AI and Machine Learning in Online Advertising

An Overview from 30,000 Feet

10/08/2025

Diko Ko / Director of Engineering, Toss USA / diko@toss.im

Agenda

Market Scale and Historical Context

Where ML/AI Plays in Online Ads

ML/AI for DSP Ad Retrieval

Conclusion

Online Advertising Market Landscape

- Global Market Size (2025 Estimate): ~**$740B** (source: eMarketer)
- Major Segments
  - Search Advertising
  - Retail / Commerce Media
  - Display Advertising
  - CTV
  - Social Media Ads
- Estimated Breakdown (eMarketer, et. al.)
  - Retail Media / Commerce Ads: 20-25%, ~$160B
  - Non-Retail Search Ads: 20-25%, ~ $160B
  - Display / Social / Other (CTV): 50-60%, ~$420B

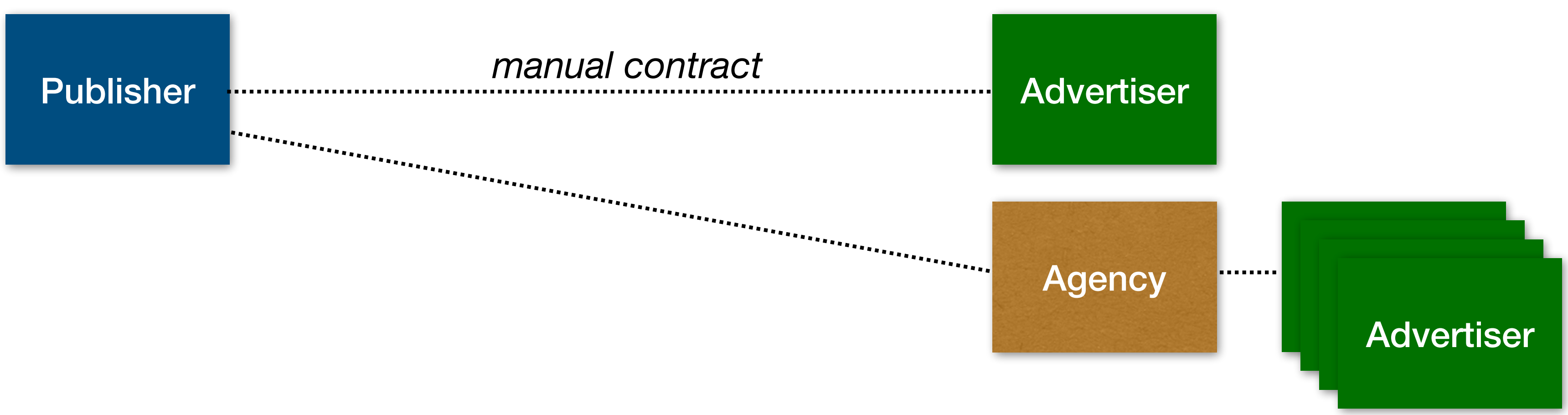
Historical Evolution

- 1990s: Banners, early search ads (Yahoo!, Altavista)
- 2000s: Google AdWords — CPC, measurable ROI
- 2010s: Programmatic buying, Real-time Bidding (RTB)
- 2020s: Privacy-first Era—retail media, contextual AI

Direct-Sales Online Ads (1990s-early 2000s)

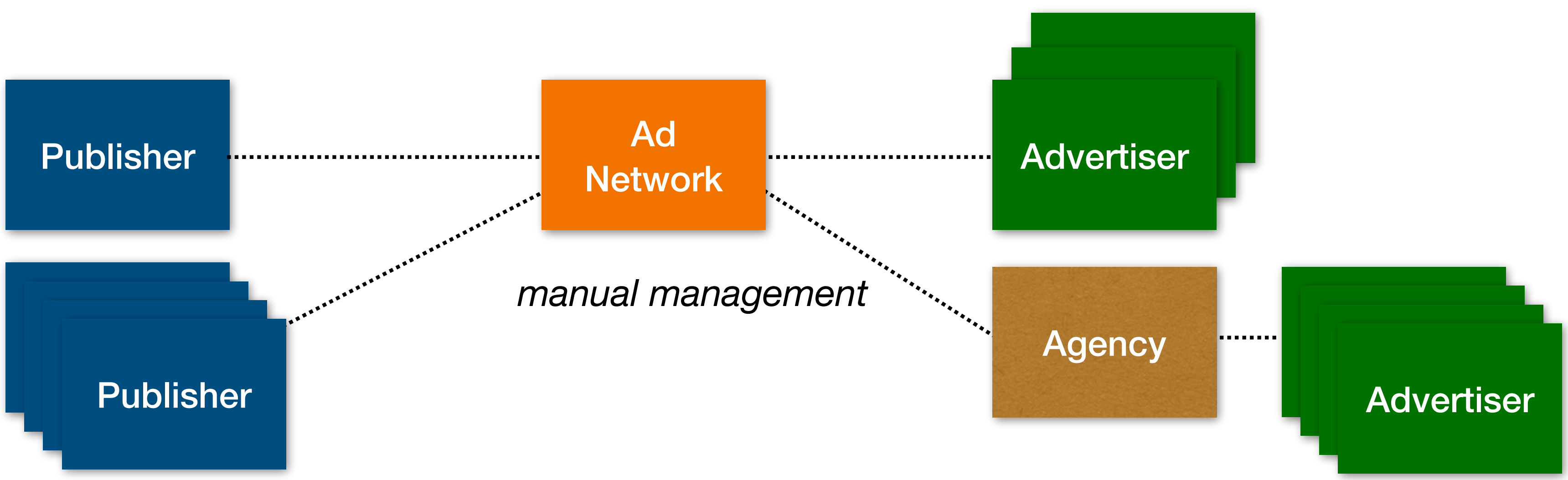
- In the early days, publishers (e.g., Yahoo, AOL) sold banner ads directly to advertisers or agencies.
  - Manual deal contracts
  - CPM (Cost Per Mille = 1K impressions)

- Characteristics
  - Limited targeting (contextual—based on the site's them)
  - Manual workflow: sales teams, insertion orders
  - High barriers for small advertisers (high upfront)



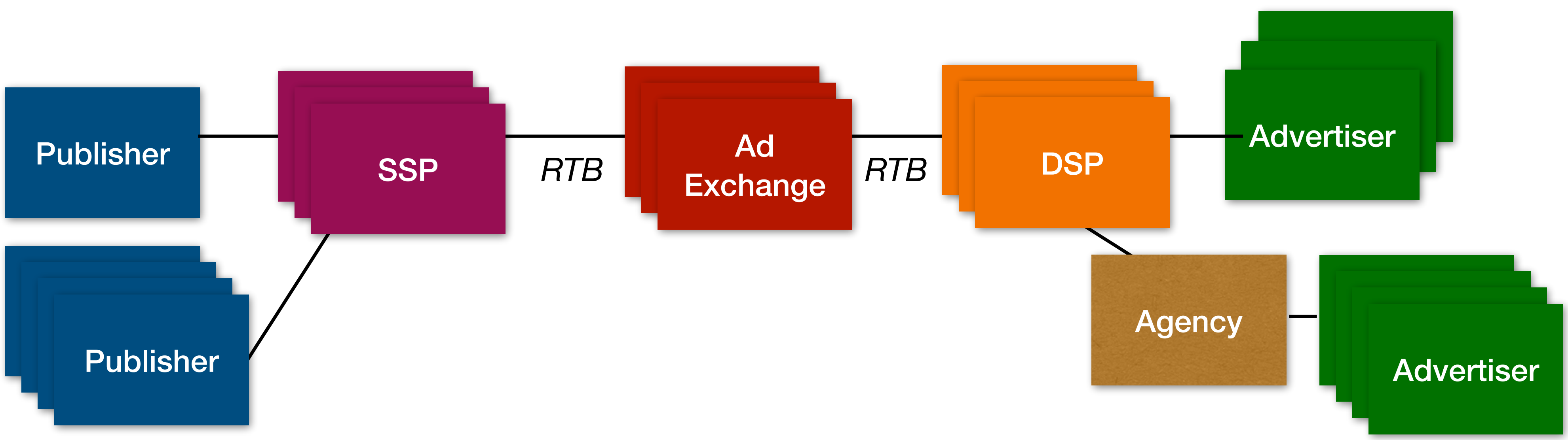
Ad Networks (2000s)

- Problems
  - Publishers had unsold “remnant” inventory, and advertisers wanted cheaper scale across many sites
- Solution
  - Ad networks aggregated inventory from multiple publishers and sold it in bulk
- Characteristics
  - More efficient than 1:1 direct deals
  - Some basic targeting (e.g., demographics, categories)
  - Black-box pricing and transparency issues



Ad Exchanges, DSPs, and SSPs (late 2000s-)

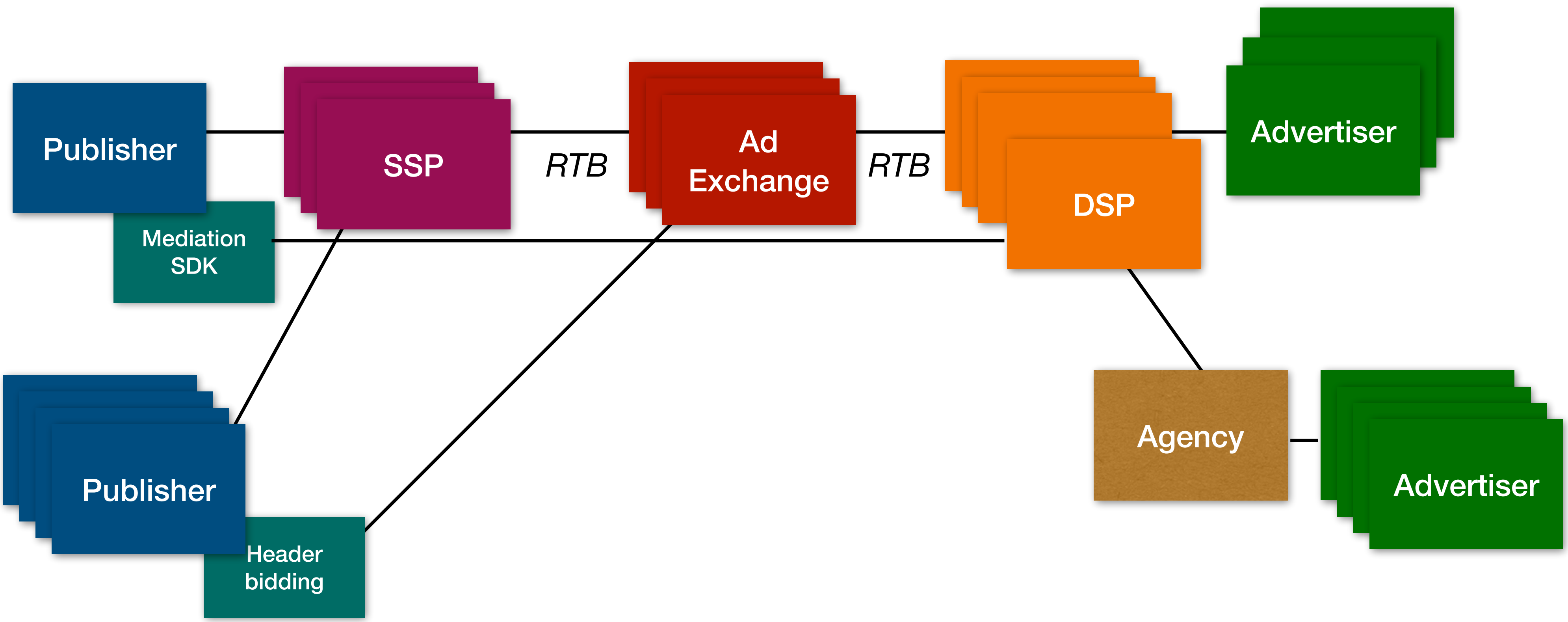
- **Ad Exchanges:** Marketplaces where ad impressions are auctioned in real time.
  - Enabled real-time bidding (RTB) at impression level instead of bulk manual buys
  - e.g., Google AdX, InMobi, Magnite, PubMatic, etc.
- **Supply-Side Platforms (SSPs):** Help publishers connect inventory to multiple exchanges and maximize yield
  - e.g., Google Ad Manager, (Google AdMob), Magnite, PubMatic, etc.
- **Demand-Side Platforms (DSPs):** Help advertisers and agencies buy inventory programmatically across exchanges.
  - Enable advanced targeting using data (cookies, ADIDs, behavioral signals)
  - e.g., The Trade Desk, Moloco, etc.



Mediation & Header Bidding (mid 2010s-)

- **Mediation Platforms:** Especially in mobile apps, SDK mediation allowed publishers to plug in multiple ad networks/SSPs so that each impression could be auctioned to the highest bidder.
  - e.g.,AppLovin MAX, ironSource, Google AdMob mediation, etc.

- **Header bidding:** Especially in web, publishers added code to their sites allowing multiple exchanges to bid simultaneously before the ad server decided the winner.
  - e.g., Prebid.js, Amazon TAM, Google Open Bidding, etc.



OpenRTB

- Now, Online Ads is **“Programmatic”**
  - OpenRTB is a standardized protocol for programmatic ad buying and selling
  - Created by IAB Tech Lab (Interactive Advertising Bureau)
  - To enable seamless communication between DSPs, SSPs, and Ad Exchanges for real-time ad auctions.
  - v 1.0: Feb, 2011
  - v 2.6: Sep, 2023

- Key Benefits
  - Standardization: consistent format across all partners
  - Transparency: common auction rules and bid attributes
  - Efficiency: reduces integration time
  - Extensibility: supports video, native, app, and CTV formats

Where ML/AI Plays?

- Basically, we will use ML/AI to match the "Best" user and campaign.
  - Supply Side (Publisher and SSP): to maximize yield
  - Exchange: for efficiency
  - Demand Side (DSP and Advertiser): to maximize ROI

Where ML/AI Plays?: Publisher

- Goals: Maximize yield, maintain content relevance, ensure brand safety
    - **Content classification (CV/NLP):** detect ad-appropriate or brand-safe content.
    - **Ad layout optimization:** predict best-performing placements per user/session.
    - **User engagement prediction:** forecast scroll depth, dwell time -> refine inventory value
    - **Fraud detection:** identify invalid traffic patterns using anomaly models

Where ML/AI Plays?: SSP

- Goals: Maximize publisher revenue while maintaining efficiency and transparency
  - **Bid landscape forecasting:** predict expected bid ranges to guide floor pricing
  - **Auction optimization:** use ML to choose between header bidding, direct deals, programmatic guaranteed, private auctions, or open auctions.
  - **Traffic quality scoring:** detect low-quality impressions or click fraud
  - **Latency prediction:** optimize routing between exchanges and demands to meet SLA.
  - **Ad quality filtering:** CV/NLP to block offensive creatives

Where ML/AI Plays?: Ad Exchange

- Goals: Fair, fast, and efficient auction across buyers and sellers
  - **Auction dynamics simulation:** predict clearing prices under different auction types
  - **Bid shading:** estimate second-price vs. first-price bid adjustment for DSPs.
  - **Fraud/anomaly detection:** detect unusual bidding patterns or replay attacks
  - **Latency optimization:** predict network and compute cost to meet <*100ms* SLA

Where ML/AI Plays?: DSP

- Goals: Maximize ROI for advertisers through efficient targeting, bidding, and creative
  - **Retrieval:** candidate ad selection using embedding similarity
  - **Ranking:** CTR/CVR/ROI prediction with DNN models
  - **Budget pacing:** predictive control to spend evenly or aggressively by performance
  - **Bid shading:** estimate second-price vs. first-price bid adjustment for DSPs.
  - **Creative optimization:** A/B and multi-armed bandit for headline/image variations
  - **User modeling:** sequence models (RNN/Transformers/GNN) for behavior and intent prediction

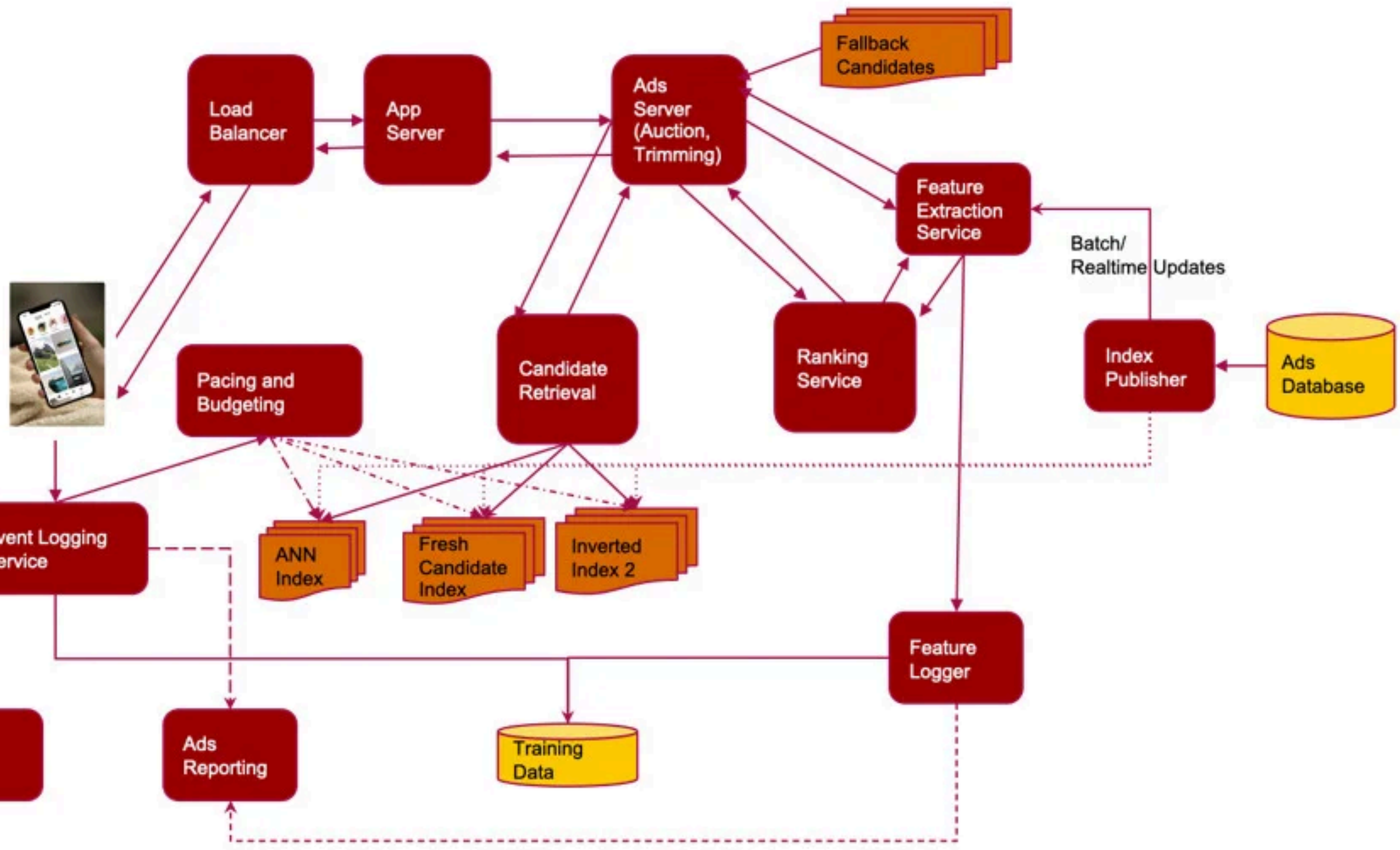
Where ML/AI Plays?: Advertiser

- Goals: Reach the right users with the right message at optimal cost
    - **Audience segmentation:** cluster users by intent, demographics, LTV
    - **Attribution modeling:** data-driven contribution analysis
    - **Media mix modeling:** regression/causal ML to allocate spend across channels
    - **Creative generation:** LLMs, diffusion models for personalized copy or images
    - **Performance prediction:** forecast campaign ROI and LTV
    - **Brand lift analysis:** statistical inference models on awareness & recall metrics

e.g., DSP—Retrieval

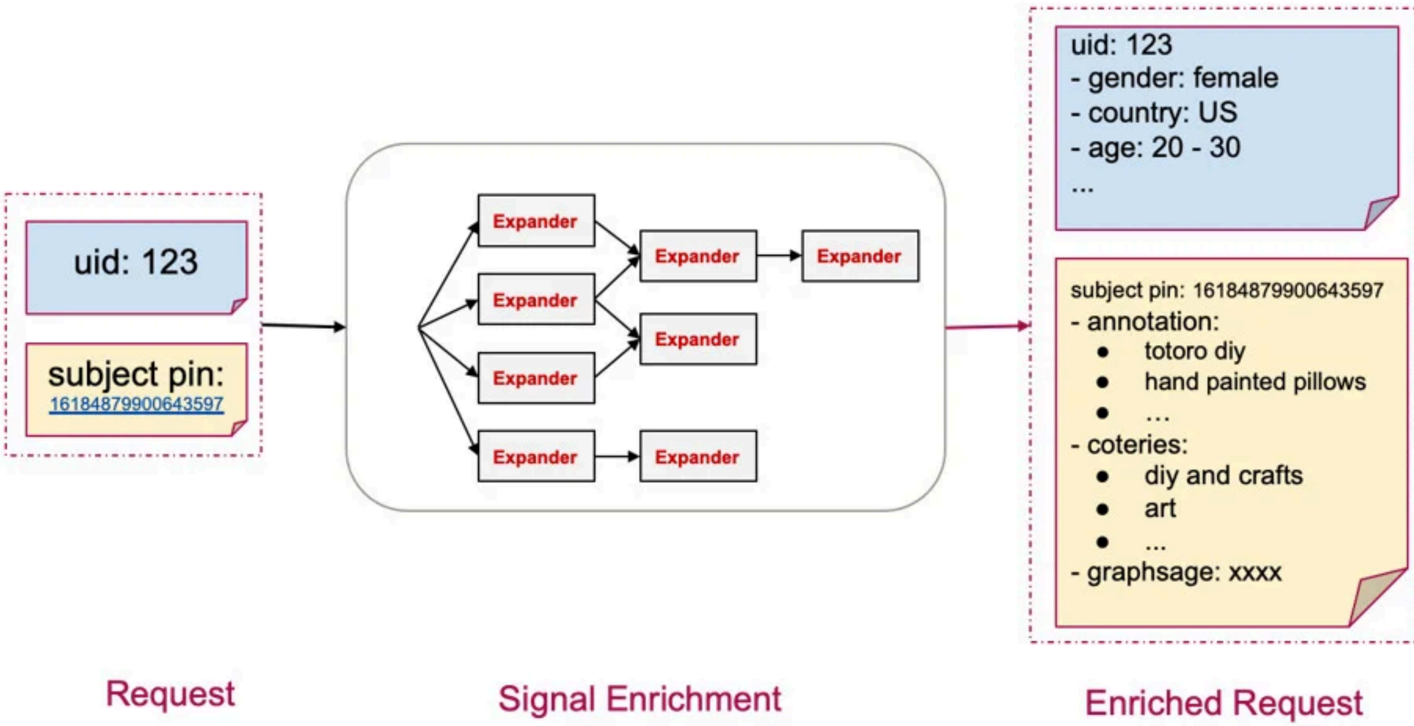
"Unpacking How Ad Ranking Works at Pinterest," Anthony Alford
https://www.infoq.com/articles/pinterest-ad-ranking-ai/

Pinterest Ads Serving Infrastructure High Level Overview



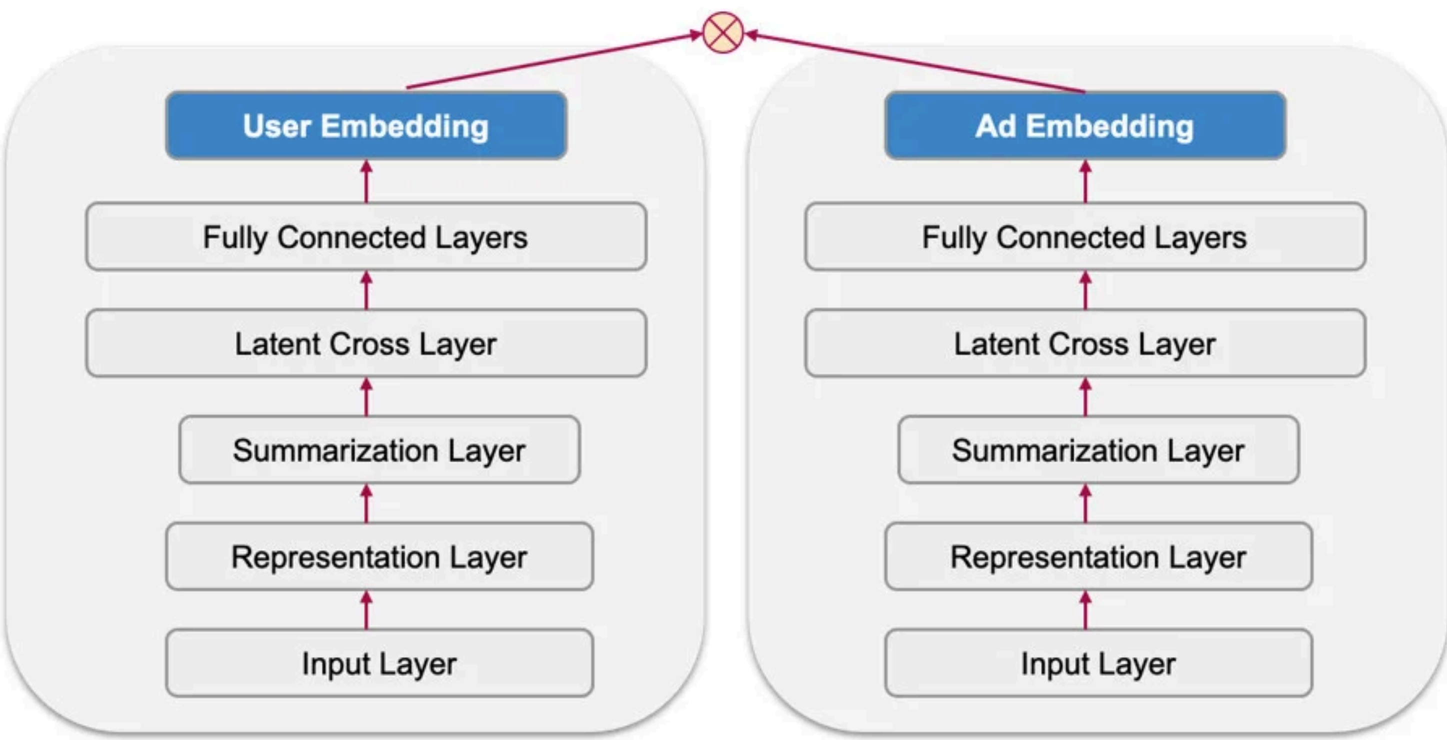
e.g., DSP—Retrieval (cont'd)

- Retrieval: to select the best ads candidate with the best efficiency—uses very lightweight ML model
- Signal Enrichment
  - uses several graph-based expanders to fetch extra features from key-value feature stores



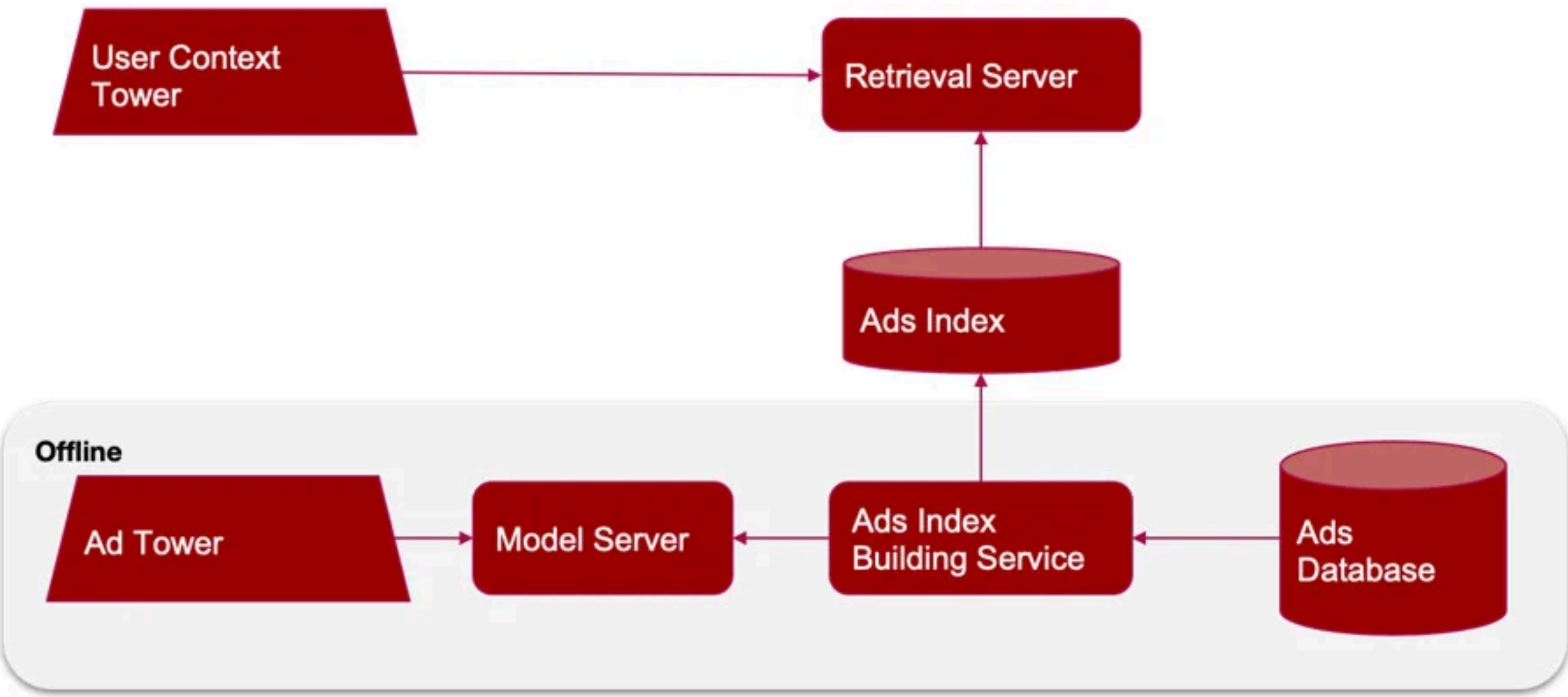
e.g., DSP—Retrieval (cont’d)

- Two-Tower Model
  - separate user and item information.



e.g., DSP—Retrieval (cont'd)

- Retrieval Serving Architecture



Conclusion—The Landscape Ahead

- From RTB to Everywhere
  - The RTB transformed online advertising into a fully programmatic ecosystem.
  - Every impression is a mini marketplace where algorithms decide what ad to show, how much to bid, and to whom.
- No One-Size-Fits-All Solution
  - There is no boilerplate template for how ML/AI should be applied in advertising.
  - Each participant (publisher, SSP, exchange, DSP, advertiser) faces different objectives, constraint, and data realities.
- The Future
  - Programmatic advertising is no longer about automation—it's about intelligent decision-making at scale (and realtime).
  - The next decade belongs to **adaptive**, **transparent**, and **human-aligned** AI in advertising