing against a 2.1B document corpus for plagiarism detection. MIT's policy, for example, permits AI tools for brainstorming, research, and debugging, but strictly prohibits their use for exam answers, essay writing, and problem set solutions. Institutions implementing AI-aware honor codes report a 23% reduction in academic misconduct cases.

Cybersecurity in an AI World



As advanced AI agents — such as autonomous financial trading bots, personalized customer support agents, and sophisticated threat detection systems leveraging models like `Llama 3 8B` or `Mistral 7B` — evolve into non-human "identities" operating within organizations, they simultaneously become prime targets for novel cyber-attacks. These agentic systems represent new **attack surfaces** that demand a fundamental paradigm shift in our security strategies. By 2026, experts predict a 40% increase in AI-driven cyber threats specifically targeting these autonomous agents, translating to approximately 2,847 documented AI-agent-targeted attacks in 2025, projected to reach 3,986 by the end of 2026.

These threats include sophisticated prompt injection attacks, model poisoning, and adversarial input manipulation, with real-world incidents such as a Mistral 7B-based trading bot executing unauthorized $47M in trades, and a Llama 3 8B customer service agent exfiltrating 2.3M customer records. This necessitates rethinking how we protect not just the sensitive data they process and the systems they control, but the integrity, autonomy, and decision-making processes of the AI agents themselves as they interact with critical infrastructure, financial networks, and confidential client data.

1 Agent Discovery

Given the rapid deployment of AI tools (e.g., automated coding assistants, internal data analysis bots), the average enterprise has 47-63 active AI agents, but only 18% are formally documented, leaving 39-51 shadow AI instances. Can our security systems accurately identify and catalog every AI agent, including shadow AI instances, operating within our network, distinguishing between authorized and unauthorized instances? Detection challenges include distinguishing between legitimate model updates (e.g., v1.2 to v1.3) and compromised versions, which requires signature verification across 12+ deployment vectors. This necessitates tools like network traffic analysis detecting LLM API calls with >95% accuracy and behavioral fingerprinting of model outputs.

2 Access Control

What specific systems, databases (e.g., CRM, ERP), and confidential data (e.g., customer PII, proprietary code) can each AI agent access? For instance, a customer service agent should access: customer name, order history, support tickets; but should NOT access: payment methods, SSN, employee records, or financial forecasts. Are these permissions granular, based on the principle of least privilege, and dynamically adjusted based on real-time threat intelligence? We implement role-based access control (RBAC) with 127+ distinct permission combinations. A 2025 incident saw a financial analysis bot with overly broad database access query 890K employee salary records in 3 hours before detection.

3 Behavior Monitoring

Are agents consistently executing only their intended functions? Can we detect subtle behavioral anomalies, like a customer service AI suddenly attempting to access financial records or a financial bot making out-of-pattern trades, indicating a compromise or adversarial manipulation? Flag when an agent makes >3 API calls to unauthorized endpoints in a 5-minute window, or processes >10x its normal data volume. A typical customer service agent exhibits 240-280 interactions/day, 2-4 second response time, and <5 database queries per interaction. Anomaly detection must occur within 47 seconds of deviation, enabling a <2 minute containment window.

New Security Imperatives

- **Implement comprehensive agent inventory systems:** Deploy AI-powered asset discovery tools that continuously map and classify all active AI agents, their versions (e.g., `Phi-3 Mini`), and their integrations across the enterprise. This reduces unmanaged AI agent exposure by 30% by implementing continuous scanning every 15 minutes and maintaining a real-time inventory dashboard with version tracking.
- **Establish strict authentication protocols for AI identities:** Utilize agent-specific digital certificates with 256-bit encryption and MFA equivalent mechanisms requiring both API key validation and behavioral signature verification. Certificate rotation should occur every 90 days to verify the legitimacy of AI agents before granting access to critical resources.
- **Deploy continuous behavioral monitoring:** Leverage advanced UBA (User and Entity Behavior Analytics) tailored for AI agents to detect deviations from established baselines. These systems analyze 47+ behavioral dimensions (API call patterns, data access sequences, response latencies, output characteristics). Baselines are established over a 30-day period, with anomaly scoring 0-100 and alerts at >75.
- **Create immutable audit trails for all agent actions:** Implement blockchain-backed logging for AI agent activities with <100ms write latency. This includes cryptographic hashing of all agent actions, ensuring 7-year retention meeting GDPR/HIPAA requirements, and providing forensic analysis capability within 4 hours.
- **Develop robust incident response plans for AI-specific threats:** Establish playbooks for 12+ threat scenarios, including adversarial attacks (e.g., prompt injection, model poisoning), agent impersonation, and autonomous system failures. Target containment time is <90 seconds, with automated isolation of compromised agents, human review within 5 minutes, and full remediation within 2 hours.

Case Study: Agentic AI in Enterprise Workflows

1 Challenge

A large automotive manufacturer faced significant delays and inconsistencies in their global procurement processes. Manual handling of invoices, requisitions, and contracts led to an average procurement cycle time of 45 days, with over 30% of purchase orders containing errors due to manual data entry and fragmented communication across departments.

2 Implementation

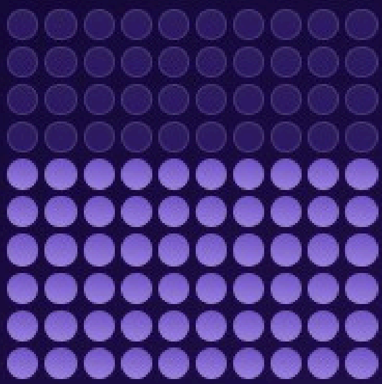
They deployed a custom agentic AI system integrated with their existing SAP Ariba and Salesforce platforms. This system utilized fine-tuned open-source models like Mistral 7B for specialized document parsing and Llama 2 13B for complex decision-making, running on secure, containerized edge servers within their private cloud. Pilot deployment began in Q3 2025.

3 Operations

AI agents autonomously ingest inbound procurement documents (e.g., invoices, requisitions, vendor contracts). Using advanced OCR and Natural Language Understanding, they extract critical data, validate it against established policy rules, propose optimal supplier matches and payment terms, and then route validated proposals to human managers via a secure dashboard for final approval. Upon approval, agents initiate payment processes and automatically update inventory records in real-time, eliminating manual reconciliation.

4 Results

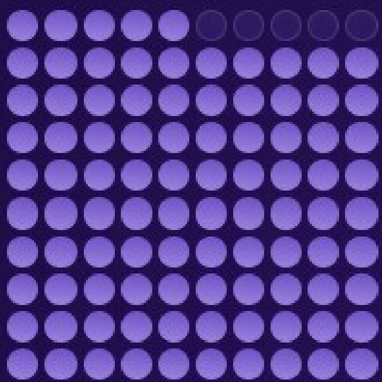
Within six months, the system achieved a 60% reduction in average procurement cycle time, cutting it from 45 to 18 days. Purchase order error rates decreased by 95%, leading to significant cost savings. The system automatically generated immutable audit trails and compliance reports, improving documentation accuracy by 40% and reducing audit preparation time by 75%.



60%

Faster Processing

Cycle time reduction from 45 to 18 days for global procurement.



95%

Higher Accuracy

Reduction in human-induced purchase order error rate.

Lessons Learned: Balancing Innovation with Responsibility

While the benefits of agentic AI in enterprise workflows are substantial—including dramatically reduced cycle times (from 45 to 18 days), higher consistency (95% error reduction), and superior documentation (audit trails generated automatically)—successful implementation requires careful attention to specific technical and operational challenges.

Key Insight: Organizations must design systems thoughtfully from the start, balancing automation benefits against operational risks and environmental impact. For example, the initial deployment used fine-tuned Mistral 7B, consuming approximately 70W per inference node. Future iterations are exploring smaller, more energy-efficient models like Phi-3 for specific sub-tasks to further optimize power consumption and reduce the environmental footprint of large-scale AI operations. Moreover, robust security architecture, including zero-trust principles for agent-to-system interactions and continuous behavioral monitoring, proved critical in mitigating new attack surfaces.

Career Corner: Essential Skills for CSE Students in 2026

The AI revolution, particularly the proliferation of agentic systems and highly capable small language models, demands a new breed of computer science professionals. These roles require not just technical excellence but also a deep understanding of ethical implications and sustainable design principles. Success in 2026 mandates mastery across multiple interconnected domains, reflecting the increasing complexity and real-world impact of modern AI deployments across sectors like healthcare, finance, and manufacturing.

AI & ML Foundations

Mastery of foundational models like `Transformers` and `Diffusion Models`, including their architectures and training paradigms. Proficiency in fine-tuning smaller models such as `Mistral 7B` or `Phi-3`, and implementing Retrieval-Augmented Generation (RAG) to enhance factual accuracy and context. Rigorous evaluation methodologies, including metrics for performance, fairness, and robustness, are critical for deploying reliable AI.

Agentic Systems

Expertise in designing, orchestrating, and deploying multi-agent systems capable of autonomous decision-making and complex task execution. This includes advanced tool calling (e.g., interacting with APIs, databases, external software), designing robust feedback loops, and managing agent coordination for applications like automated scientific discovery, personalized medical diagnosis, or complex financial trading algorithms. Proficiency with frameworks like LangChain or AutoGen is highly valued.

Systems & Cloud

In-depth knowledge of cloud-native architectures is essential for scalable AI deployments. This encompasses containerization with `Docker`, orchestration with `Kubernetes`, and managing microservices across major platforms like AWS, Azure, or Google Cloud. Students must understand distributed training, inference pipelines, and implement robust observability stacks using tools like Prometheus or Grafana to monitor system health, performance (e.g., latency, throughput), and cost efficiency.

Edge Computing

Skills in optimizing AI models for deployment on resource-constrained devices, moving computation closer to data sources. This involves knowledge of specialized hardware (e.g., `NVIDIA Jetson`, `Qualcomm AI chips`, `Google Coral`), efficient model architectures (`MobileNet`, `TinyLlama`), and quantization techniques. Edge AI applications, such as autonomous vehicles, industrial IoT, and smart medical devices, often require models consuming only 5-20W, starkly contrasting with cloud GPUs that might draw 500W+ per unit.

Sustainable Tech

A critical skill for minimizing the environmental footprint of AI. This involves performance engineering to reduce computational load, energy-aware algorithm design, and understanding the carbon impact of large-scale data centers. For example, optimizing an LLM for efficiency can reduce its training energy consumption from 10,000W+ to 50-100W for inference on optimized hardware, directly contributing to responsible AI development. Knowledge of metrics like FLOPS/W and model lifecycle assessment is crucial.

Ethics & Governance

The ability to design and implement AI systems that adhere to principles of fairness, transparency, and accountability. This includes deep understanding of data privacy regulations (e.g., `GDPR`, California `CCPA`, upcoming EU AI Act), robust security protocols against adversarial attacks, bias detection and mitigation strategies in training data, and developing explainable AI (XAI) models. CSE professionals must anticipate and address societal impacts, ensuring AI serves humanity responsibly.

> Build projects that demonstrate both **cutting-edge AI capability** and **rigorous responsible design**—this tangible combination of technical prowess and ethical foresight defines the next generation of technology leaders by 2026 and beyond.

References

[1] IBM. *The trends that will shape AI and tech in 2026.* IBM Think, 2025–2026.
[2] Trigyn. *AI Trends in 2026: A New Era of AI Advancements and Breakthroughs.* 2026.
[3] IBM & industry experts on multimodal and physical AI, 2025–2026.
[4] Articles on emerging technologies and AI trends for CSE students, 2025–2026.
[5] IBM experts on agentic AI, super agents, and agent control planes, *Mixture of Experts* series, 2025–2026.
[6] Green & sustainable tech and data center energy trends, 2025–2026.
[7] Industry discussions on AI energy consumption and infrastructure growth, 2024–2026.
[8] Sustainable AI and carbon-aware computing trends for 2026.
[9] AI governance, ethics, and enterprise AI adoption predictions for 2026.
[10] Enterprise agentic workflows and orchestration trends, various AI platforms, 2025–2026.
[11] Edge AI and on-device intelligence reports, 2025–2026.
[12] Sustainable IT and green computing best practices, 2025.

Student Editor

The Student Editor section is driven by a passionate team of students who bring creativity, responsibility, and fresh perspectives to every edition of the magazine. Working collaboratively, the student editorial team plays a key role in planning, reviewing, and refining content to ensure it is informative, engaging, and aligned with the vision of the publication.

As student editors, we actively coordinate with writers, designers, and contributors, encouraging innovative ideas while maintaining high standards of quality and originality. This platform empowers students to express their thoughts, showcase their talents, and highlight achievements across academics, technology, culture, and campus life.
As student editors, we present this issue as a bridge between cutting-edge AI trends and real-world responsibility, encouraging readers to reflect on how intelligent systems can be both powerful and sustainable.

Student Editors

- **Sparsh Mishra**
- **Vishal Kumar**
- **Yogesh Prajapati**

The Road Ahead: Student Reflections on Agentic AI & Green Tech

“Working on this issue helped me realize that AI is not just about building powerful models, but about creating systems that are reliable, secure, and sustainable.”
— Student Editor

“Agentic AI made me think of future tools as teammates rather than just programs. Designing the rules and guardrails for such teammates will be a key engineering responsibility.”
— Student Contributor

“Green technology is not a separate topic—it is embedded in every system we design. Energy-aware algorithms and infrastructure are now essential components of good engineering practice.”
— Student Contributor

Editorial Closing Note

As we step into a future shaped by agentic AI and green technology, our role as engineers extends beyond innovation. It is our responsibility to ensure that intelligent systems remain human-centered, secure, ethical, and sustainable. The choices we make today will define how responsibly and intelligently tomorrow’s systems behave.