Applied Agentic AI for Organizational Transformation

Glossary

Module 1

A/B Testing: An experimentation method where two variations (A and B) are compared to see which performs better—commonly used in digital marketing, UX design, and AI-driven personalization.

Agentic AI: An AI system that autonomously sets and pursues goals with limited human supervision, coordinating multiple specialized agents to complete complex tasks.

AI Agent: A software entity capable of independently performing tasks or making decisions on behalf of a user or system.

AI Assistant: An interactive software agent, often powered by generative and predictive AI, that supports users with tasks, information retrieval, and decision-making.

API (Application Programming Interface): A set of rules and tools that allow software systems to communicate. In AI, APIs are commonly used to access models (like GPT or DALL·E) from other applications.

Caching: A technique where frequently used data or responses are temporarily stored for faster access, reducing the need to recompute or fetch the same information repeatedly.

Chatbot: A conversational software application designed to interact with users via text or speech, often leveraging Generative AI and natural language processing to simulate human dialogue.

Chain-of-Thought Prompting: A prompting method that guides language models to produce step-by-step reasoning chains, thereby simulating the human thought process for problem solving.

Data Centers: Large facilities that house servers and infrastructure used to store, process, and run cloud-based services—including AI models and digital tools.

Deep Learning: A branch of machine learning using neural networks with many layers to learn hierarchical representations of data, excelling at tasks like image and speech recognition.

Digital Transformation: The process by which organizations use digital technologies—like AI, cloud computing, and automation—to fundamentally change operations, services, or customer experiences.

Discriminator (GANs): The second part of a GAN, the discriminator evaluates data and decides whether it is real (from training data) or fake (from the generator), pushing the generator to improve.

Fine-Tuning: The process of taking a pre-trained AI model and training it further on a smaller, domain-specific dataset to improve its performance for a particular task or brand.

Generative Pre-trained Transformer (GPT) Model: A transformer-based language model pre-trained on large amounts of text and fine-tuned for specific tasks

Generator: In a Generative Adversarial Network, the generator is the neural network that tries to create fake (but realistic) data—like images—to fool the discriminator.

Graphics Processing Units (GPUs): Specialized processors originally designed for rendering graphics but now essential for training and running AI models due to their ability to handle massive parallel computations.

Hallucination: The phenomenon where AI models generate incorrect or false information that appears plausible but has no basis in the training data.

Large Language Models: Deep-learning models trained on extensive datasets to understand and generate human-like language across diverse contexts.

Long Short-Term Memory (LSTM): A special kind of RNN (see definition below) designed to remember longer sequences by controlling what to keep or forget—often used in natural language tasks before transformer models became dominant.

Machine Learning: A subset of AI where algorithms learn from data to make predictions or decisions without explicit programming

Mode Collapse: A limitation in GANs where the generator starts producing only a narrow set of outputs, reducing creativity and diversity in the results.

Natural Language Processing: The field of AI concerned with enabling machines to understand, interpret, and generate human language.

Neural Networks: Computational models composed of interconnected layers of nodes ("neurons") that process input data and adjust internal weights to recognize patterns and make classifications

Predictive AI: AI systems focused on forecasting future outcomes—such as customer behavior or equipment failures—by analyzing historical data and identifying trends.

Pre-Training: The initial training phase where a large language model ingests vast, general-purpose data (e.g. books, articles) to learn broad patterns before any specialization

Probablistic: Describing systems or models that use probability distributions to represent uncertainty and make predictions based on likelihoods.

Prompt Engineering: The practice of designing and refining model inputs (prompts) to elicit desired outputs, ensuring clarity in instructions and context.

Recurrent Neural Networks (RNNs): A type of neural network designed to handle sequential data by passing memory from one step to the next—used in early AI models for tasks like speech and language.

Reinforcement Learning: A learning paradigm where AI models learn optimal behaviors by receiving rewards or penalties from interactions within an environment

Self-Attention: A mechanism within transformer architecture where each token in a sequence weighs every other one in parallel to decide which ones matter most for understanding its meaning

Tokens: Units of text (often 3–4 characters or 0.75 words) that language models process. Both your input and the AI's output are measured in tokens, which affect cost and model limits.

Turing Test: A test proposed by Alan Turing in 1950 to determine if a machine's conversational behavior is indistinguishable from that of a human evaluator.

Vectors: A mathematical representation—often a list of numbers—that encodes features of data (such as words or images) for processing by machine-learning models.

Weights: Numeric parameters in neural networks that are adjusted during training to minimize errors and determine the strength of connections between neurons.