

Day 1: Embeddings & Semantic Search

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
<ul style="list-style-type: none">■ Improving GenAI Search■ Vector Embeddings Overview	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Vertex Embeddings• Pinecone• FAISS• HuggingFace

Day 1: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 2: Embedding Task Types

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Embedding Task Types	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Gemini Embeddings• SentenceTransformers• OpenAI Embeddings

Day 2: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 3: Vector Databases & ANN Search

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Vector Search Overview	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Qdrant• Weaviate• Milvus• Vertex Vector Search

Day 3: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 4: RAG Fundamentals

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Vertex RAG Engine	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Vertex RAG Engine• LangChain• LlamaIndex

Day 4: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 5: Chunking & Document Structuring

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Chunking Guide	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Cohere Chunking• Document AI• Haystack

Day 5: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 6: Hybrid Search

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Hybrid Search Intro	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Elastic + Vectors• BM25• Hybrid Pinecone

Day 6: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 7: Prompt Engineering

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Prompt Engineering Guide	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Gemini• GPT-4• Claude• PromptLayer

Day 7: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 8: Grounding & Hallucination Reduction

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Responsible Prompting	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• NeMo Guardrails• Google Safety APIs• Anthropic Safety

Day 8: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 9: Agentic AI

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Agentic RAG	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• LangGraph• Vertex Agents• OpenAI Assistants

Day 9: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 10: Guardrails & Safety

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ AI Safety Guardrails	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• DLP APIs• Guardrails AI• PII Classifiers

Day 10: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 11: Evaluation & Observability

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Evaluating GenAI Systems	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Arize Phoenix• WhyLabs• DeepEval

Day 11: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 12: AI for Financial Services

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ AI in Financial Services	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• GCP FS Suite• Snowflake• BigQuery

Day 12: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 13: AI Strategy & Architecture

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Google AI Blog	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Kubernetes• BigQuery• Vertex AI

Day 13: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization

Day 14: Final Integration Review

<i>Blog Links</i>	<i>Task</i>	<i>Tech Stack</i>
■ Enterprise Architecture	■ Complete today's reading and document 5 insights.	<ul style="list-style-type: none">• Cloud Run• Monitoring• CI/CD for AI

Day 14: Comprehensive Buzzword Glossary

- Semantic Search — Meaning-based retrieval
- Dense Retrieval — Embedding-driven search
- Sparse Retrieval — Keyword/BM25 search
- Hybrid Search — Combining dense + sparse
- Embedding Space — Vector representation
- Cosine Similarity — Measuring vector angle
- Dot Product Scoring — Vector magnitude method
- ANN — Approximate nearest neighbor
- HNSW — Hierarchical graph index
- Chunk Overlap — Preserving context
- Tokenization — Splitting into tokens
- Context Window — Max LLM tokens
- Grounding — Evidence-based output
- Hallucination — Unsupported LLM answer
- RAG — Retrieval-augmented generation
- Reranking — Cross-encoder scoring
- Fusion Scoring — Weighted blending
- Vector Normalization — Unit scaling
- Sharding — Splitting vector index
- Embedding Drift — Distribution shift
- Chain-of-Thought — Structured reasoning
- Toolcalling — LLM executing tools
- Agent Workflow — Multi-step reasoning
- Guardrails — Enforced constraints
- Safety Classifier — Blocks harmful text
- Faithfulness — Evidence consistency
- Relevance@K — Retrieval quality metric
- Precision — Correctness of hits
- Recall — Coverage of relevant docs
- Latency Optimization — Speed tuning
- Vector Quantization — Compression
- Dimensionality Reduction — Lowering vector size
- Semantic Clustering — Grouping meaning
- Embedding Task Types — Similarity/classification
- Metadata Filtering — Attribute filters
- BM25 — Lexical ranking algorithm
- RAG Pipeline — Retrieve + generate
- Cold Start Problem — Missing embeddings
- Prompt Engineering — Input optimization