ICONIC DATA

PRODUCT REVIEW

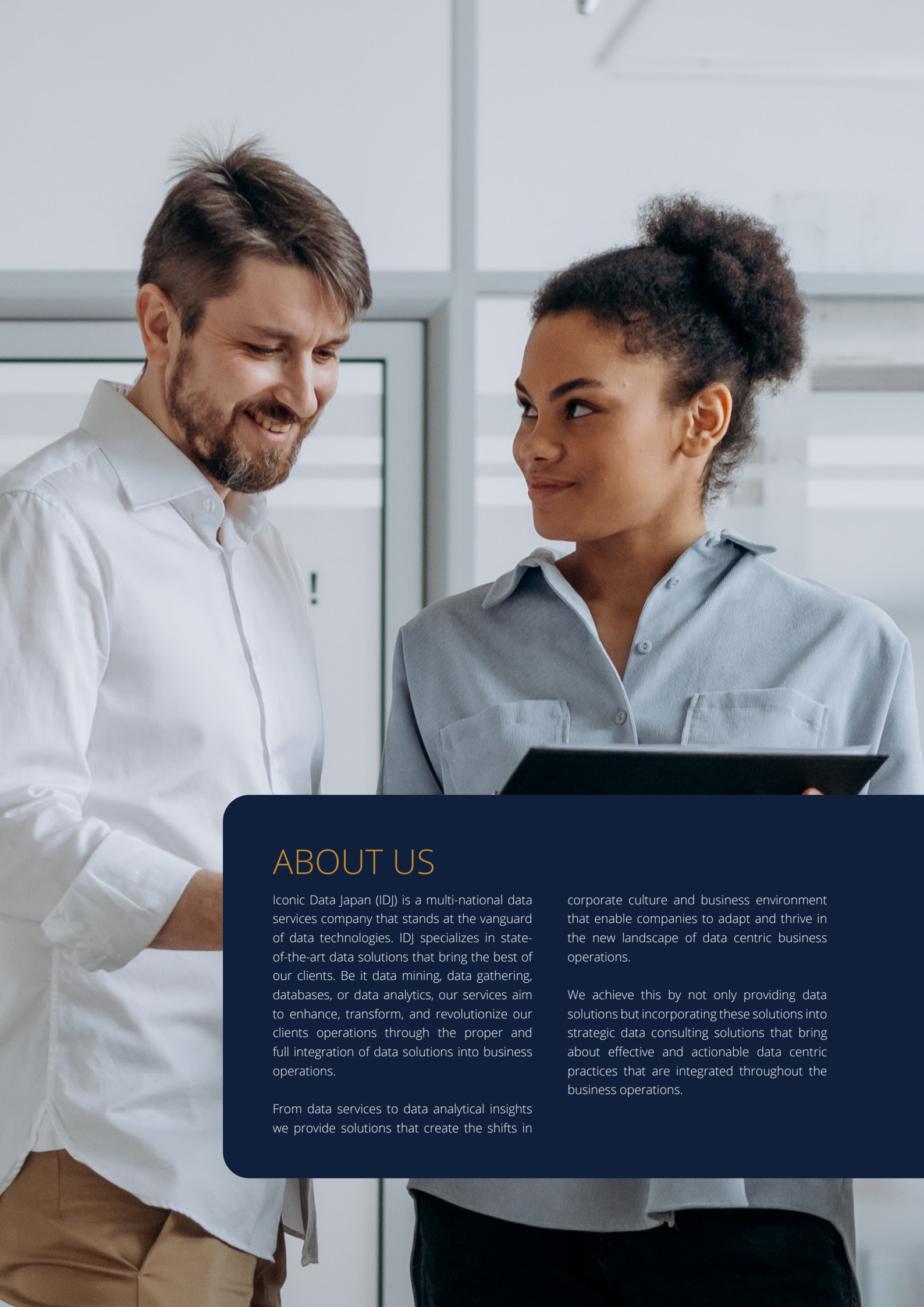
# TiDB

## THE LARGE SCALE RELATIONAL DATABASE SOLUTION

2022

Piloting tomorrow's creativity

https://www.iconic.inc



# ABOUT US

Iconic Data Japan (IDJ) is a multi-national data services company that stands at the vanguard of data technologies. IDJ specializes in state-of-the-art data solutions that bring the best of our clients. Be it data mining, data gathering, databases, or data analytics, our services aim to enhance, transform, and revolutionize our clients operations through the proper and full integration of data solutions into business operations.

From data services to data analytical insights we provide solutions that create the shifts in corporate culture and business environment that enable companies to adapt and thrive in the new landscape of data centric business operations.

We achieve this by not only providing data solutions but incorporating these solutions into strategic data consulting solutions that bring about effective and actionable data centric practices that are integrated throughout the business operations.

# REPORT
# CONTENT


02 EXECUTIVE SUMMARY

04 PRODUCT INTRO

05 FEATURES

06 SPECIFICS/ DRAWBACKS

08 COMPETITORS

09 CONCLUSION



PRODUCT REVIEW - 2022 - TIDB
ICONIC DATA, DATATECH



AUTHORS: FRANCISCO R. VASSALLO
CONTRIBUTORS: RUCHI MITTAL
AI IMAGES: DALL-E
EDITOR: SUSANNE HELLE

# TIDB AS A SOLUTION

# Executive Summary

TiDB is a new database solution that is targeted primarily to clients that need to handle very large databases, with very large frequency of queries. If those two problems describe your business then this database solution is very attractive, as it also features a number of other features that make it stand out from its older competitors.

Particularly in Japan a fully supported (in Japanese) database solution of this kind is particularly attractive and to an extent an anomaly in a market that leans very heavily on the community.

It also features custom tuned tools for a range of services from data migration to data transformation. Depending on the client's need these could eliminate the need for expensive ETL tools.

As well as its real-time data analytics system do give it a notable advantage over competitors in the same product category. (Large Databases and Fast Queries)

## TIDB

# A DATABASE SOLUTION

**TiDB is a NewSQL database system whose primary target audience is clients with very large databases, and/or databases that require processing securely at very high speeds queries.**

TiDB focuses on scalability, database clustering, and its ability to automatically scale horizontally (across nodes/instances/machines), another focus of TiDB is its speed in being able to handle a large number of reads and writes at high speed. This puts it in a very specific category of comparison, as it is not really a general database product, but is meant to serve a specific market. It is important to note that merely based on its scalability and its ability to process at speed TiDB is not particularly novel, other options in the market provide fairly similar features in this particular respect (Vitess and CRDB). It is instead a combination of these market-defining focuses and additional features that set TiDB as a truly attractive product for its target audience.

> TIDB IS A GOOD SOLUTION FOR THOSE THAT CAN TAKE ADVANTAGE OF IT'S FEATURES

## TIDB STANDOUT FEATURES:

- MySQL Compatible
- Real-Time Analytics
- High Level of Guarantees on data reads and writes
- An ecosystem of tools
- Support

## AUDIENCE

**Clients with large databases and the need to for high intensity queries would benefit the most.**

The type of customers that would most benefit from this Database solution are:

- Those with Large Databases (2TB+ Tables)
- Those with High Write Rates (Less than 20ms)
- Already use SQL databases
- Require high consistency

## Competitors:

- Vitess
- CRDB
- PHXSQL

## SPEED

**Core to the solution that TiDB is trying to be is speed, and the lack of delay in queries.**

TiDB feature similar optimizations to those that clickhouse has. These optimizations would normally make Clickhouse an ideal database solution for big data problems like registering clicks from popular websites, most database solutions cannot handle the speed to process the large amount of data coming in so fast, so dirty reads and writes are quite common without a specialized solution.

## SCALE

**TiDB's ability to incorporate horizontal scaling, with vertical scaling is market defining.**

TiDB features full scaleability in that its mark-out feature is horizontal scalability (across nodes) but it also features the much more common vertical scalability. This is present in many database solutions including ones that are meant for smaller scale databases, DynamoDB for example has this feature built into AWS services. This vertical scalability scales individual nodes based on traffic and need, but most of these solutions do not automatically scale across nodes.

# TIDB
# SPECIFIC FEATURES

## MYSQL COMPATIBILITY

MySQL compatibility for TiDB means existing queries work with this SQL database out of the box, and with little to no need for adaptation. It also means existing tools and services that support MySQL databases also support TiDB.

## ECOSYSTEM OF TOOLS

TiDB also enjoys an ecosystem of tools that facilitate various aspects of database management, and increase its attractiveness. These tools range from Data Migration tools like TiDB Data Migration, and TiDB Lighting, to database management packages like TiUp, and a fully up-to-date Operator. This is not an extensive list but is meant to give an overview of the tools available, but is meant instead to highlight that there exists a range of solutions that are up to date and supported. So for example Vitess has an Operator tool, this lacks support which results in issues while trying to set it up with Kubernetes.

## REAL-TIME ANALYTICS

In real-time analytics TiDB is using HTAP, which is also known as hybrid transaction/analytics processing, transactions—like updating a database—are combined with analytics—like identifying prospective customers. Both workloads i.e. Online Transactional Processing (OLTP) or Online Analytical Processing (OLAP) workloads may be supported by an HTAP database, which offers speed and simplicity. TiDB is using Storage engines, Data consistency, Data isolation, and the MPP computing engine of HTAP.

## HIGH LEVEL GUARANTEES ON DATA READS AND WRITES

Unlike some other architectures TiDB retains the higher level “repeatable read” system of baseline MySQL. This isolation level system allows increased guarantees of reads and writes.

**ECOSYSTEM**

**REAL-TIME ANALYTICS**

**HIGH LEVEL GUARANTEES**

# DRAWBACKS

## 01 TAKES GETTING USED TO

TiDB despite being highly compatible with MySQL still requires some getting used to. (Using what's best practice in MySQL, like Auto-Increment of Primary key, can result in not intended results like all write traffic going to one instance in the cluster.)

## 02 LAYER PROBLEMS

Change Data Capture (CDC) tools and processes that rely on the SQL layer, have some issues since TiCDC only receives changes at a higher TiKV level. There are workarounds but these can be inelegant.

## 03 EXPENSIVE IF NOT EXPLOITED

TiDB is relatively expensive if used for small datasets, further, its also harder to deploy compared to simpler smaller solutions. So this is a database solution with a specific market, and to that market it serves pretty well.

# TIDB COMPETITORS

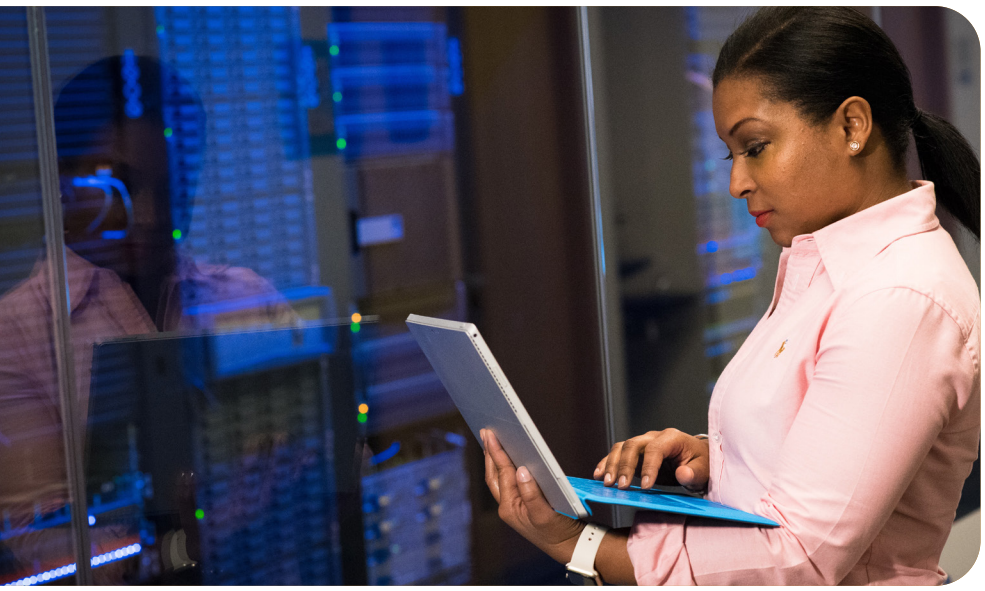
In General terms, compared to competitors in its market category we find that TiDB performs quite well, and is quite competitive

## VITESS

- **Vitess' support for certain setups of databases may require significant changes. (Not a niche issue, this happens with Databases that contain foreign keys.)**
- **Vitess uses lower-level dirty reads protection, while MySQL and TiDB feature higher-level guarantees of the data read. ("read committed" and "repeatable read")**
- **Cross Shard Support is not recommended in Vitess.**
- **Lack of proper support for Kubernetes from the Vitess Operator**

## CRDB

- **Maybe attractive to those that use PostgreSQL databases. Migration to MySQL could be lengthy and costly.**
- **Does not have the ecosystem and supporting tools that TiDB has. (TiDB Data Migration, and TiDB Lightning)**



## PHXSQL

- **Unmaintained, and lacking support.**
- **Can not handle a high amount of queries/writes per second.**

TiDB is a highly competitive options as long as you are in the market for the type of database it is trying to be.

# COMPETITOR SUMMARY

**HEAD TO HEAD**

| | | TIDB | MYSQL | VITESS | CRDB | PHXSQL |
|---|---|---|---|---|---|---|
| **Support** | | High | Low | Medium | Medium | Low |
| **Compatibility** | Overall | High | High | Medium | Medium | Low |
| | MySQL | Yes | Yes | Yes | No | Yes |
| | PostgreSQL | No | No | No | Yes | No |
| | Kubernetes | Yes | Yes | Yes* | Yes | No |
| **Ecosystem** | | ★★★★★ | ★★★★ | ★★★ | ★★ | ★ |
| **Guarantee** | | ★★★★ | ★★★★ | ★★★ | ★★★★★ | ★★★★ |

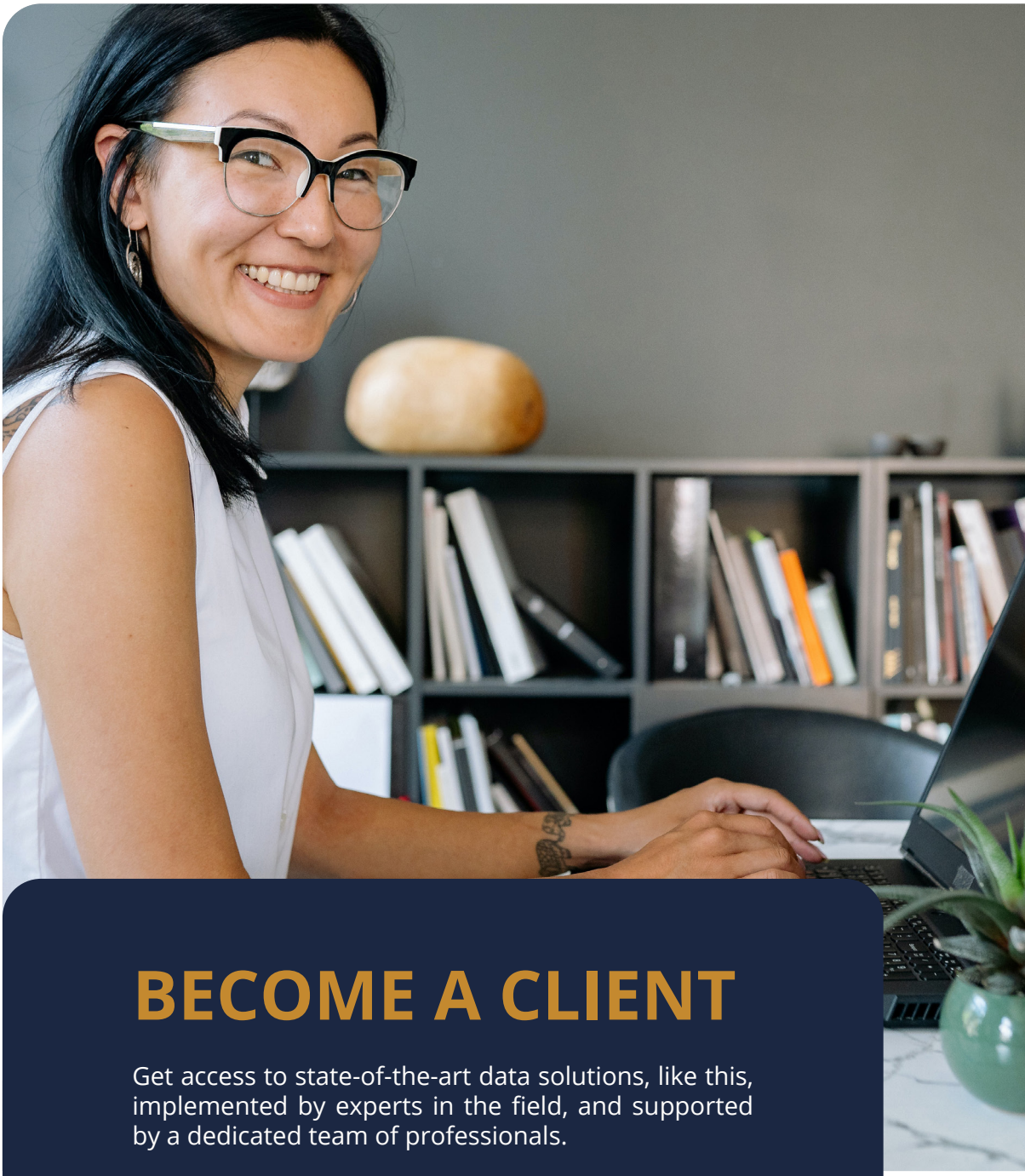
* Not recommended.

## CONCLUSION

If in need of a large scale and high read frequency database solution TiDB offers a very competitive solution, save for very specific cases. These specific cases include if one would want to create such a database as or from a PostgreSQL database. This involves substantial additional costs, or is impossible to set up with TiDB. TiDB also defaults to an isolation level that is not the highest, this instead achieved by CRDB, this is still an option in TiDB but if one were to desire it by default CRDB is also the way to go in this case.

On the contrary it is in the more traditional space of SQL cluster relational databases that TiDB excels at, by offering support in multiple languages, and offering additional tailored tools. This brings it, in this category as a top contender in the space.

ICONIC DATA



# BECOME A CLIENT

Get access to state-of-the-art data solutions, like this, implemented by experts in the field, and supported by a dedicated team of professionals.

To learn more, visit www.iconic.inc