



# Open-Source vs. Commercial Vendor Software in the Enterprise

Presented by: William McKnight

#1 Global Influencer in Big Data Thinkers360

President, McKnight Consulting Group

3 X Inc 5000

 /in/wmcknight

www.mcknightcg.com  
(214) 514-1444



# The Open Source Definition



- Free Redistribution
- Source Code
- Derived Works
- Integrity of the Author's Source Code
- No Discrimination Against Persons or Groups
- No Discrimination Against Fields of Endeavor
- Distribution of License
- License Must Not Be Specific to a Product
- License Must Not Restrict Other Software
- License Must Be Technology-Neutral

# What is Open Source



- Open source software is software that is distributed with its source code, giving anyone the right to use, study, change, and distribute the software in any form.
- Open source software is often developed and maintained by a community of volunteers, but it can also be developed by companies and organizations.
- Open source software is used in a wide variety of applications, including operating systems, web browsers, office suites, and games.
- Open source software is often more reliable and secure than proprietary software because it is constantly being reviewed and updated by the community.
- Open source software is often more affordable than proprietary software because it is free to use and distribute.
- Open source software is often more customizable than proprietary software because anyone can modify the source code to meet their specific needs.

# What is Open Source



- Open source software is often more innovative than proprietary software because the community is constantly developing new features and improvements.
- Open source software is often more transparent than proprietary software because the source code is available for anyone to review.
- Open source software is often more ethical than proprietary software because it is developed and maintained in a collaborative and open manner.
- Open source software is often more supportive of user privacy than proprietary software because it is not controlled by a single company that may be motivated to collect and sell user data.
- Open source software is often more accessible to people with disabilities than proprietary software because it is often developed with accessibility in mind.
- Open source software is often more sustainable than proprietary software because it is not controlled by a single company that may go out of business.

# Stages of Open Source (CNCF)

- Sandbox
- Incubation
- Graduation

# Open Source Licensing Models

- Permissive licenses allow users to do almost anything with the software, including using it in commercial products.
- Copyleft licenses require that any derivative works of the software also be released under the same license. This means that if you use code from a copyleft-licensed project in your own project, your project must also be released under the same license.

# Popular Open Source Tools in Analytics

- MySQL
- PostgreSQL
- SQLite
- MariaDB
- MongoDB
- Cassandra
- Redis
- Elasticsearch
- CouchDB
- Neo4j
- CockroachDB
- ArangoDB

# More Recent Examples

- LLaMA
- gRPC
- Apache Iceberg indexes
- Jupyter kernel extensions
- DuckDB
- LLMs





# Benefits of Open Source

- Cost Savings
- Community Support
- Transparency
- Flexibility and Customization
- No Vendor Lock-In
- Community-Driven Innovation
- Cross-Platform Compatibility
- Licensing Freedom
- Ecosystem and Integrations
- Reliability

# Downsides of Open Source

- Documentation Quality
- Integration Challenges
- Security Concerns
- Learning Curve
- Limited Features
- Ongoing Maintenance
- Community Dependence
- License Complexity
- Lack of Vendor Accountability
- Hidden Costs

# How to Contribute to Open Source

- Write code: This is the most common way to contribute to open source software. You can help to fix bugs, add new features, or improve the performance of existing code.
- Test code: Testing is essential for ensuring the quality of open source software. You can help by testing new code, existing code, and code that has been modified.
- Write documentation: Documentation is important for helping users understand how to use open source software. You can help by writing new documentation, updating existing documentation, or translating documentation into other languages.
- Report bugs: If you find a bug in open source software, you can report it to the developers so that they can fix it.
- Help with support: You can help to answer questions from other users of open source software, or you can help to troubleshoot problems that they are having.
- Donate money: If you don't have the time or skills to contribute code, you can still help open source software by donating money to developers.

# Contributing Code to Open Source

- Choose a project
- Set up a development environment
- Fork the repository
- Clone the fork
- Create a branch
- Edit the code
- Commit the changes to the fork
- Create a pull request



# Features Typically Close Sourced

- Enterprise-Grade Support
- Advanced Security Features
- Vendor Integration and Ecosystem
- Proprietary Algorithms or Models
- Customization Without Development Resources
- Certifications and Compliance
- Proprietary File Formats or Protocols
- Advanced Reporting and Analytics
- Warranty and Liability Protections

# Databases

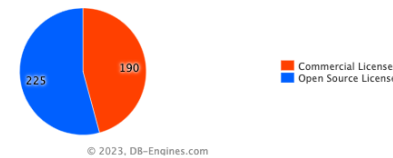
- Open Source databases are increasingly popular
- Oracle, SQL Server, and DB2 will stay for the legacy applications, and open source is considered for new applications, or modernization
- Some have migrated to PostgreSQL or MySQL

## Popularity of open source DBMS versus commercial DBMS

Based on the measurements we are using for our [ranking of database management systems](#), we analyze the popularity of open source tools compared to the popularity of commercial systems.

A system is classified as open source, when the source code is freely available and may be used and modified according to respective licenses.

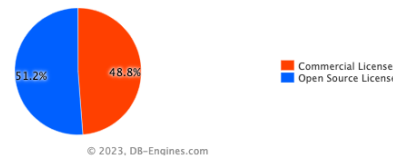
### Number of systems, October 2023



DB-Engines lists 415 different database management systems.

This pie chart shows the number of open source and commercial systems.

### Popularity scores, October 2023

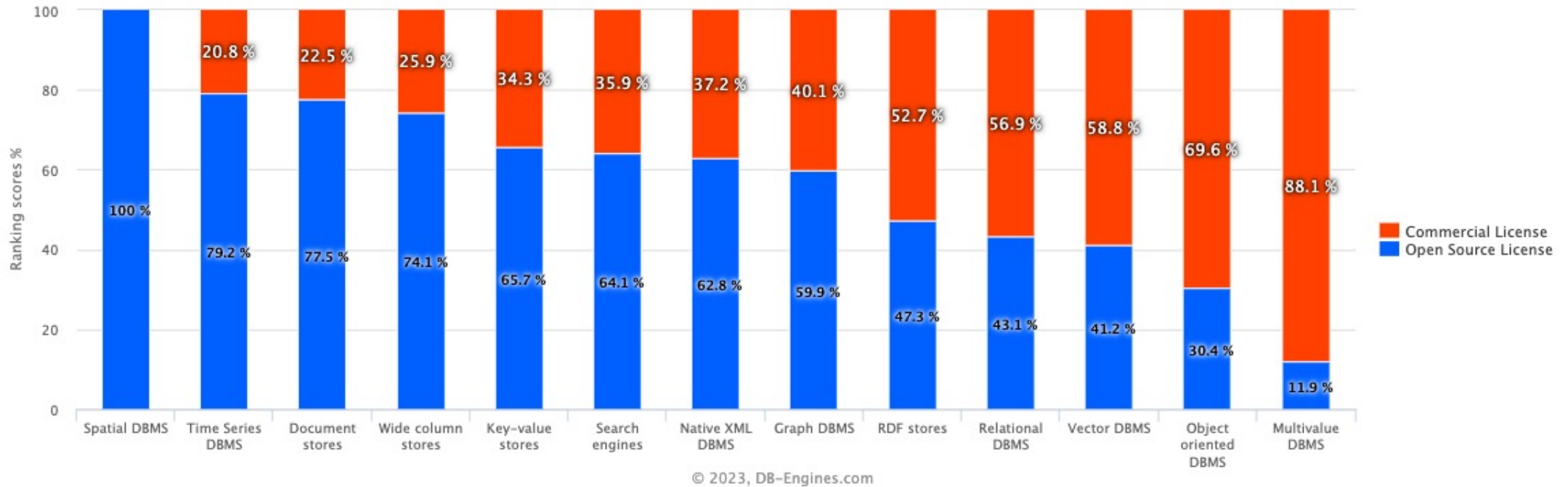


This pie chart compares the popularity of open source and commercial systems. It is calculated by summing up the popularity (i.e. the ranking scores) of all individual systems.

Source: db-engines.com

# Popularity by database model

## Popularity broken down by database model, October 2023



This chart demonstrates that the popularity of open source tools heavily depends on their database model.

Source: db-engines.com

# Enterprises Using Both

- Kaiser Permanente: Kaiser Permanente uses a mix of open- and closed-source software in its integrated healthcare system, including Cerner Corporation for its EHR system, Red Hat Enterprise Linux for its operating system, and OpenVista for its ambulatory care EHR system.
- Mayo Clinic: Mayo Clinic uses a mix of open- and closed-source software in its hospitals and clinics, including Epic Systems for its EHR system, Red Hat Enterprise Linux for its operating system, and R for its research and development.
- JPMorgan Chase: JPMorgan Chase uses a mix of open- and closed-source software in its banking and investment operations, including Oracle Financial Services Suite for its core banking system, Linux for its operating system, and Apache Hadoop for its big data processing.
- Pfizer: Pfizer uses a mix of open- and closed-source software in its drug discovery and development process, including SAP for its enterprise resource planning (ERP) system, Linux for its operating system, and R for its data analysis.
- Merck: Merck uses a mix of open- and closed-source software in its drug discovery and development process, including IBM Watson for its artificial intelligence platform, Linux for its operating system, and R for its data analysis.



# “Open Source is Free”

- What will you do when you encounter an issue or when a security leak is exposed? Will you modify the source yourself, recompile and deploy it? Of course not! You need a support contract, you need consultants, or you need dedicated teams. This is not free.
- it's generally unacceptable to run your company's core business with software on which you cannot fix bugs or apply new security patches.
- if you don't accept all responsibilities, you probably need commercial support to use it at the enterprise level.
- the savings on a commercial license can be invested in support, training, and team, you will get more freedom, trust, and control over the software that processes your data

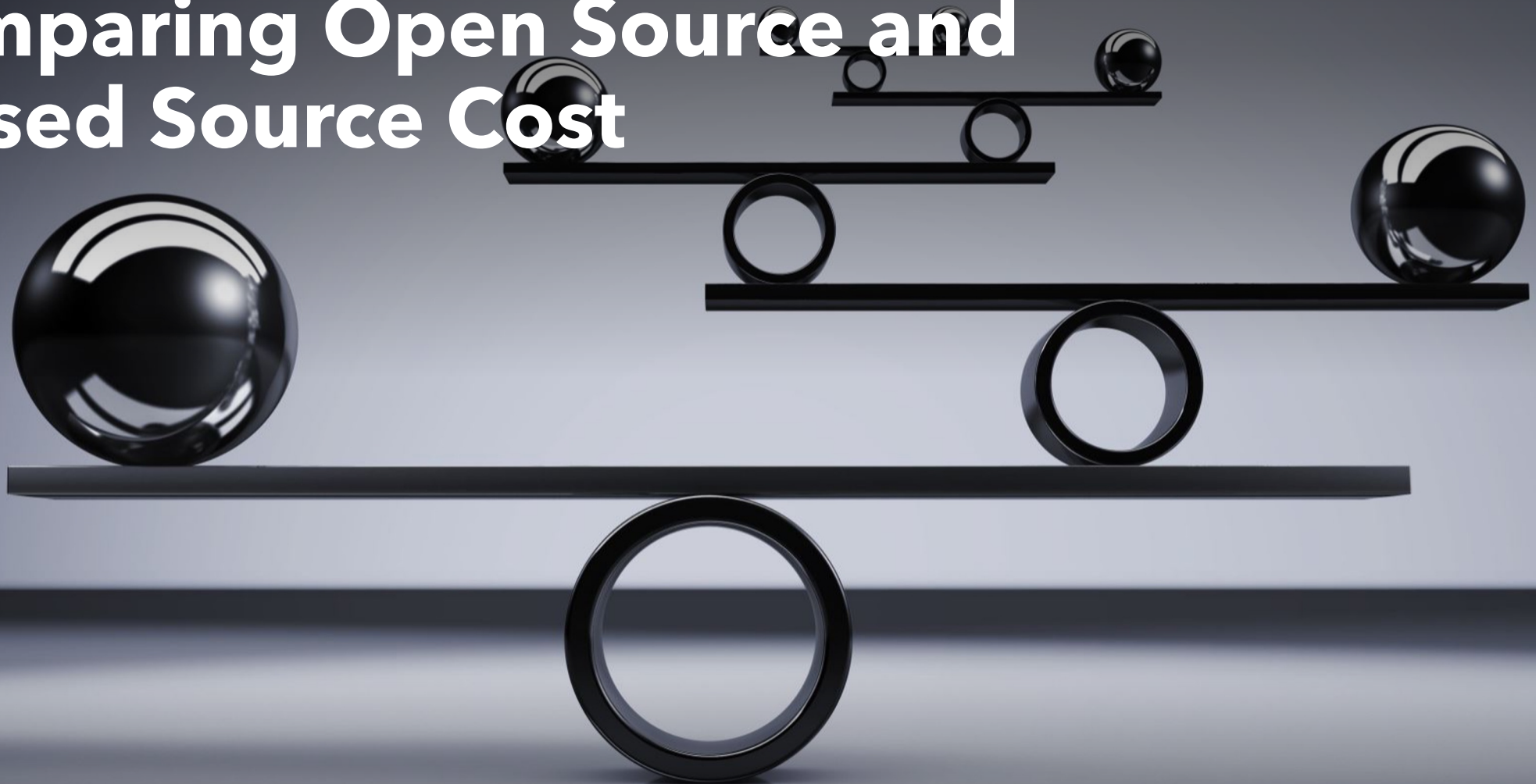
# Security of open-source software vs. commercial vendor software

- Open Source
  - Transparency
  - Community/Lack of Vendor Support
  - Complexity
- Commercial Vendor Software
  - Vendor Support
  - Rigorous Security Testing
  - Lack of Transparency
  - Cost

# Support for open-source software vs. commercial vendor software

- The quality and quantity of community support for open-source software can vary depending on the project. Some projects have very active communities, while others have less active communities.
- Commercial vendor software often comes with SLAs that guarantee a certain level of support. This can be important for businesses that need a high
- Consider if your team has the technical expertise to troubleshoot problems with open-source software
- An example, for PostgreSQL, you have the mailing lists, slack channel, telegram, and of course Stack Overflow.

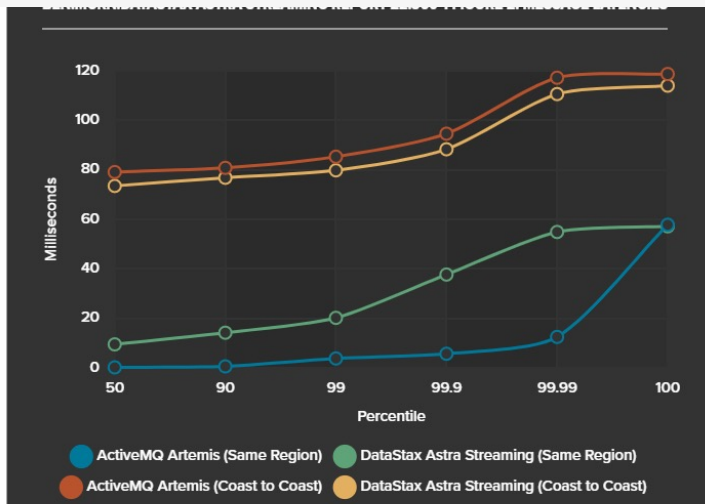
# Comparing Open Source and Closed Source Cost



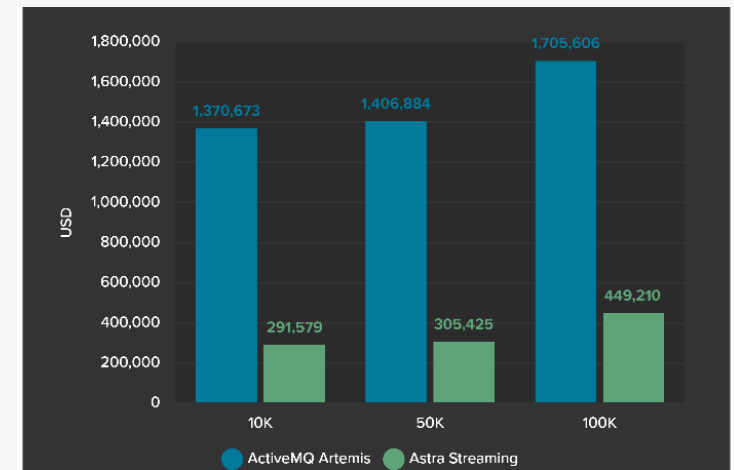
# Benchmark: Astra Streaming with Starlight for JMS and Self-Managed OSS ActiveMQ Artemis JMS

- With Astra Streaming, architecture is consolidated and simplified along with many of the management, administration, and disaster recovery functions inherent in a self-managed platform.
- The performance and resiliency of the fully managed Astra Streaming can keep up with ActiveMQ Artemis(open source) without the burden of scaling out infrastructure (or scaling down when demand is light). You simply pay for what you use, and DataStax manages the back end.
- We found that in situations where messages per second throughput rapidly and frequently spikes and varies (bursting), Astra Streaming was 2x cheaper in infrastructure costs and up to 4x cheaper in total cost of ownership.
- We found that modernizing a streaming JMS data environment to fully managed Astra Streaming would have many benefits and capability enhancements, including real-time data integration, analytics, and AI/ML applications.

# Benchmark Artifacts



ACTIVEMQ ARTEMIS		ASTRA STREAMING	
DBA FTE	1	DBA FTE	0.25
Infrastructure FTE	1	Infrastructure FTE	0
Total FTE	2	Total FTE	0.3
Total \$/Year	\$366,000	Total \$/Year	\$45,750
3-Year Total	\$1,098,000	3-Year Total	\$137,250



# Decision Making Factors

- Cost
- Features
- Customization
- Support
- Security
- Community
- Vendor lock-in
- Transparency
- Innovation
- Licensing
- Long-term support
- Team Expertise
- Compliance Requirements

# The Future of Vendor Software

- Enterprise-level features, like security, partitioning, and parallelism, better instrumentation, troubleshooting tools, performance, and resilience options are added to closed source databases.
- More unbalanced situations
- Vendors will provide a smaller subset of their features for free
- Enterprise features, like backups, scalability, and encryption come with a commercial license

The Register®

## What is Google doing with its open source teams?

Nothing good – the recent layoffs hit its best and brightest leaders hard

[Steven J. Vaughan-Nichols](#)

Fri 27 Jan 2023 / 10:34 UTC

**OPINION** Remember when Google's motto was "Don't be Evil"? I do. Even though Google dumped that phrase from its code of conduct in 2018, many of us still thought Google was a bit better than other companies.

We were wrong. Those who were fired last week found out from emails, discovering they no longer had corporate access and their ID badges no longer worked. Class act, Google.

How were they chosen? Good question. It has been widely reported that some of the firing was done by an algorithm. Those fired included staffers who had just received high performance reviews or held top managerial positions with annual compensation packages of \$500,000 to \$1 million. But, as I looked at who was being fired, something struck me. Many of those shown the doors were the best of the best in Google's open



# Recommendations

- Evaluate your needs carefully before choosing an open source solution
- Be aware of the risks associated with open source software
- Make sure you have the resources to support open source software
- Open source software is often free to use, but it is important to have the resources to support it
- Be prepared to keep your open source software up to date
- Be prepared to contribute to the open source community
- Ruggedized across the criteria: Linux, Kubernetes, Docker, Prometheus, Grafana, Apache Spark, Elasticsearch, Kibana, Nginx, HAProxy



# Open-Source vs. Commercial Vendor Software in the Enterprise

Presented by: William McKnight

#1 Global Influencer in Big Data Thinkers360

President, McKnight Consulting Group

3 X Inc 5000

 /in/wmcknight

www.mcknightcg.com  
(214) 514-1444

