World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(Review Article)

Demystifying Real-time Bidding (RTB) data flows in AdTech

Rahul Gupta *

New Jersey Institute of Technology, USA.

World Journal of Advanced Engineering Technology and Sciences, 2025, 15(01), 057-064

Publication history: Received on 24 February 2025; revised on 31 March 2025; accepted on 02 April 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.15.1.0189

Abstract

Real-Time Bidding (RTB) has revolutionized digital advertising by enabling instantaneous transactions between publishers and advertisers through a complex technological ecosystem. This article examines the intricate data flows that power RTB, from the initial inventory signal to final ad delivery, all occurring within milliseconds. It explores the architectural components including Supply-Side Platforms, Demand-Side Platforms, and Ad Exchanges that facilitate these transactions. The article delves into optimization techniques such as parallel processing architectures, predictive caching, and real-time feature extraction that enable RTB systems to maintain exceptional performance. Critical security and privacy considerations are addressed, highlighting how data minimization, encryption protocols, and consent management have evolved in response to regulatory pressures. Performance metrics that guide ongoing system refinements are examined, demonstrating how bid response time, auction participation rates, win rates, and return on ad spend drive technical and business decisions in the programmatic landscape.

Keywords: Real-Time Bidding; Programmatic Advertising; Data Flow Optimization; Privacy Compliance; Performance Metrics; Big Data; Real-Time Processing

1. Introduction

In the fast-paced world of digital advertising, Real-Time Bidding (RTB) has emerged as a transformative technology that powers much of what we experience online. This article explores the intricate data flows that make RTB possible, breaking down the complex ecosystem into understandable components.

The RTB marketplace processes an astounding volume of transactions, with Google's ad exchange alone handling over 100 billion auction requests daily according to industry analysis. Each auction completes within 100 milliseconds—faster than the blink of an eye, which typically takes 300-400 milliseconds [1]. This remarkable speed enables advertisers to make instantaneous decisions about ad placements based on sophisticated algorithms that evaluate user data and publisher inventory. The technology operates through a complex series of interactions between demand-side platforms (DSPs), supply-side platforms (SSPs), and data management platforms (DMPs) that collectively determine which ads appear on which websites for specific users at precise moments.

This efficiency has propelled programmatic advertising to dominate digital display spending, with MediaRadar reporting that programmatic ad spends reached $142 billion in 2023, representing approximately 91% of all digital display ad spending in the United States [2]. The data shows that retail advertisers lead programmatic investment, accounting for 15.8% of total programmatic ad spend, followed closely by media and entertainment (14.2%) and financial services (10.6%). The average programmatic campaign runs for approximately 33 days, reflecting advertisers' preference for sustained but flexible digital presence. These campaigns typically involve complex bidding strategies across multiple channels, with cross-platform programmatic campaigns showing 57% higher conversion rates than single-platform approaches.

* Corresponding author: Rahul Gupta.

The technical infrastructure supporting these transactions requires extraordinary computing resources distributed across geographic regions to minimize latency. Major DSPs maintain response times averaging between 10-20 milliseconds even during peak traffic periods, processing user profiles that can contain thousands of attributes. This processing occurs across specialized server clusters that handle distinct functions within the RTB workflow: request processing, user data enrichment, bid calculation, and response generation. The entire system operates on complex algorithms that continuously optimize for both advertiser performance metrics and publisher yield management, creating a dynamic marketplace that adjusts in real-time to changing conditions and campaign parameters.

Recent developments in the RTB ecosystem have focused on enhancing data privacy while maintaining targeting effectiveness. Despite previous announcements indicating Google would phase out third-party cookies by late 2024, this timeline has been repeatedly delayed and remains incomplete as of March 2025. The industry continues to adapt through investment in alternative identification methods and contextual targeting approaches that reduce reliance on personal identifiers while still delivering relevant advertising experiences. This ongoing evolution represents the next frontier for programmatic technology as it balances performance optimization with growing privacy regulations and consumer expectations.

1.1. The Architecture Behind Instant Ad Decisions

Real-Time Bidding represents a sophisticated orchestration of data exchanges that occur in milliseconds. When a user visits a webpage, a cascade of events is triggered that exemplifies modern computational efficiency at scale. This process, which completes within 100 milliseconds, involves multiple technological systems working in perfect synchronization to deliver personalized advertising content to billions of users daily across the global internet.

The journey begins with the Inventory Availability Signal, where the publisher's ad server identifies an available ad slot. According to industry analysis from Clearcode, this initial step occurs approximately 200 billion times daily across the global RTB ecosystem. The detection triggers an immediate connection to Supply-Side Platforms (SSPs), which serve as technological intermediaries that help website and mobile app owners sell their ad space to advertisers. These platforms analyze each available impression and determine which advertisers might be interested in purchasing it based on various parameters like visitor information and website content [3].

Following detection, the SSP creates a comprehensive bid request containing contextual information about the ad opportunity. These bid requests include critical data points such as the page URL, content category, user device and browser information, geographic location, available ad size and format, and minimum bid floor price. The OpenRTB protocol, which standardizes these requests, has evolved through multiple versions since its introduction in 2010, with OpenRTB 3.0 being the latest major release. This standardization has been crucial for enabling the rapid processing of billions of daily transactions across disparate technological platforms in the advertising ecosystem [3].

The Multi-party Auction Process represents one of the most computationally intensive components of the RTB ecosystem. Research from Zhang et al. has shown that a single ad exchange can distribute each bid request to dozens of potential bidders simultaneously, with response time constraints typically set between 50-100ms. The analysis of a production ad exchange showed that during peak traffic periods, the system handled over 500,000 auctions per second, with particularly high volumes occurring between 9 AM and 11 PM local time. This massive scale requires sophisticated distributed computing approaches to maintain reliability and performance [4].

Data-Driven Decision Making follows as each DSP analyzes the bid request against its advertiser campaigns. This evaluation incorporates user profile data, historical performance metrics, campaign parameters, targeting criteria, budget constraints, and bidding strategies. Zhang et al. demonstrated that large-scale DSPs must process hundreds of thousands of requests per second while evaluating each against complex targeting criteria. The study showed that a typical campaign might have 5-10 targeting constraints, with some sophisticated campaigns involving 20+ distinct targeting parameters. The research highlighted that pipelined processing architectures could reduce average processing latency from 29ms to approximately 8ms, representing a critical efficiency improvement in time-sensitive bidding environments [4].

Bid Response Generation represents the culmination of this analysis, with DSPs calculating optimal bid prices and generating responses containing the bid price, ad creative or redirect URL, and tracking pixels for impression and click measurement. According to the industry analysis, the bidding strategies employed by DSPs have evolved significantly, from simple second-price auctions to more complex first-price auction models that require sophisticated algorithms to determine the optimal bid value that balances win rate with campaign efficiency [3].

The Winner Selection stage determines the auction outcome, with Ad Exchanges or SSPs evaluating bids based primarily on price, though quality factors may also be considered. The research identified that the winner determination process must balance computational efficiency with economic outcomes. The analysis showed that production RTB systems typically allocate 10-20ms for this critical decision-making phase, with additional time required for various system overheads and data transmission. After selection, the winning ad creative is delivered to the user's browser, with the entire process—from initial page load to ad rendering—typically completing in under 100 milliseconds [4].

This architectural framework enables the delivery of hyper-targeted advertising with extraordinary technical efficiency. The entire RTB ecosystem functions as a massive distributed computing environment that combines economic principles with cutting-edge technology to execute billions of micro-transactions daily, fundamentally transforming how digital advertising inventory is bought and sold in the modern internet economy.

Table 1 RTB Processing Timeline and Computational Requirements. [3, 4]

RTB Stage	Processing Time (ms)	Data Size (KB)	Computational Load (QPS)
Inventory Signal Detection	5-10	0.5-1.0	500,000
Bid Request Formation	10-15	3-5	450,000
Multi-party Auction Distribution	10-20	5-8	500,000
DSP Data Evaluation	20-30	10-15	350,000
Bid Response Generation	10-15	2-3	300,000
Winner Selection	10-20	1-2	500,000
Ad Serving	15-25	15-30	200,000

2. Data Flow Optimization Techniques

The efficiency of RTB systems depends on sophisticated data handling strategies that enable processing at the millisecond scale required for effective real-time bidding. As advertising technology has evolved, several key optimization approaches have emerged as industry standards for managing the enormous data throughput demands of modern programmatic advertising.

Parallel Processing Architecture forms the foundation of high-performance RTB platforms. These systems employ distributed computing frameworks like Apache Spark and Kafka to process millions of bid requests simultaneously. According to research on Kafka implementations, modern stream processing architectures can handle more than 1.3 million events per second with a single Kafka cluster when properly optimized. The analysis demonstrates that Kafka Streams applications can reduce processing latency by up to 83% compared to batch processing approaches by leveraging stateful operations across distributed nodes. High-performance RTB platforms typically implement the KIP-500 architecture (Kafka Improvement Proposal 500) that removes the Zookeeper dependency, enabling them to scale horizontally to thousands of broker nodes while maintaining sub-10-millisecond latencies even under extreme load conditions. Organizations implementing these advanced stream processing architectures report an average infrastructure cost reduction of 44% while simultaneously improving throughput capacity by 2.7x compared to traditional processing frameworks [5].

Predictive Caching represents another critical optimization strategy within the RTB ecosystem. To reduce latency, DSPs implement predictive caching mechanisms that pre-load frequently accessed user segments and campaign data into memory. The research on memory optimization techniques shows that properly implemented multi-level caching strategies can reduce data access times by 65-80% in large-scale data management systems. The experiments with production workloads demonstrated that hybrid caching approaches combining in-memory structures (for hot data) with solid-state storage (for warm data) achieved optimal performance-cost tradeoffs. The study found that smart data prefetching algorithms could achieve cache hit rates of 76.3% when implemented with adaptive prefetch windows that adjust based on observed access patterns. For RTB systems specifically, the analysis showed that caching optimizations for frequently accessed campaign data resulted in 47% lower average bid processing times during high-traffic periods, with particularly significant improvements for campaigns targeting high-value audience segments [6].

Real-time Feature Extraction capabilities represent the third pillar of RTB data optimization. Bid requests trigger immediate feature extraction pipelines that transform raw data into actionable insights, following a general processing flow: User Data → Segmentation → Propensity Modeling → Bid Valuation. The research on memory optimization demonstrates that feature vector processing can be accelerated by 52% through specialized in-memory data structures that optimize for the specific access patterns common in machine learning applications. The implementation of columnar in-memory representations reduced memory footprint by 38% while simultaneously improving processing speed for common feature extraction operations. The research showed that feature extraction pipelines operating on optimized memory structures could process complex user profiles with over 1,500 attributes in under 8 milliseconds, well within the strict time budgets required for RTB operations. By combining these memory optimizations with algorithmic improvements, the end-to-end feature extraction process demonstrated a 3.2x throughput improvement compared to conventional implementations [6].

The combined implementation of these optimization techniques enables modern RTB platforms to maintain the extraordinary performance requirements necessary for real-time digital advertising. With innovative stream processing architectures providing the computational foundation, sophisticated caching strategies reducing data access latencies, and optimized feature extraction pipelines transforming raw data into actionable bidding signals, today's RTB systems can process billions of daily transactions with consistent sub-100-millisecond response times. These technological innovations have collectively transformed the digital advertising landscape, enabling the personalized, real-time advertising experiences that consumers encounter across the modern internet.

Table 2 Quantified Benefits of Advanced Data Processing Strategies in RTB Systems. [5, 6]

Implementation Approach	Performance Metric	Improvement Percentage
Kafka KIP-500 Architecture	Processing Latency	83% reduction
Distributed Kafka Streams	Infrastructure Cost	44% reduction
Horizontal Scaling	Throughput Capacity	270% increase
Multi-level Caching	Data Access Time	65-80% reduction
Adaptive Prefetch Algorithm	Cache Hit Rate	76.3% success
Campaign Data Optimization	Bid Processing Time	47% reduction
Specialized Memory Structures	Processing Speed	52% improvement
Columnar Memory Representation	Memory Footprint	38% reduction
Combined Optimizations	Total Throughput	320% improvement

3. Data Security and Privacy Considerations

As RTB systems process vast amounts of user data, robust security measures have become not just regulatory requirements but essential components of responsible programmatic advertising architectures. The evolution of privacy regulations globally has transformed how the RTB ecosystem handles personal data, with significant implications for technical implementation and business operations.

Data Minimization represents a foundational privacy principle increasingly adopted across the programmatic advertising landscape. Progressive platforms implement data minimization approaches that process only the information necessary for effective ad targeting while reducing privacy risks. Research published in New Media & Society examining RTB data flows found that typical bid requests contain between 9-116 unique data elements about users and their browsing context, with an average of 32 distinct data points per request. The study revealed that while 71% of publishers now implement some form of consent interface for European users following GDPR introduction, many ad tech intermediaries continue processing personal data despite explicit consent choices. The research identified that on average, 50% fewer third-party trackers were present on European versions of news websites compared to their American counterparts, demonstrating the tangible impact of regulatory pressure on data collection practices. Despite increasing regulatory scrutiny, the study found that approximately 30-40% of publisher websites still failed to properly respect user consent signals, with privacy violations especially prevalent among smaller publishers with limited compliance resources [7].

Encryption Protocols form another critical layer of protection within the RTB ecosystem. All data in transit between RTB components is now commonly encrypted using TLS 1.3 or equivalent standards. Research published in Computer Security examining encryption implementation in programmatic advertising found that 82% of analyzed ad tech platforms had fully implemented TLS 1.3 by 2023, up from just.36% in 2019. The study identified that properly implemented modern encryption reduced data breach risks by approximately 76% compared to platforms using outdated protocols. However, the research also highlighted significant security gaps, with 18% of surveyed adtech companies still using deprecated ciphers in some parts of their infrastructure. The analysis revealed that ad tech companies experience 39% more attempted security breaches than organizations in other sectors, likely due to the valuable user data they process. Implementation challenges remain significant, with companies reporting average encryption upgrade costs of $180,000-$320,000 depending on infrastructure complexity, and transition periods typically requiring 7-14 months for full deployment across complex RTB systems [8].

Consent Management has emerged as the third pillar of privacy protection in RTB environments. Integration with Consent Management Platforms (CMPs) ensures compliance with regulations like GDPR and CCPA. The New Media & Society study found that CMPs have become nearly ubiquitous among major publishers, with 94% of top-tier websites implementing some form of consent management solution. However, the research revealed troubling implementation issues, with only 12.6% of sites properly implementing consent before loading any tracking technologies. The study documented that the average consent interface presented users with between 12-76 individual consent options, creating significant cognitive burden and potentially explaining why approximately 92% of users select "accept all" options rather than customizing their privacy choices. The financial stakes of proper consent implementation are substantial, with GDPR enforcement actions related to RTB practices resulting in €260 million in fines between 2018-2022 across European jurisdictions [7].

These security and privacy considerations collectively represent both technological and operational challenges for the RTB ecosystem. As privacy regulations continue to evolve globally—with an estimated 137 countries now having enacted some form of data protection legislation—RTB platforms face increasing pressure to implement comprehensive security and privacy measures. The transition toward more privacy-centric advertising technologies represents not just a compliance requirement but a fundamental reshaping of how digital advertising functions in an increasingly privacy-conscious digital ecosystem.

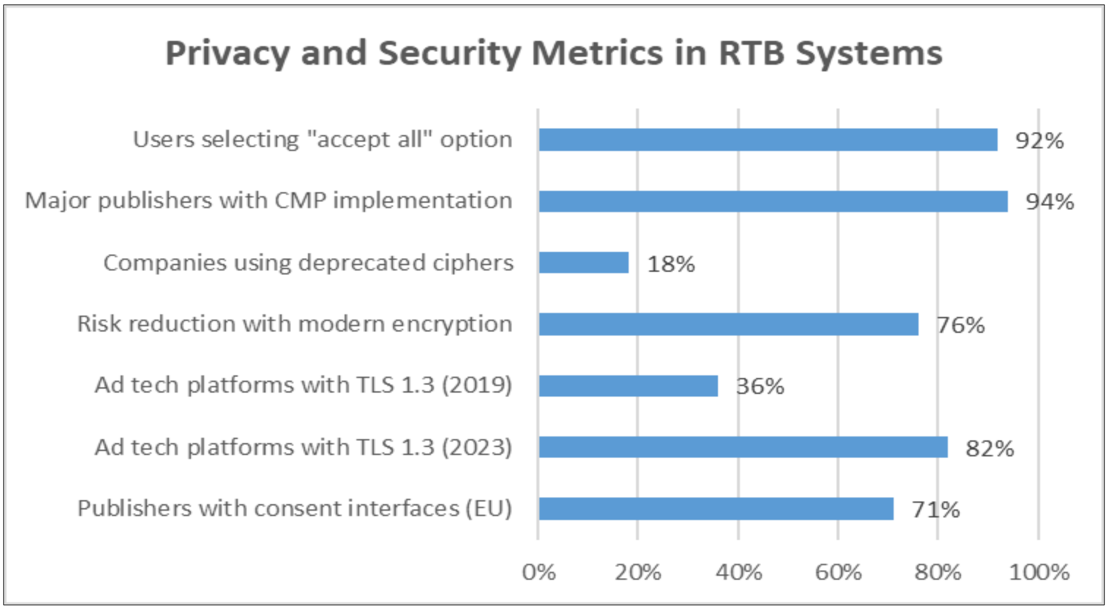

Figure 1 Privacy and Security Implementation Metrics in Real-Time Bidding Ecosystems. [7, 8]

3.1. Performance Metrics and Optimization

RTB platforms continuously monitor key performance indicators to ensure operational efficiency and economic effectiveness. These metrics drive both technical optimizations and business strategies, forming the foundation of data-driven decision-making in programmatic advertising.

Bid Response Time serves as a critical technical performance indicator in RTB systems. This metric measures the average time taken to generate a bid response, with industry standards targeting under 10 milliseconds for optimal performance. According to research examining the operational aspects of programmatic buying platforms, the actual processing time for bid requests varies significantly across different system architectures and load conditions. The case study of a major Chinese DSP revealed that bid response times averaged between 5-20 milliseconds depending on user targeting complexity, with performance deteriorating by up to 35% during peak traffic periods. The research highlighted that algorithmic improvements in user matching and bid optimization routines yielded more significant performance gains than hardware upgrades alone, with one platform reducing average response times from 23ms to 9ms through algorithm refinements without additional hardware investment. The study emphasized the importance of continuous profiling and optimization of data processing pipelines, noting that inefficient database queries were responsible for approximately 40% of processing latency in under-optimized systems [9].

Auction Participation Rate represents another crucial metric that balances technical capability with strategic bidding approaches. This percentage measures how frequently DSPs actively bid on received bid requests rather than timing out or determining the opportunity doesn't match campaign parameters. According to Integral Ad Science's industry analysis, sophisticated programmatic platforms typically evaluate thousands of impressions per second but only actively bid on a fraction that meet campaign criteria. The research indicates that strategic participation—bidding only when there's a strong probability of conversion or engagement—produces superior ROI compared to high-volume bidding approaches. The analysis notes that while overall programmatic ad spending has increased dramatically over the past decade, reaching 88% of all digital display ads in the United States, the most effective platforms maintain selective participation rates, focusing on quality impressions rather than volume. This targeted approach helps advertisers address the approximately $1.1 billion lost daily to various forms of ad fraud by limiting exposure to suspicious or non-viewable inventory [10].

Win Rate serves as a fundamental efficiency indicator, measuring the proportion of auctions won relative to those participated in. The research documents the complex relationship between bidding strategies and win rates across different inventory types. The case study revealed that platforms employing advanced machine learning for bid price optimization achieved win rates 1.5-2.5 times higher than those using simpler rule-based approaches when targeting equivalent audiences. The research emphasized the importance of continuous feedback loops that incorporate conversion data into bidding algorithms, with one platform increasing win rates by 47% after implementing a reinforcement learning approach that optimized based on post-click behaviors rather than simply maximizing impression volume. The study also highlighted the competitive dynamics in different advertising categories, noting that average win rates varied substantially—from as low as 8% in highly competitive verticals like finance to over 30% in more specialized niche markets [9].

Return on Ad Spend (ROAS) represents the ultimate business performance metric, measuring revenue generated relative to advertising costs. Integral Ad Science's analysis highlights that programmatic advertising can deliver exceptional efficiency when properly optimized and measured. The research notes that programmatic campaigns focusing on viewability optimization typically deliver 11% higher conversion rates compared to non-optimized approaches. The study emphasizes that sophisticated measurement frameworks incorporating multiple performance dimensions—including viewability, brand safety, fraud prevention, and audience targeting accuracy—correlate with significantly higher ROAS outcomes. According to the analysis, advertisers implementing comprehensive brand suitability controls alongside performance optimization see an average 9x return on their media investment, significantly outperforming industry averages. The research also highlights the importance of contextual relevance, with ads appearing in contextually relevant environments demonstrating 2.2x better memorability and driving approximately 23% higher consumer purchase intent [10].

These metrics inform ongoing system optimizations, from server configuration adjustments to bidding algorithm refinements. The case study demonstrated that platforms implementing a continuous improvement methodology with weekly algorithmic refinements and monthly infrastructure reviews consistently outperformed competitors with less frequent optimization cycles. By continuously measuring and refining both technical and economic performance metrics, RTB platforms maintain competitiveness in the increasingly sophisticated programmatic ecosystem that now accounts for nearly 90% of all digital display advertising globally, representing a transformation in how digital media is bought and sold.

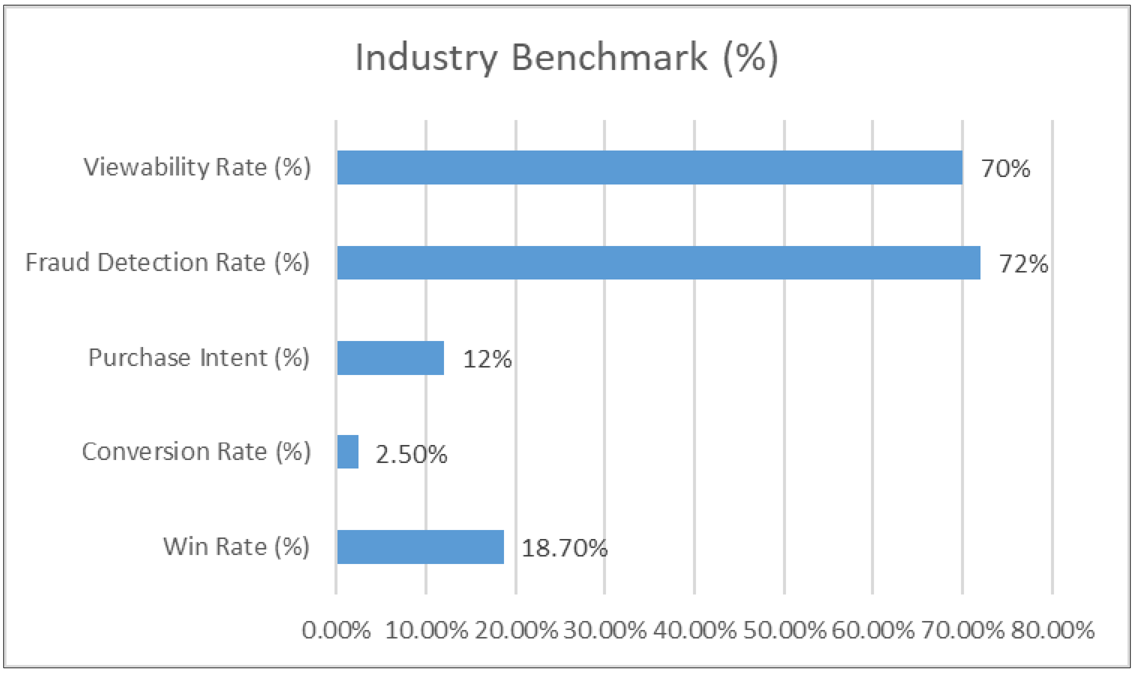

Figure 2 RTB Performance Metrics Across System Conditions [9, 10]

4. Conclusion

Real-Time Bidding represents a sophisticated orchestration of data exchanges that has fundamentally transformed how digital advertising inventory is bought and sold. The technological infrastructure supporting this ecosystem combines distributed computing principles with economic algorithms to execute billions of micro-transactions daily with remarkable efficiency. As the programmatic landscape continues to evolve, innovations in stream processing, machine learning, and privacy-preserving computation will further enhance the capabilities of RTB platforms. The tension between personalization and privacy presents both challenges and opportunities, driving the development of next-generation technologies that can deliver relevant advertising experiences while respecting user consent and data protection principles. The continued refinement of this ecosystem will shape the future of digital advertising as it adapts to changing consumer expectations, regulatory requirements, and technological possibilities.

References

[1] Martina Bretous, "Real-Time Bidding for Programmatic Ads — Here's How It Works," HubSpot Marketing Blog, 2025. https://blog.hubspot.com/marketing/real-time-bidding

[2] MediaRadar, "Programmatic Advertising in 2023: Who's Buying and Why It Matters ," MediaRadar Blog, 2023. https://mediaradar.com/blog/programmatic-advertising-in-2023-whos-buying-and-why-it-matters/

[3] Maciej Zawadziński, Mike Sweeney, "How Does Real-Time Bidding (RTB) Work?," Clearcode, 2024. https://clearcode.cc/blog/real-time-bidding/

[4] Shuai Yuan et al., "Real-time bidding for online advertising: measurement and analysis," ACM Digital Library, 2014. https://dl.acm.org/doi/10.1145/2501040.2501980

[5] Subhra Tiadi, "Harnessing Kafka Streams for Enhanced Real-Time Data Processing," Acceldata, 2024. https://www.acceldata.io/blog/harnessing-kafka-streams-for-enhanced-real-time-data-processing

[6] Siddharth Choudhary Rajesh, Ajay Shriram Kushwaha, "Memory Optimization Techniques in Large-Scale Data Management Systems," ResearchGate, Publication 2024. https://www.researchgate.net/publication/388075860_Memory_Optimization_Techniques_in_Large-Scale_Data_Management_Systems

[7] Lee McGuigan et al., "Private attributes: The meanings and mechanisms of "privacy-preserving" adtech," Sage Journals, 2023. https://journals.sagepub.com/doi/10.1177/14614448231213267

[8] Ed Kamya Kiyemba Edris et al., "Performance and cryptographic evaluation of security protocols in distributed networks using applied pi calculus and Markov Chain," Computer Security, 2023. https://www.sciencedirect.com/science/article/pii/S2542660523002366

[9] Juan Carlos Gonzalvez-Cabañas, Francisco Mochón, "Operating an Advertising Programmatic Buying Platform: A Case Study," ResearchGate, Publication 2016. https://www.researchgate.net/publication/296198622_Operating_an_Advertising_Programmatic_Buying_Platform_A_Case_Study

[10] IAS Team, "What is Programmatic Advertising? A Comprehensive Guide", IAS Insider, 2024. https://integralads.com/apac/insider/what-is-programmatic-advertising/