The Ultimate Data & AI Taxonomy — Master List

> A comprehensive, hierarchical index of concepts, keywords, and specialties across Data Science, AI/ML/DL, GenAI, and the full product lifecycle (data → modeling → training → evaluation → deployment → ops → governance). Use as a checklist to master the field.

---

0) How to use this list

- **Levels**: (F) Foundational, (C) Core, (A) Advanced/Specialized, (O) Ops/Platform, (G) GenAI-specific.
- Each bullet packs **keywords/synonyms** in parentheses. Treat each as a search seed.
- Starred ★ items are **must-know** for senior roles.

---

1) Mathematical & Computational Foundations (F)

- **Probability & Statistics ★** (random variables, distributions, expectation, variance, covariance, Bayes' rule, conditional independence, CLT, LLN, sufficiency, MLE, MAP, conjugate priors, hypothesis testing, p-values, CIs, power analysis, multiple testing/Bonferroni, false discovery rate, bootstrap, jackknife, robust stats)
- **Linear Algebra ★** (vectors, matrices, norms, inner products, eigendecomposition, SVD, PCA, orthogonality, projections, low-rank, Kronecker, tensor algebra)
- **Calculus & Real Analysis** (limits, continuity, derivatives, gradients, Jacobian/Hessian, chain rule, Taylor, convexity, Lagrange multipliers, measure basics)
- **Optimization ★** (convex/non-convex, gradient descent/SGD, momentum, Nesterov, Newton/Quasi-Newton, L-BFGS, proximal methods, coordinate descent, line search, KKT, duality, ADMM)
- **Information Theory** (entropy, cross-entropy, KL/JS divergence, mutual information, MDL)
- **Numerical Methods** (floating point, conditioning, stability, autodiff, mixed precision, ODE/PDE solvers basics)
- **Graph Theory & Networks** (graphs, paths, cuts, centrality, spectral methods)
- **Computation & Systems** (complexity classes, data structures, caching, vectorization, SIMD, GPU/TPU basics, memory hierarchy)

---

2) Data Management & Engineering (C, O)

- **Data Modeling & Storage** (relational modeling, star/snowflake schemas, OLTP vs OLAP, columnar stores, time-series DBs, graph DBs, document stores, key-value stores)
- **Data Lakes & Warehouses** (data lakehouse, Parquet/ORC, Delta/Iceberg/Hudi, partitioning, Z-ordering, table formats)
- **ETL/ELT Pipelines ★** (batch vs streaming, CDC, micro-batching, orchestration)
- **Streaming & Eventing** (Kafka/Pulsar, event time vs processing time, watermarks, windowing)
- **Data Quality ★** (profiling, constraints, great expectations-style checks, anomaly detection in data, freshness/completeness/consistency/uniqueness/validity, data contracts)

- **Metadata & Lineage** (data catalogs, lineage graphs, column-level lineage, schema evolution)
- **Master/Reference Data** (MDM, deduplication, entity resolution/record linkage)
- **Governance & Compliance** (data classification, retention, access policies, consent, right to be forgotten)
- **Privacy & PII Handling** (tokenization, hashing, pseudonymization, DLP scanning)
- **Security** (encryption at rest/in transit, IAM, KMS, key rotation, secrets management)
- **File Formats** (CSV/JSON/Avro/Parquet/ORC, image/audio/video codecs; TFRecords, WebDataset)
- **Big Data Compute** (Spark/Ray/Flink concepts, shuffle, skew, join strategies)

---

3) Data Acquisition, Labeling & Preparation (C, G)

- **Data Collection** (instrumentation, telemetry, logging schemas, sampling strategies)
- **Labeling** (annotation tools, guidelines, gold sets, inter-annotator agreement/Cohen’s kappa)
- **Weak Supervision** (label functions, programmatic labeling, Snorkel-like paradigms)
- **Distant/Self-Supervision** (heuristic/metadata-derived labels, pseudo-labeling)
- **Active Learning** (uncertainty sampling, diversity sampling, core-set, human-in-the-loop)
- **Synthetic Data & Augmentation ★** (image/audio/text augmentations, MixUp/CutMix, back-translation, paraphrasing, diffusion-synthesized data)
- **Data Cleaning** (missing data—MCAR/MAR/MNAR, imputation, outlier detection, dedup, normalization/standardization, winsorization, feature hashing)
- **Feature Engineering ★** (target encoding, interactions, binning, polynomial features, time-based features, aggregations, lag/rolling windows)
- **Feature Stores** (offline/online parity, point-in-time correctness, backfills)
- **Imbalance Handling** (class weights, resampling, SMOTE/ADASYN, focal losses)
- **Bias & Leakage Checks** (train/test contamination, proxy variables, survivorship bias)
- **Train/Val/Test Splits** (stratification, group splits, temporal splits, nested CV)

---

4) Classical Machine Learning (C)

- **Supervised Learning ★**
- Regression (linear, ridge/lasso/elastic net, GLMs, Poisson/Gamma, quantile/regression with pinball loss)
- Classification (logistic regression, naive Bayes, kNN, SVMs—C/ν, kernels, margins)
- Trees & Ensembles ★ (CART, random forests, gradient boosting—GBDT/XGBoost/LightGBM/CatBoost concepts, stacking/blending)
- **Unsupervised Learning** (k-means/k-medoids, GMM/EM, hierarchical clustering, DBSCAN/HDBSCAN, spectral clustering)
- **Dimensionality Reduction** (PCA/SVD, ICA, NMF, t-SNE/UMAP, autoencoders as DR)
- **Anomaly/Novelty Detection** (one-class SVM, isolation forests, LOF, density estimation)
- **Metric Learning** (triplet/contrastive losses, Siamese nets, prototypical networks)
- **Graphical Models** (Bayesian networks, MRF/CRF, HMMs, factor graphs)
- **Bayesian Inference** (priors/posteriors, MCMC—Metropolis-Hastings, HMC/NUTS; variational inference)

- **Causal Inference ★** (Pearl's do-calculus, DAGs, backdoor/frontdoor, propensity scores, IVs, DiD, RDD, synthetic controls, uplift modeling)
- **Time Series** (ARIMA/SARIMA, ETS, state-space, Kalman/HMM, VAR, STL, TBATS, multivariate, seasonality, trend, change-point detection)
- **Reinforcement Learning (classical)** (MDP, policy/value functions, DP, TD, Q-learning, policy gradients, actor-critic, exploration vs exploitation)
- **Recommender Systems ★** (explicit/implicit feedback, MF, ALS/BPR, content-based, session-based/sequential, bandits, re-ranking)
- **Information Retrieval Basics** (indexing, TF–IDF, BM25, query likelihood, NDCG/MAP)

5) Deep Learning Core (C, A)

- **Neural Network Essentials ★** (perceptron, MLP, activations—ReLU/GELU/SiLU, initialization—Xavier/He, normalization—Batch/Layer/Group, dropout, residual connections, skip connections)
- **Optimization in DL ★** (SGD/momentum, Adam/AdamW, RMSProp, learning-rate schedules—cosine/one-cycle, warmup, early stopping, gradient clipping, weight decay, label smoothing)
- **Loss Functions** (CE/BCE, focal, hinge, Huber, triplet/contrastive, Dice/IoU, CTC, Wasserstein, InfoNCE)
- **CNNs** (conv/pool, dilations, depthwise separable, attention in CNNs)
- **Sequence Models** (RNN/LSTM/GRU, temporal conv nets, attention, Transformers ★)
- **Transformers Internals ★** (self-attention, multi-head, Q/K/V, positional encodings—sinusoidal/RoPE/ALiBi, pre-norm vs post-norm, residual streams, FFN/GLU)
- **Efficiency Tricks** (Flash Attention, memory-efficient attention, KV-cache, paged attention, speculative decoding, prefix caching, continuous batching)
- **Regularization & Robustness** (mixup/cutmix, stochastic depth, data augmentation policies, adversarial training/FGSM/PGD, uncertainty—ensembles/MC Dropout)
- **Model Compression** (pruning—magnitude/structured, distillation—soft labels/logit matching, low-rank adapters, tensor decomposition)
- **Quantization** (PTQ/QAT, INT8/INT4/FP8, GPTQ/AWQ/NF4, activation-aware methods)
- **Distributed Training ★** (data/model/pipeline parallel, ZeRO/sharded, FSDP/DP, gradient checkpointing, parameter/optimizer offload, mixed precision—FP16/BF16)
- **Hardware** (GPU/TPU tiles, memory bandwidth vs compute, NUMA, interconnects—NVLink/InfiniBand, storage I/O, SSD/NVMe, CPU vectorization)

6) Natural Language Processing (beyond LLMs) (C)

- **Text Processing** (tokenization, stemming/lemmatization, n-grams, stopwords)
- **Traditional NLP** (POS tagging, parsing, NER, coref, topic modeling—LDA/NMF, keyword extraction—TF–IDF/TextRank)
- **Embeddings** (Word2Vec/CBOW/Skip-gram, GloVe, FastText, subword models)
- **Seq2Seq** (encoder–decoder, attention before Transformers)
- **Speech & Audio NLP** (ASR basics, diarization, keyword spotting)

7) Computer Vision (C, A)

- **Core Tasks** ★ (classification, detection—anchor-based/anchor-free, segmentation—semantic/instance/panoptic, keypoints/pose, tracking, OCR, super-resolution, image retrieval)
- **Architectures** (ResNet, DenseNet, EfficientNet, MobileNet, Vision Transformers/ViT, Swin, DETR family)
- **3D/Geometry** (stereo, SfM, SLAM, point clouds, meshes, NeRF/3DGS, differentiable rendering)
- **Video** (temporal conv, video transformers, action recognition, tracking)
- **Medical/Industrial CV** (DICOM, multi-modal fusion, defect detection)

---

8) Time Series, Forecasting & Anomaly Detection (C)

- **Classical & ML** (ARIMA/SARIMA, ETS, Prophet-like ideas, gradient-boosted TS, feature-based TS)
- **Deep TS** (temporal fusion transformers, sequence-to-sequence, attention-based)
- **Anomaly/Change-point** (CUSUM, Bayesian change-point, matrix profiles)
- **Evaluation** (MAE/MAPE/SMAPE, pinball loss, coverage/CRPS)

---

9) Recommenders, Search & IR (C, A)

- **Retrieval** (BM25, dense retrieval, dual encoders, ColBERT-style late interaction, hybrid retrieval, lexical expansion)
- **Indexing** (HNSW, IVF-Flat/IVF-PQ, scalar/product quantization, ANN libraries)
- **Ranking** (pointwise/pairwise/listwise—LambdaRank/LambdaMART, learning-to-rank)
- **Session/Sequential & Contextual** (GRU4Rec, Transformers4Rec, bandits/contextual bandits)
- **Evaluation** (Recall\@k/Precision\@k, NDCG/MAP, CTR/CVR/Lift, calibration)

---

10) Generative AI & LLMs (G, A)

- **Tokenization & Vocab** (BPE, Unigram LM, SentencePiece, byte-level)
- **Pretraining Objectives** (causal LM, masked LM, seq2seq denoising, contrastive multimodal)
- **Architectural Variants** (decoder-only, encoder–decoder, Mixture-of-Experts—Sparse MoE, routing/load balancing, MoE inference)
- **Scaling Laws** (model/data/compute scaling, Chinchilla-like data–param tradeoffs)
- **Long-Context Techniques** (RoPE/ALiBi/YaRN/NTK scaling, chunked attention, memory layers)
- **Multimodal** (VLMs/VLAs: image+text; audio+text; video+text; 3D+text; speech TTS/ASR integration)
- **Instruction Tuning (SFT)** ★ (datasets, preference mixtures, data dedup/quality)
- **Parameter-Efficient Finetuning (PEFT)** ★ (LoRA/QLoRA/Adapters, prefix/prompt tuning)
- **Preference Optimization** (RLHF, RLAIF, DPO/IPO/ORPO/KTO/SimPO; reward models; pairwise/ranking-based)
- **RAG** ★ (chunking strategies—fixed/sliding/semantic/hybrid; embedding models; retrievers—sparse/dense/hybrid; re-rankers/cross-encoders; index structures—HNSW/IVF/PQ; document segmentation; query expansion—HyDE; answer synthesis; citations/grounding)

- **Agents & Tool Use** (function/tool calling, program synthesis, planners—ReAct/ToT/GoT, memory/episodic stores, multi-agent orchestration, tool routers, safety-aware planning)
- **Inference Systems ★** (KV cache management, paged attention, streaming, speculative decoding, batching, beam/greedy/sampling—top-k/p, temperature, nucleus, repetition penalties)
- **Compression & Deployment** (quantization—AWQ/GPTQ/NF4, distillation, pruning; server frameworks; offline vs online serving; Triton/TensorRT-like compilation concepts)
- **Safety & Trust** (jailbreaks/prompt injection, data exfiltration, output filtering, toxicity/bias/groundedness evals, watermarking/signatures, provenance/C2PA, content labeling)
- **LLM/GenAI Evaluation** (perplexity, pass\@k, BLEU/ROUGE/METEOR, BERTScore, MMLU/GSM8K/HumanEval/MBPP/HellaSwag/Winogrande/TruthfulQA/BIOsafety; judge models; pairwise/tournament)

---

11) Training Mechanics & Recipes (C, A, G)

- **Data Curriculum** (curriculum/self-paced learning, data curation, deduplication, filtering)
- **Batching & Samplers** (class-balanced, temperature sampling, dynamic bucketing)
- **Initialization & Schedules** (warmup, cosine/one-cycle, restarts, EMA)
- **Stability** (gradient clipping, loss scaling, anomaly detection, numerical stability)
- **Regularization** (weight decay, dropout, stochastic depth, label smoothing, early stopping)
- **Multi-task & Transfer** (MTL, adapters, fine-tuning protocols, domain adaptation)
- **Distributed/Parallel** (DDP/FSDP, pipeline, ZeRO, sharding, checkpointing, elastic training)
- **Mixed Precision** (FP16/BF16/FP8; autocast; loss-scaler)
- **Checkpointing** (periodic, best-k, EMA weights, state dicts, sharded checkpoints, safetensors)

---

12) Evaluation & Validation (C)

- **Validation Design ★** (temporal splits, group-aware, nested CV)
- **Metrics by Task**
- Regression (R^2, RMSE/MAE, MAPE/SMAPE, pinball loss/quantiles)
- Classification (accuracy, precision/recall/F1, ROC/PR AUC, log loss, MCC, Brier score)
- Ranking/IR (Precision\@k/Recall\@k, MAP, MRR, NDCG, coverage, diversity)
- Clustering/DR (silhouette, Davies–Bouldin, ARI/NMI, trustworthiness/continuity)
- Segmentation/Detection (IoU/Dice, mAP, AP@[.5:.95])
- Time Series (MAE/MAPE/SMAPE, MASE, CRPS, coverage, calibration)
- LLM/GenAI (perplexity, BLEU/ROUGE/BERTScore, exact match/F1, task leaderboards, groundedness/hallucination, toxicity, bias, jailbreak resistance)
- **Statistical Testing ★** (A/B tests, sequential tests, CUPED, variance reduction, non-parametric tests, multiple hypothesis control)
- **Calibration & Uncertainty** (Platt/Isotonic, temperature scaling, conformal prediction, prediction intervals)
- **Error Analysis** (slices, confusion matrix, counterfactuals, adversarial probes)

---

13) Productionization & Serving (O)

- **Serving Patterns ★** (batch, online, streaming, real-time/bidirectional—WebSockets/gRPC)
- **Interfaces** (REST/gRPC, message queues, serverless, edge, mobile)
- **Model Packaging & Artifacts** (ONNX, TorchScript, XLA/TFRT, MLIR, TVM; model registries, versioning, signatures, schemas)
- **Inference Optimization ★** (tensor RT/graph compilers, operator fusion, KV cache, quantization, distillation, dynamic shapes, pinned memory)
- **Scalability** (horizontal vs vertical scaling, autoscaling/HPA, sharding, load balancing, multiplexing, request batching, caching—embedding/prefix)
- **Latency/Throughput Engineering** (p50/p95/p99, tail latencies, backpressure, circuit breakers)
- **Canary/Shadow/Blue-Green** (progressive delivery, rollback strategies)
- **Monitoring & Observability ★** (metrics/logs/traces, dashboards, SLO/SLA, golden signals, anomaly detection, drift monitoring—data/concept; feature skew)
- **Feedback Loops & CT/CI** (continuous training, online learning, human-in-the-loop validation)
- **Cost Engineering** (GPU utilization, spot vs on-demand, throughput per dollar, right-sizing)

---

14) MLOps & LLMOps (O)

- **Experiment Tracking ★** (runs, params, artifacts, lineage, reproducibility)
- **Data & Model Versioning** (DVC-like concepts, semver for models, model cards, datasheets)
- **Pipelines & Orchestration ★** (Airflow/Kubeflow/Ray/Flyte patterns; DAGs; caching; retries; idempotency)
- **Feature Stores** (offline↔online consistency, TTLs, point-in-time joins)
- **Model Registry & Promotion** (staging/prod, approvals, rollbacks)
- **Testing** (unit/integration, data tests, regression tests, golden sets, load tests, chaos)
- **Security in MLOps** (supply chain, SBOM, signed artifacts, provenance, secret rotation)
- **LLM-Specific Ops** (prompt/versioning, prompt tests, eval harnesses, safety filters, tool catalogs, agent orchestration, memory stores)

---

15) Safety, Privacy, & Security (C, O, G)

- **Privacy** (k-anonymity/l-diversity/t-closeness, differential privacy—DP-SGD, PATE; federated learning; secure aggregation)
- **Security** (model stealing, inversion, membership inference, data poisoning, backdoors/trojans)
- **Adversarial Robustness** (FGSM/PGD/CW, certified defenses, randomized smoothing)
- **LLM Threats** (prompt injection, indirect injection, jailbreaking, data exfiltration, prompt leaking)
- **Cryptography for ML** (homomorphic encryption, MPC, TEEs/SGX/SEV)
- **Watermarking & Provenance** (model/output watermarks, C2PA/content credentials)

---

16) Governance, Ethics & Regulation (O)

- **Responsible AI Frameworks** (fairness, accountability, transparency, interpretability, safety)
- **Fairness Metrics** (demographic parity, equalized odds/opportunity, calibration within groups)
- **Explainability** (global vs local; SHAP/LIME, feature attribution, counterfactuals, surrogate models)
- **Documentation** (model cards, data statements, risk assessments, TIAs)
- **Regulatory Landscapes** (GDPR/CCPA/DPDP, EU AI Act risk tiers, sectoral rules—HIPAA/PCI)
- **Audits & Red-Team** (bias audits, security red-team, model eval audits)

---

17) Domains & Modalities (A)

- **NLP** (text mining, knowledge graphs, IE, QA, MT, summarization)
- **Vision** (medical imaging, remote sensing, retail, industrial)
- **Speech/Audio** (ASR/TTS, enhancement, music IR)
- **Structured/Tabular** (credit risk, churn, fraud, marketing mix)
- **Geospatial** (GIS, raster/vector, tiling, coordinate systems, route optimization)
- **Healthcare** (ICD/SNOMED, de-identification, survival analysis, causal safety)
- **Finance** (market microstructure, risk, compliance, time-series extremes)
- **Cybersecurity** (IDS/IPS, anomaly detection, threat intel, SOC automation)
- **IoT/Edge** (embedded ML, quantization for MCUs, TinyML)
- **Robotics & Control** (SLAM, planners, model-based RL, sim2real)

---

18) Retrieval-Augmented Systems & Knowledge (G)

- **Ingestion** (connectors, crawling, sitemap parsing, rate limiting, deduplication, boilerplate removal)
- **Chunking** (semantic/sentence-aware, windowed overlap, table/diagram handling, code-aware chunking, multimodal chunking)
- **Embedding Stores** (vector DBs, HNSW/IVF/PQ, disk vs memory indices, hybrid—BM25+dense)
- **Retrievers** (kNN, MMR/recency/semantic diversity, learning-to-retrieve)
- **Re-ranking** (cross-encoders, late interaction)
- **Grounding & Citation** (span highlighting, provenance, quote extraction)
- **Knowledge Graphs** (RDF/property graphs, entity linking, graph embeddings)
- **Evaluation** (retrieval precision/recall, faithfulness/groundedness, answer quality, latency/throughput)

---

19) Product Thinking, Experimentation & Analytics (C)

- **Problem Framing & Scoping ★** (objectives, constraints, success metrics/KPIs, ROI)
- **Cohorting & Segmentation** (RFM, clustering for personas, LTV)
- **A/B/n & Multi-armed Bandits** (exploration policies, regret, Thompson sampling)
- **Causal Uplift & Personalization** (heterogeneous treatment effects)
- **Analytics Engineering** (dbt-like concepts, semantic layers, metrics layers)

- **Storytelling & Viz** (EDA, dashboards, experiments readouts, uncertainty communication)

---

20) Lifecycle & Project Management (O)

- **Frameworks** (CRISP-DM, OSEMN, ML Canvas, DS lifecycle, MLOps loop)
- **Backlogs & Roadmaps** (epics/stories, prioritization—RICE/ICE, risks/assumptions)
- **Documentation & Runbooks** (playbooks, on-call, incident response)
- **Release Management** (model approvals, sign-offs, change control, versioning policies)
- **Postmortems & RCA** (blameless postmortems, five whys, fishbone)

---

21) Tooling Ecosystem (neutral concepts) (O)

- **Languages** (Python, R, SQL, JVM basics, C++ interop)
- **DL Framework Concepts** (autograd graphs, modules/layers, jit/compilers, distributed runtimes)
- **Data Tools Concepts** (orchestration, feature store, labeling, catalog, lineage, monitoring)
- **Serving & Inference Runtimes** (model servers, inference compilers, vector DB serving, LLM serving engines)
- **Visualization** (grammar of graphics, interactive viz, large-scale viz)

---

22) Edge, Mobile & Embedded ML (A)

- **On-Device ML** (TFLite/CoreML/ONNX-runtime concepts, NNAPI/Metal, GPU vs NPU)
- **Compression for Edge** (quantization/pruning/distillation pipelines, sparsity)
- **Latency/Power Trade-offs** (thermal throttling, offline mode, model updates)

---

23) Robotics, Control & Advanced RL (A)

- **RL Algorithms** (DQN family, DDPG/TD3, PPO/TRPO/SAC, distributional RL)
- **Model-Based RL** (world models, MPC, planning with learned dynamics)
- **Imitation & Offline RL** (behavior cloning, D4RL concepts, CQL/IQL)
- **Simulators & Toolchains** (sim2real gap, domain randomization)

---

24) Emerging Topics & Trends (A, G)

- **Diffusion Models ★** (denoising diffusion, latent diffusion, guidance, control—ControlNet)
- **Video Generation & Editing** (temporal consistency, keyframe control)
- **3D/NeRF & Gaussian Splatting** (scene representation, novel view synthesis)
- **Program Synthesis & Code LLMs** (compilers context, static/dynamic analysis integration)
- **AutoML & NAS** (search spaces, weight sharing, Hyperband/BOHB)

- **Neurosymbolic & Reasoning** (symbolic integration, logic, theorem proving)
- **Bio/Protein ML** (structure prediction, diffusion for molecules)

25) Checklists (Quick Scan)

- **Senior Must-Know (★):** probability/statistics, linear algebra, optimization, feature engineering, trees/boosting, Transformers, training stability, distributed training, evaluation & statistical testing, serving patterns, inference optimization, monitoring, RAG, instruction tuning/PEFT, preference optimization, security/privacy basics, responsible AI, A/B testing, product metrics.
- **Ops Must-Know:** experiment tracking, data/model versioning, orchestration, registries, CI/CD, canary/shadow, observability, drift, cost engineering, incident response.
- **Safety Must-Know:** jailbreak/prompt injection defenses, groundedness evals, privacy/PII handling, compliance checklists.

26) Glossary Seeds (non-exhaustive but dense)

- **Attention** (self, cross, causal, linear-time, multi-query)
- **Batching** (static, dynamic, continuous)
- **Calibration** (probability calibration, temperature scaling)
- **Contrastive** (InfoNCE, NT-Xent, SimCLR, CLIP)
- **Decoding** (greedy, beam, top-k, nucleus, repetition penalty)
- **Drift** (data, concept, covariate, prior probability shift)
- **Explainability** (SHAP, LIME, saliency, counterfactuals)
- **Groundedness** (faithfulness, citation support)
- **Indexing** (HNSW, IVF, PQ, OPQ)
- **KV Cache** (reuse, paging, eviction)
- **Long-Context** (RoPE, ALiBi, NTK scaling, YaRN)
- **MoE** (sparse experts, router, load balancing, token-level routing)
- **Normalization** (Batch/Layer/Group, RMSNorm)
- **Quantization** (PTQ, QAT, INT8/4, FP8, AWQ, GPTQ, NF4)
- **Regularization** (dropout, weight decay, label smoothing, mixup, cutmix)
- **Retrieval** (BM25, dense, hybrid, re-ranking)
- **Safety** (toxicity filters, jailbreak detection, prompt shields)
- **Speculative Decoding** (draft–verify, assisted generation)

Final Note

This taxonomy is meant as your **master checklist**. As you study, attach links, examples, math, and code under each bullet. When you're ready, we can export this as CSV/Notion or expand any section into a deep-dive guide with math, pseudo-code, and best practices.