https://doi.org/10.54318/eip.2025.da.387

THE ROLE OF NON-RELATIONAL DATABASES IN THE OPTIMIZATION OF MARKET MATCHING

Danilo Đurđić[1], e-mail: danilo.djurdjic@ekof.bg.ac.rs
Aleksandra Zečević[1], e-mail: aleksandra.zecevic@ekof.bg.ac.rs

Abstract: *The paper examines the role of NoSQL databases in optimizing market matching processes, especially with regard to supply and demand coordination. Traditional relational databases are reliable with structured data, but they are not much use in handling large volumes and diverse types of data on today's dynamic markets. Scalable and flexible NoSQL databases allow Netflix to efficiently manage and analyze big volumes of unstructured and semi-structured data. NoSQL databases apply advanced algorithms and real-time data analytics, thus allowing for personalized recommendations that improve customer satisfaction and enhance user retention. The findings of the study point to the conclusion that the usage of NoSQL solutions has become imperative for companies intending to maintain their competitiveness by optimal data management and market matching.*

Key words: NoSQL databases, market matching, network effects.

ULOGA NERELACIONIH BAZA PODATAKA U OPTIMIZACIJI TRŽIŠNOG UPARIVANJA

Apstrakt: *Ovaj rad istražuje ulogu NoSQL baza podataka u optimizaciji procesa tržišnog uparivanja, posebno u kontekstu usklađivanja ponude i tražnje. Tradicionalne relacione baze podataka, iako pouzdane za strukturisane podatke, nisu pogodne za rad sa velikim količinama i raznovrsnim tipovima podataka koji se generišu na današnjim dinamičnim tržištima. NoSQL baze podataka, sa svojom skalabilnošću i fleksibilnošću, omogućavaju kompanijama poput Netflix-a efikasno upravljanje i analizu ogromnih količina nestrukturiranih i polustrukturiranih podataka. Primenom naprednih algoritama i analitike u realnom vremenu, NoSQL baze olakšavaju kreiranje personalizovanih preporuka, čime se poboljšava zadovoljstvo i zadržavanje korisnika. Studija zaključuje da je usvajanje NoSQL rešenja ključno za kompanije koje teže konkurentnosti kroz optimizaciju upravljanja podacima i procesa tržišnog uparivanja.*

Ključne reči: NoSQL baze podataka, tržišno uparivanje, mrežni efekti.

JEL klasifikacija: C63, D47, L86, O33.

[1] Univerzitet u Beogradu, Ekonomski fakultet, Kamenička 6, 11000 Beograd, Srbija

1. Introduction

In the modern digital era, effective data management and use have become crucial to achieving and maintaining a competitive advantage in the business environment. Numerous internal and external factors determine company's success in the market[2], including how companies manage their data. Traditional relational databases, although clearly structured and reliable, are often not an efficient solution for huge volumes and different types of data generated on today's dynamic markets. This is where NoSQL databases come in, offering scalable, flexible and high-performance solutions for managing large datasets and driving business intelligence.

NoSQL databases are a collective term for various database technologies, developed to overcome the shortcomings of relational databases. They are widely accepted in various industries. These systems are particularly valued for their ability to process unstructured and semi-structured data, supporting horizontal scaling and easier data analysis. These features make them ideal for applications that require real-time analytics and a personalized user experience. As stated in the paper[3], reducing uncertainty regarding the common value component in-creases efficiency, which aligns with the goal of NoSQL databases to provide businesses with timely, accurate insights for more effective decision-making.

Effective market matching, which involves the process of coordinating supply and demand in the most efficient way, depends on data processing speed and flexibility. In this context, NoSQL databases play a key role in enabling busi-nesses to process vast volumes of different types of data, from user preferences and behavior to market trends and product inventories. This ability to process data in real time gives businesses the agility and performance needed to quickly adapt to market changes, thereby optimizing their operations and improving customer satisfaction.

This paper investigates the role of NoSQL databases in optimizing the process of matching supply and demand and improving company performance. The case study of Netflix, which has successfully implemented NoSQL solutions, shows how these databases bring superior market matching ca-pabilities. The paper will also discuss the challenges and limitations associated with the adoption of NoSQL databases and point to future trends and reco-mmendations for busi-nesses considering this approach.

As the digital economy continues to evolve, the importance of effective data management systems cannot be overstated. NoSQL databases represent a

[2] Čivić et al. (2022), p. 49
[3] Trifunović (2023), p. 10

significant step forward in this area as they provide the tools companies need to thrive in an increasingly data-driven world. The research objective is to highlight the transformative potential of NoSQL databases in the context of market matching.

2. Theoretical basis of non-relational databases

In the dynamic data management environment, non-relational databases have emerged as a disruptive innovation, designed to overcome the limitations of traditional relational databases. As businesses increasingly strive to optimize market matching, knowing the theoretical bases of non-relational databases becomes crucial. This section explores the basic concepts of non-relational databases, market matching mechanisms, and how non-relational databases help companies efficiently manage large and diverse datasets. This encourages a personalized user experience, which is crucial for success on modern digital platforms.

2.1. Overview of NoSQL databases

Relational database management systems (RDBMS) have emerged as the dominant technology for storing structured data. As previous studies point out [4], a large number of web and business applications apply this principle. However, the emergence of the Internet of Things and the accelerated process of glo-balization in the past few decades have increased the need for unstructured data from various sources such as modern communication devices, sensors, digital media and IoT devices. Many companies, which base their operations on the analysis of customer preferences, store and process such data in order to imple-ment customized sales systems based on customer preferences and thus achieve competitive advantage. It is in such environments, characterized by huge volumes of dynamic data, that RDBMS systems have shown their shortcomings. As[5] states, relational database systems require a well-defined structure based on mathe-matical principles. Edgar Codd[6] states that maximum utility of relational data-bases is possible using the principles of relational algebra. However, this strictly defined, rigid database structure proved to be unsuitable for managing unstructured data, which is especially pronounced in modern applications that require greater flexibility and the ability to process different types of data. This problem has become apparent with the development of IoT technologies, where large volumes of dynamic, unstructured data is generated, requiring a more suitable solution.

[4] Strauch et al. (2011), p. 6
[5] Tiwara (2011), p. 5
[6] Codd (1979), p. 397

In order to overcome this problem, a system of non-relational databases, or NoSQL, was developed. As Thakur and Gupta[7] point out, the term NoSQL refers to database systems that do not use SQL to query their data. Although non-relational databases support SQL queries, non-relational databases queries are significantly different from RDBMS. The term NoSQL was first used by Carlo Strozzi in 1998 to distinguish non-relational databases from traditional data storage systems[8]. Originally, the term NoSQL was used to emphasize that non-relational databases do not rely on SQL items. The term NoSQL was originally intended to show a complete distinction from relational databases. However, as the purpose of non-relational databases is broader than that, the term "not only SQL" is used today.

2.2. Comparison of relational and non-relational databases

The development of non-relational databases can be attributed in part to the shortcomings of traditional RDBMS. The main problem that relational databases face when processing large data volumes (big data) is the speed of query execution, as well as memory space[9]. Song et al.[10] list the main problems of relational databases:

- Relational databases do not show a high degree of scalability. They focus on vertical scalability, because horizontal scalability brings problems with data synchronization and merging of different tables.
- Tabular data display often does not allow all types of data to be saved. The reason for this is that encapsulation of unstructured data inside a table is not possible.
- SQL is very complex and unintuitive with unstructured data.

It can be concluded that relational databases are not adapted to work with unstructured data. Therefore, databases that require a more flexible data organization structure and are more efficient in processing large volumes of data are necessary, and NoSQL databases satisfy these criteria[11]. These factors find support in the following:

- Highly scalable architecture allows for faster and more efficient data processing.
- Some of the most famous NoSQL databases are in-memory databases, which means data storage takes place in the computer's main

[7] Thakur and Gupta (2021), p. 119
[8] http://www.strozzi.it/cgi-bin/CSA/tw7/I/en_US/nosql/Home%20Page
[9] Song et al. (2014), p. 1
[10] Jatana et al. (2021), p. 3
[11] Kabakus and Kara (2017), p. 520

memory, such as RAM. This enables faster data reading and writing, thus helping applications cope with large data volumes[12].

- NoSQL databases do not require a strict data schema, which is particularly suitable for working with unstructured data.
- Low latency and high performance for working with large datasets.
- More efficient data sharing, which can multiply the efficiency of data processing [13].

However, relational databases are by no means a matter of the past. NoSQL can be applied in specific areas and can be considered an excellent complement to relational databases in practice. The main disadvantages of non-relational databases are[14]:

- Immaturity of non-relational databases implies insufficient standardization. This lack of standardization can be seen as a multi-layered pro-blem because it affects different aspects, such as data structuring rules, spe-cific languages for databases, as well as performance optimization rules. This can lead to major problems on the user side.
- The lack of a single query language prevents uniformity in working with different NoSQL databases, forcing developers to master different approaches and syntaxes depending on the system used. This can make it difficult to move between different databases and increases the complexity of learning and implementation.
- Insufficient standardization of database structuring rules. Different NoSQL databases may apply different approaches to data organization, which leads to incompatibilities between systems and makes it difficult to integrate them.
- Heavy maintenance. Due to different approaches used by different NoSQL databases, maintaining these systems can be extremely complex. The lack of standardization means that each database can have its own specific optimization and management requirements, which increases maintenance costs and requires more technical knowledge.

Considering all the above, it is clear that the advantages and disadvantages of relational and NoSQL databases are key to understanding their application in different contexts. Taking into account the fact that relational and non-relational databases have their specific advantages and disadvantages, which ultimately determine the area of their application, it can be concluded that the choice between these databases will depend entirely on user needs and the specific requirements of their application.

[12] Hemmatpour et al. (2020), p. 5
[13] Choi et al. (2014), p. 356
[14] Nayak et al. (2013), p. 19

2.3. Types of non-relational databases

The previous chapter established one of the key advantages of non-relational databases – solutions that exceed the capabilities of relational databases. Since they target different spheres of data management (i.e. have diverse purposes), four main types of NoSQL databases have been developed in modern business[15]:

- Wide column store or column families. This type of database is signi-ficantly different from traditional relational databases that require strict data organization schemas. With the column store database, there is no clearly defined scheme; each record (row in the table) can contain a specific number of columns, which can be completely different in structure and number of columns from other rows in the table. Similar to relational databases, column family databases function as two-dimensional arrays, with each record having multiple key/value pairs defined. When it comes to data storage, Wide column databases store all values within one column, which makes it possible to efficiently record large volumes of data without affecting other columns. Another important feature is simple scaling, which enables very efficient and fast work with big data.
- Document Store. These are document-oriented databases and are used to store and process data that is saved in the form of a document. There are several types of Document Store databases. What they have in common is that they handle data stored in XML, YAML and JSON[16], and CSV formats. Each document is defined by a unique key, which helps in identification.
- Key Value / Tuple Store. These databases enable data storage without the need to define a precise structure. Data is structured in such a way that there is a key, which is most often defined as a String value[17] and added real data of different types: string, number, bit, array, and objects. This type of database is widespread and is the basis for storing and processing data in an increasing number of applications. Also, key value databases are becoming more and more popular as integrated systems within complex applications that require intensive data processing, as well as in the areas of machine learning, and within larger systems that require support for more advanced data models[18].
- Graph Databases. Graph databases do not require a precise schema. They store their data in the form of graphs. This includes hubs, edges and bound properties. Hubs are basic elements within graphs that can relate to real-world objects such as organizations, individuals, and events. Edges are connections between hubs, i.e. they create a network that connects hubs to each other.

---

[15] Meier and Kaufmann (2019), p. 202
[16] Maté et al. (2012), p. 2210
[17] Meier and Kaufmann (2019), p. 203
[18] Idreos and Callaghan (2020), p. 2667

Finally, Bound Properties represent added properties to hubs or edges that can determine the strength or type of connection between them. Researching and establishing connections between the objects of this type of database makes it possible to gain significant knowledge about patterns of behavior or the features of various events.

Standard classification of NoSQL databases refers to document, key-value, column-family, and graph databases[19]. In addition, the mentioned categories are further segmented by subtypes and specialized databases. For example, the document database category has variations such as multimodel databases, which support multiple data models at the same time, and object-oriented data-bases, which are designed for object-oriented programming. The key-value databases have been expanded to include network and cloud databases, which are optimized for distributed computing environments, and XML databases, which are adapted to XML data structures. Similarly, graph includes broad column/family databases, highlighting their ability to process large volumes of data across many columns and graph databases, which are exceptional at mana-ging relationships between data points. The incorporation of specialized data-bases, such as multidimensional and multivalued, highlights the diversity within the NoSQL ecosystem, illustrating its ability to address a wide range of data management requirements and application scenarios. This detailed breakdown provides a nuanced overview of how NoSQL databases can be used to meet specific organizational needs, particularly in optimizing market matching and improving company competitiveness. Therefore, the next section will focus on the practical application of the databases mentioned above, with special reference to their role in market matching optimization.

3. The role of non-relational databases in market matching processes

The ability to effectively apply market matching and to maintain a company's competitiveness increasingly depends on sophisticated data management solutions. Market matching is a very complex process that requires a careful definition of the strategy by means of which companies match their offer of products or services with customer demand. This requires them to understand customer needs, preferences and behaviors, as well as broader market dynamics. The data volume to be processed in these circumstances is enormous. For this reason, NoSQL databases can play a transformative role in this area. Through efficient processing and analysis of huge and diverse datasets, NoSQL databases allow businesses to gain deeper insight into market trends and customer behavior. This, in turn, supports more accurate and dynamic market matching, ensuring that product and service offerings are in perfect harmony

[19] Acharya et al. (2019), p. 53

with market demand. Effective market matching, therefore, ensures companies demand management, supply chain optimization and revenue maximization, which can potentially significantly boost their competitive advantage[20].

Furthermore, the scalability and flexibility of NoSQL databases is key to maintaining competitiveness, as companies can quickly adapt to new information and trends, ensuring a continuous and efficient supply. NoSQL databases opti-mize data management processes, thus helping businesses reduce operational costs, improve decision-making and innovate faster than their competitors.

A real-world example, a Netflix case study, will illustrate how companies can use NoSQL databases to optimize their offer and improve business per-formance and competitiveness.

This purpose requires the application of the Two-Sided Markets Theory, which studies markets where two different groups of users communicate through a platform that facilitates their interactions. These platforms, such as Netflix and Spotify, connect content producers (such as movie studios and music artists) with consumers who consume that content. A key concept in this theory is the network effect, where the value of a platform increases with the number of users on both sides of the market. The more users who watch movies and series on Netflix or listen to music on Spotify, the greater the motivation for content producers to market their content on these platforms, and vice versa. This creates a positive feedback loop that can rapidly increase market size and value. Netflix and Spotify must carefully balance attracting and retaining users on both sides. The costs associated with attracting new users and maintaining a balance between supply and demand are key to the success of these platforms. For example, platforms often use algorithms for personalized recommendations to increase user sati-sfaction and keep them on the platform for longer. In other words, they use NoSQL databases to analyze user data in real time, providing personalized recommendations that increase user engagement and satisfaction, thereby optimizing content offer. In this context, business informatics and databases play a key role in the collection, processing, and analysis of data necessary for the optimization of these processes.

Market matching requires efficient analysis and processing of large data volumes. NoSQL databases allow companies to dive deeper into market trends and customer behavior for more accurate matching of supply and demand. As mentioned above, companies such as Spotify and Netflix use NoSQL databases to analyze user data in real time, which allows them to provide personalized recommendations and optimize content offerings. The scalability and flexibility

[20] Alpár (2010), p. 166

of NoSQL databases are key to maintaining competitiveness, allowing businesses to adapt more quickly to new information and trends. In short, NoSQL databases play a key role in enabling businesses to stay ahead of the competition by providing tools to optimize markets and improve business performance.

3.1. Technological support in data collection and processing

Data collection and processing are the basic processes that allow companies to get to know the market trends, optimize their offer and maintain competitiveness. The process can be broadly categorized into several stages: data collection, data storage, data processing, and data analysis.

Data collection means collecting information from different sources. These sources can be broadly categorized into transactional data, customer data, web and mobile analytics, social media data, and external data. Transaction data is collected from business transactions, such as sales, purchases, and payments, often recorded in real time and including details such as transaction amounts, dates, and parties involved. Customer data is collected from interactions and include personal data, purchase history, preferences, feedback, and behavioral data. It can be obtained from CRM systems, surveys, social media, and customer service interactions. Web and mobile analytics collects data about website and mobile app interactions, including clickstream data, page views, session duration, navigation paths, and user demographics. It uses tools such as Google Analytics and Adobe Analytics. Social media data comes from platforms such as X (formerly known as Twitter), Facebook, and Instagram and provide insight into public sentiment and trends. External data is obtained from third-party vendors, including market research, economic reports, and industry-specific databases. It helps businesses understand broader market conditions and trends.

Once collected, data must be stored securely and must be accessible. Different data storage solutions are used depending on data nature and volume. Relational database management systems such as MySQL, PostgreSQL, and Oracle are suitable for structured data that fits into predefined schemas. NoSQL database management systems such as MongoDB, Cassandra, and Couchbase handle unstructured or semi-structured data, offering flexibility and scalability. Data lakes, centralized repositories such as Hadoop and Amazon S3, store vast volumes of raw data in its native format, whether structured, semi-structured, or unstructured. Storage services on cloud platforms such as AWS, Google Cloud, and Microsoft Azure provide scalable, on-demand storage solutions capable of handling different volumes and types of data.

Data processing means transforming raw data into a usable format. This includes data cleaning, which removes or corrects inaccurate, incomplete, or irrelevant data to ensure quality and reliability. Data integration combines data from different sources into a single view, often merging databases, reconciling conflicting data, and ensuring consistency. Data transformation converts data into the required format or structure, including data normalization, encoding cate-gorical variables, and data aggregation into summary statistics. Data enrichment enhances a dataset by adding information from external sources, such as adding demographic data to customer records or adding market data to sales transactions.

After processing, the data is analyzed in order to obtain meaningful insights. Descriptive analytics summarizes historical data to understand past events, generating reports, dashboards, and visualizations. Diagnostic analytics investigates the root causes of historical outcomes using techniques such as data mining and root cause analysis. Predictive analytics uses statistical models and machine learning algorithms to predict future outcomes based on historical data, including sales forecasting, customer behavior prediction, and risk identification. Prescriptive analytics recommends actions based on data insights, using optimization algorithms, simulation models, and decision analysis to come up with the best course of action.

Various tools and technologies are used during data collection and processing. ETL (Extract, Transform, Load) tools such as Apache NiFi, Talend, and Informatica automate data extraction from sources, transformation into appropriate formats, and loading into storage systems. Big data platforms such as Apache Hadoop, Spark, and Flink are frameworks for efficient large-scale data processing. Data warehouses such as Amazon Redshift, Google BigQuery, and Snowflake offer scalable and fast query performance for large datasets. Machine learning frameworks such as TensorFlow, Scikit-Learn, and PyTorch facilitate the development and deployment of machine learning models.

The integration of various data sources and the use of advanced analytical tools are key to a comprehensive understanding of customer preferences and habits. This integration allows companies to create a single view of the customer, combining data from multiple touch points to gain deeper insights. In addition, Python has emerged as a powerful data analysis tool due to its extensive libraries, for instance Pandas, NumPy, and SciPy, which facilitate data manipulation, statistical analysis, and machine learning[21].

4. Providing desired content to users on Netflix

The importance of matching supply and demand in streaming platforms is gaining more and more importance in the conditions of accelerated deve-

[21] Python's versatility and ease of use make it an invaluable asset for businesses aiming to analyze customer data and derive actionable insights.

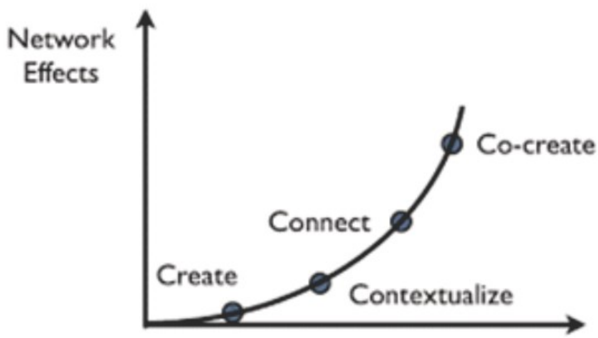
lopment of diverse digital platforms. Each platform is expected not only to commit to attracting new users, but to finding ways to retain them (by consistently offering content that meets their preferences). Effective supply (content) and demand (user preferences) matching increases user satisfaction and engagement, which in turn drives subscription renewals and overall platform growth. This section aims to show how Netflix successfully manages these complex processes using advanced data analytics and a two-sided market model.

4.1. Netflix as a two-sided market

Two-sided markets, or two-sided platforms, are economic models that facilitate interaction between two different groups of users. For Netflix, these two groups are content creators (such as movie studios and independent producers) and content consumers (subscribers). Netflix acts as an intermediary connecting these groups, creating value for both parties by offering a platform where supply meets demand.

In the Netflix context, the platform invests heavily in producing and acquiring a wide range of content to cater to the diverse tastes of its global audience. At the same time, it uses sophisticated recommendation algorithms to ensure that users can easily find content that matches their preferences, thus improving their viewing experience and encouraging longer engagement on the platform.

Figure 1. Four stages of network effects on digital platforms

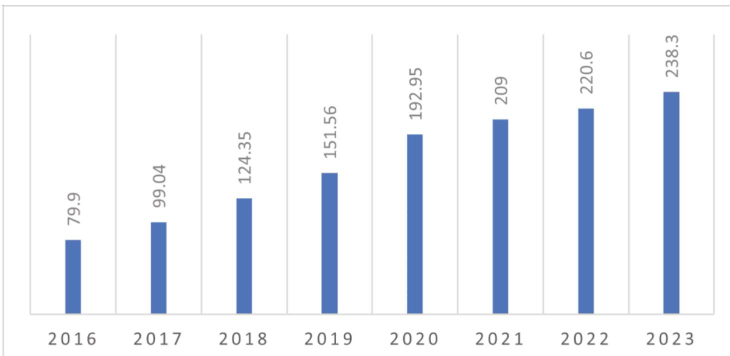


Source: https://corporatefinanceinstitute.com/resources/economics/what-is-network-effect/

Network effects are a key component of two-sided markets, where the value of the platform increases as more users and content providers join. Figure 1 shows the process of network effects, starting with content creation, followed by contextualizing user experience, then connecting users and content, and finally co-creating value on the platform.

For Netflix, positive network effects mean that as its subscriber base grows, more content creators are encouraged to partner with the platform, resulting in a richer and more diverse content library. This, in turn, attracts more users, creating a cycle of growth. Figure 2 shows the continuous growth in the number of Netflix subscribers from 2016 to 2023, which directly contributes to the platform's network effects.

Figure 2. Number of Netflix subscribers (in millions)

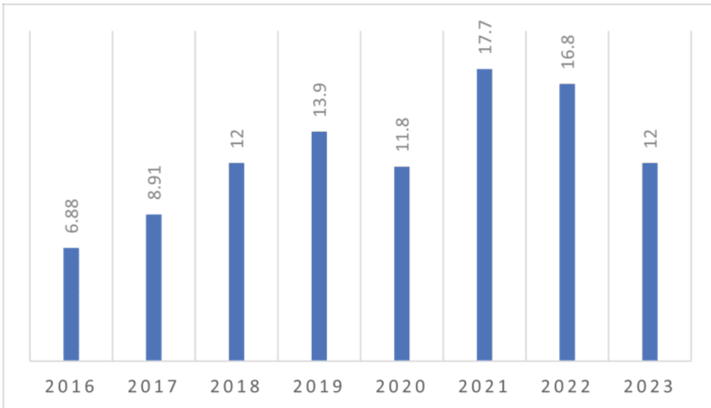


Source: https://www.businessofapps.com/data/netflix-statistics/

Netflix has to strike a balance between attracting new users and retaining existing ones. The key to success lies in the use of advanced algorithms for personalizing recommendations, which increases user satisfaction and prolongs their loyalty to the platform. In addition, content optimization in line with user needs reduces search costs and increases platform efficiency.

Figure 3 shows Netflix's total investment in content from 2016 to 2023. This figure shows how much Netflix has been investing in content acquisition and production to satisfy its growing user base. Fluctuations in content investment can be linked to changes in the number of subscribers, indicating the existence of network effects. That is, when Netflix increases its investments in the quality and variety of content, it attracts new users, thereby further increasing the value of the platform for all users. A decline in content spending in 2022 and 2023 can be attributed in part to the Hollywood screenwriters' strike, which affected the availability of new content and the attraction of new users. Due to reduced investments, network effects weakened, which could be reflected in the growth rate of the user base. Still, Netflix has continued to invest heavily in content, recognizing the importance of maintaining network effects to maintain a competitive edge and satisfy a growing user base.

Figure 3. Investment in Netflix content (billions of dollars)



Source: https://www.businessofapps.com/data/netflix-statistics/

The combination of user growth and increased content spending clearly illustrates network effects at work. As the user base expands, Netflix is able to attract more content producers, which in turn enriches the platform and attracts even more users.

Netflix relies on its own personalized recommendation system, which uses vast volumes of user data, to predict and suggest content that individual users are likely to enjoy. Netflix continuously refines these algorithms, thereby not only improving user satisfaction, but also helping users discover new content and maximizing the utility of its content library. By driving powerful network effects and leveraging advanced data analytics, Netflix ensures that its platform remains the preferred choice for both content creators and consumers, maintaining its competitive edge in the streaming industry.

4.2. Data collection and storage for recommendations

Collecting data on user preferences is key to providing personalized content recommendations. Various sources and techniques are used to collect this information. Web analytics, which includes data collected from user interactions within web platforms, such as page views, clicks, and time spent on different pages, plays a vital role. App analytics, which records user interactions within mobile or desktop applications, is another important data source. In addition, user profiles, which contain information directly provided by users, such as demographic details, preferences, and previous interactions, are invaluable for understanding user behavior.

Data collection techniques include cookies, small data pieces stored on the user's device that track user behavior during different sessions[22]. User behavior tracking involves tracking how users navigate and interact with the platform in real time, providing insight into their preferences. Feedback, in the form of user reviews, ratings and survey responses, further helps understand user preferences. This comprehensive data collection is key to content recommendation, as accurate data on user preferences enables the creation of personalized recommendations, thereby increasing user satisfaction and engagement.

In order to manage the vast volumes of data collected, Netflix implements a hybrid data storage strategy that includes both relational and NoSQL databases. To handle transactional data that requires strict consistency, such as billing information and user account details, Netflix uses traditional SQL databases. To handle big data such as user interactions and preferences, Netflix relies on Apache Cassandra, a NoSQL database management system known for its scalability and high performance. Cassandra allows Netflix to seamlessly distribute data across multiple servers, ensuring fast downloads.

In addition, Netflix uses Apache Druid for real-time analytics to ensure a high-quality user experience. Druid allows over 2 million events per second to be captured and over 1.5 trillion rows to be queried in near real-time, enabling engineers to spot anomalies and effectively optimize the streaming service. Using Druid, Netflix can monitor the performance of over 300 million devices worldwide and conduct A/B testing to assess the impact of updates and changes on the user experience[23].

Netflix also tested the scalability of the Cassandra database on AWS, achieving over a million records per second. This scalability allows it to handle massive data volumes with high performance and reliability, ensuring efficient processing and storage of user preferences and interactions.

4.3. Using data for personalized content recommendations on Netflix

After data collection, Netflix ensures accuracy and reliability through processes such as data cleaning, integration, and transformation. Data cleaning removes inaccurate and incomplete information, while integration brings together data from different sources into a consistent format. Data transformation adapts the data to the required format for further analysis, thus

[22] Tappenden and Miller (2008), p. 133
[23] https://netflixtechblog.com/benchmarking-cassandra-scalability-on-aws-over-a-million-writes-per-second-39f45f066c9e

enabling the efficient management of large amounts of information using tools such as Talend[24] and Informatica[25].

Netflix analyzes user preferences and adapts content to user needs. Descriptive analytics helps understand past events, while predictive analytics predicts future trends, enabling personalized content recommendations. Economists applied matching algorithms and created markets where they did not exist, or made them less congested, with sufficient participants and with incentives that encouraged participants to reveal their private information[26].

The paper by Gomez-Uribe and Hunt[27] describes the key algorithms that make up Netflix's recommendation system. The first is the Personalized Video Ranker (PVR), which personalizes content within genre lines by combining personalization and popularity. Then there is the Top-N Video Ranker, which generates the best personalized recommendations, focusing on the most appro-priate recommendations for a particular user. Trending Now recognizes short-term trends to display currently popular content, while Continue Watching ranks titles the user is already watching, predicting continued viewing. The algorithm Video-Video Similarity (Sims) is used in rows like "Because You Watched", recommending similar titles. The Page Generation algorithm generates the homepage, selecting relevant rows and optimizing their diversity. Evidence Selection algorithms choose what information about a video (e.g. synopsis, awards) to show users, while search algorithms help find specific titles and search for alternatives when the desired content is not available. Together, these algo-rithms enable a personalized experience and optimize user interaction with the platform[28].

Personalization relies on advanced machine learning and increases user engagement and retention on the platform[29], which is key to its success. Personalization algorithms take into account users' viewing history, searches and ratings to offer relevant recommendations. These algorithms adapt content to individual preferences, which not only increases user satisfaction but also pro-longs their activity on the platform. According to Lalle and Conati[30], perso-nalization must take into account individual differences among users to be effective. Machine learning algorithms are becoming the focal point of re-

[24] https://www.talend.com/resources/what-is-database-integration/
[25] https://thecuberesearch.com/ai-powered-metadata-informaticas-role-in-the-future-of-data-management/
[26] Trifunović (2019), p. 5
[27] Gomez-Uribe and Hunt (2015), p.13:2
[28] Gomez-Uribe and Hunt (2015), p. 13:5
[29] https://abmatic.ai/blog/impact-of-personalization-on-user-engagement-and-retention
[30] Lalle and Conati (2019), p. 329

commendation systems, enabling content personalization and improving user experience.

Thus, content personalization through advanced machine learning algorithms allows Netflix to meet individual user needs. This process results in a dynamic user experience that adapts content in real time, allowing users to quickly find what interests them. In the future, further development of personalization may improve platforms' ability to predict changes in user preferences even more, thereby further improving their experience and satisfaction with the platform.

5. Conclusion

In a world increasingly reliant on digital technologies, the effective data management and use has become crucial to maintaining a competitive advantage. Scalable and flexible NoSQL databases are a key tool in optimizing market matching, allowing companies such as Netflix to effectively match supply with demand. By using advanced algorithms and analytical tools, Netflix not only adjusts its content offering according to user preferences, but also uses network effects to increase the value of its platform.

Effective supply and demand matching and delivering content to users who want to see it is critical to the success of the platforms. Data collection and analysis allow Netflix to understand user preferences, optimize its content offering and maintain competitive advantage. Despite the challenges associated with the adoption of NoSQL databases, their role in enabling real-time analytics and content personalization confirms their value in modern business.

Recommendations for future research or implementation include further development of advanced personalization algorithms and exploration of new data analysis techniques that can further improve user experience. In particular, research could focus on integrating artificial intelligence algorithms with NoSQL databases to improve the ability to predict user preferences. Furthermore, exploring possibilities for deeper integration of NoSQL databases with different data sources, for example IoT devices and social networks, could open new opportunities for even more precise content personalization. In addition, future research can focus on the development of hybrid models that combine the advantages of relational and NoSQL databases, allowing companies to take advantage of the best of both systems in the context of data management. Research into new approaches to data storage optimization and latency reduction could further improve the performance of streaming platforms. Finally, an exa-mination of the ethical aspects of collecting and processing large amounts of user data remains crucial in order to ensure the protection of user rights in an increasingly complex digital environment.

References:

Acharya, B., Jena, A. K., Chatterjee, J. M., Kumar, R. and Le, D. N. (2019), "NoSQL Database Classification: New Era of Databases for Big Data", International Journal of Knowledge-Based Organizations , Vol. 9(1), Pp. 50-65, https://doi.org/10.4018/IJKBO.2019010105

https://thecuberesearch.com/ai-powered-metadata-informaticas-role-in-the-future-of-data-management/ [Accessed: 15/08/2024]

Alpár, F. Z. (2010), "Matchmaking Framework for B2B E-Marketplaces", Informatica Economica, 14(4).

https://netflixtechblog.com/benchmarking-cassandra-scalability-on-aws-over-a-million-writes-per-second-39f45f066c9e [Accessed: 29/08/2024]

Choi, Y. L., Jeon, W. S. and Yoon, S. H. (2014), "Improving Database System Performance by Applying NoSQL", Journal of Information Processing Systems, Vol. 10(3), Pp. 355-364, https://doi.org/10.3745/JIPS.04.0006

Čivić, B. and Ćilimković, D. (2022), "Key Features of Brand Management and Customer Behavior in the Market of Food Products in Bosnia and Herzegovina", Ekonomske ideje i praksa, (45), https://doi.org/10.54318/eip.2022.bc.323\3

Codd, E. F. (1979), "Extending the Database Relational Model to Capture More Meaning", ACM Transactions on Database Systems, Vol. 4(4), Pp. 397-434, https://doi.org/10.1145/320107.320109

Gomez-Uribe, C. A. and Hunt, N. (2015), "The Netflix Recommender System: Algorithms, Business Value, and Innovation", ACM Transactions on Management Information Systems, Vol. 6(4), Pp. 1-19, https://doi.org/10.1145/2843948

Hemmatpour, M., Montrucchio, B., Rebaudengo, M. and Sadoghi, M. (2020), "Analyzing In-Memory NoSQL Landscape", IEEE Transactions on Knowledge and Data Engineering, Vol. 34(4), Pp. 1628-1643, https://doi.org/10.1109/TKDE.2020.3002908

Idreos, S. and Callaghan, M. (2020), "Key-Value Storage Engines", in Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data, Pp. 2667-2672, https://doi.org/10.1145/3318464.3383133

https://abmatic.ai/blog/impact-of-personalization-on-user-engagement-and-retention [Accessed: 17/06/2024]

Jatana, N., Puri, S., Ahuja, M., Kathuria, I. and Gosain, D. (2012), "A Survey and Comparison of Relational and Non-relational Database", International Journal of Engineering Research & Technology, Vol. 1(6), Pp. 1-5.

Kabakus, A. T. and Kara, R. (2017), "A Performance Evaluation of In-Memory Databases", Journal of King Saud University-Computer and Information Sciences, Vol. 29(4), Pp. 520-525, https://doi.org/10.1016/j.jksuci.2016.06.007

Lallé, S. and Conati, C. (2019, March), "The Role of User Differences in Customization: A Case Study in Personalization for Infovis-Based Content", in Proceedings of the 24th International Conference on Intelligent User Interfaces, Pp. 329-339, https://doi.org/10.1145/3301275.3302283

Maté, A., Peral, J., Trujillo, J., Blanco, C., García-Saiz, D. and Fernández-Medina, E. (2021), "Improving Security in NoSQL Document Databases Through Model-Driven Modernization", Knowledge and Information Systems, Vol. 63, Pp. 2209-2230, https://doi.org/10.1007/s10115-021-01589-x

Meier, A. and Kaufmann, M. (2019), SQL & NoSQL Databases, Springer Fachmedien Wiesbaden, Wiesbaden, https://doi.org/10.1007/978-3-658-24549-8

Nayak, A., Poriya, A. and Poojary, D. (2013), "Type of NoSQL Databases and its Comparison with Relational Databases", International Journal of Applied Information Systems, Vol. 5(4), Pp. 16-19.

https://www.businessofapps.com/data/netflix-statistics/ [Accessed: 20/08/2024]

Song, Y., Zhu, Y. and Li, L. (2014, October), "Large Scale Data Storage and Processing of Insulator Leakage Current Using Hbase And Mapreduce", in 2014 International Conference on Power System Technology, Pp. 1331-1337, https://doi.org/10.1109/POWERCON.2014.6993650

Strauch, C., Sites, U. L. S. and Kriha, W. (2011), NoSQL databases. Lecture Notes, Stuttgart Media University, www. christof-strauch.de/nosqldbs.pdf

Tappenden, A. and Miller, J. (2008, April), "A Three-Tiered Testing Strategy for Cookies", in 2008 1st International Conference on Software Testing, Verification, and Validation, Pp. 131-140, https://doi.org/10.1109/ICST.2008.18

Thakur, N. and Gupta, N. (2021), "Relational and Non Relational Databases: A Review", Journal of University of Shanghai for Science and Technology, Vol. 23(8), Pp. 117-12, https://doi.org/10.51201/JUSST/21/08341

Tiwari, S. (2011), Professional NoSQL, John Wiley & Sons.

Trifunović, D. (2019), Uparivanje na tržištima bez cena [Matching in No-Price Markets], CID Ekonomski fakultet.

Trifunović, D. (2023), "Aukcije sa višedimenzionalnim signalima [Auctions With Multi-dimensional Signals]", Ekonomske ideje i praksa, (50), https://doi.org/10.54318/eip. 2023.dt.353

https://www.talend.com/resources/what-is-database-integration/ [Accessed: 17/06/2024]

https://corporatefinanceinstitute.com/resources/economics/what-is-network-effect/ [Accessed: 05/08/2024]

Primljen (Received): 27.9.2024.
Prihvaćen (Accepted): 17.10.2024.
Pre štampe (Online First) 24.2.2025.