



# Unlocking the Value of AI with Oracle

---

Daniel Morgan  
Technical Director Database and Cloud Security  
Oracle Professional Services

May 8, 2025



# Agenda

Introduction

Foundations of AI at Oracle

Business and Use Cases

AI for DBAs

AI Security

Final Thoughts

Questions

Next Steps



# daniel.d.morgan@oracle.com



- **Oracle** Consulting Services, Technical Director, Database Security



Member, Oracle Security Tiger Team



Oracle ACE Director Alumnum

- Educator



Adjunct Professor, University of Washington, Oracle Program, 1998-2009



Oracle Consultant: Harvard University

- Guest lecturer at universities & colleges in APAC, EMEA, LATAM and NA
  - Frequent conference speaker ... OpenWorld + 178 country visits to 47 countries, since 2008
  - Presented at Argonne National, Lawrence Livermore, Pacific Northwest Labs and NASA (DC)
- IT Professional
    - Primary Author: US DOD, Defense Information Systems Agency, 19c Secure Technical Implementation Guides
    - Member Oracle Database Security Partner Advisory Council 2019-2021
    - Presenter: AskTom.Oracle.com
    - Founder and Principal Contributor @ Morgan's Library ([www.morganslibrary.org](http://www.morganslibrary.org))
    - Founding Chair Washington Software Association's Database Special Interest Group
    - Oracle Database and Database Beta Tester since 1988-9

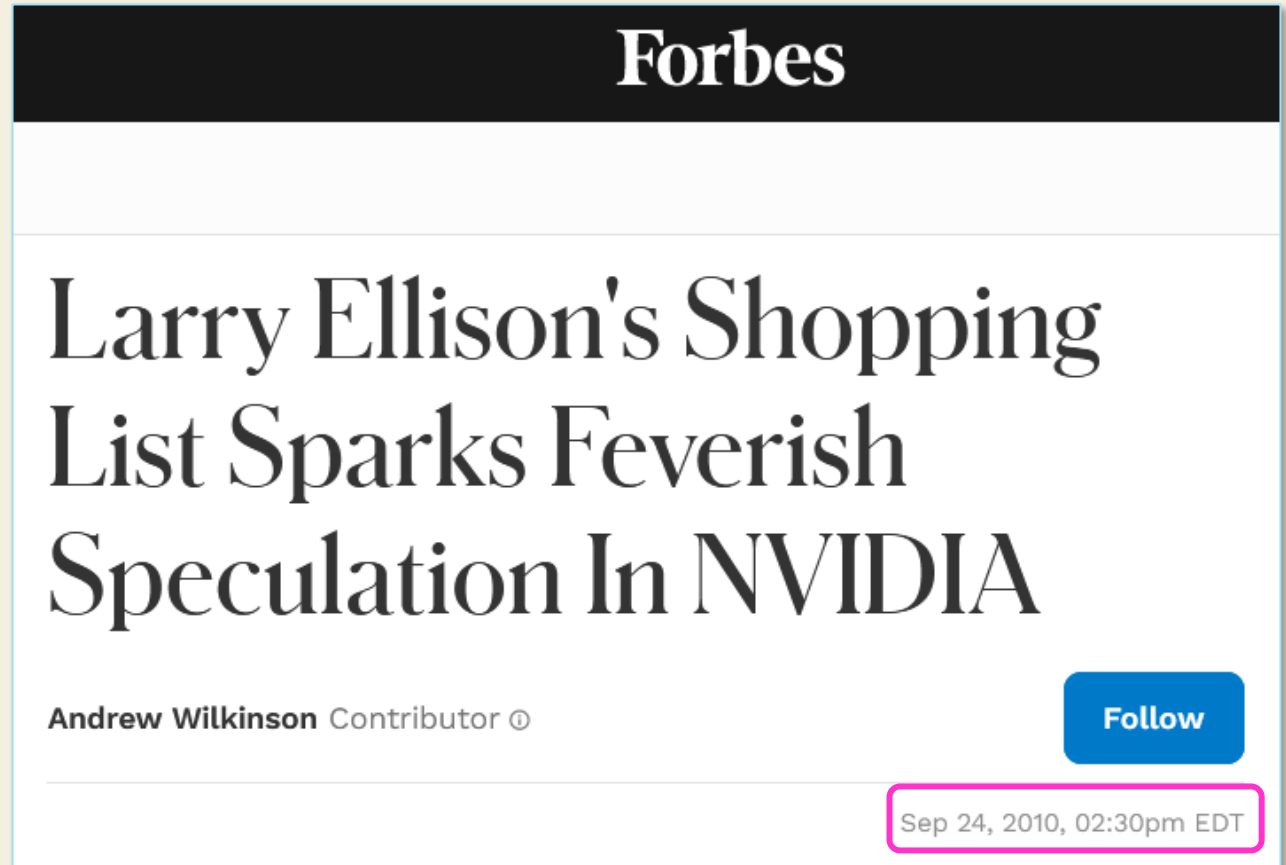
# How Long Has Oracle Been Thinking About AI?

OpenAI was founded in 2015

ChatGPT was launched in November 2022

Gemini (Google AI) March 2023

xAI in March 2023

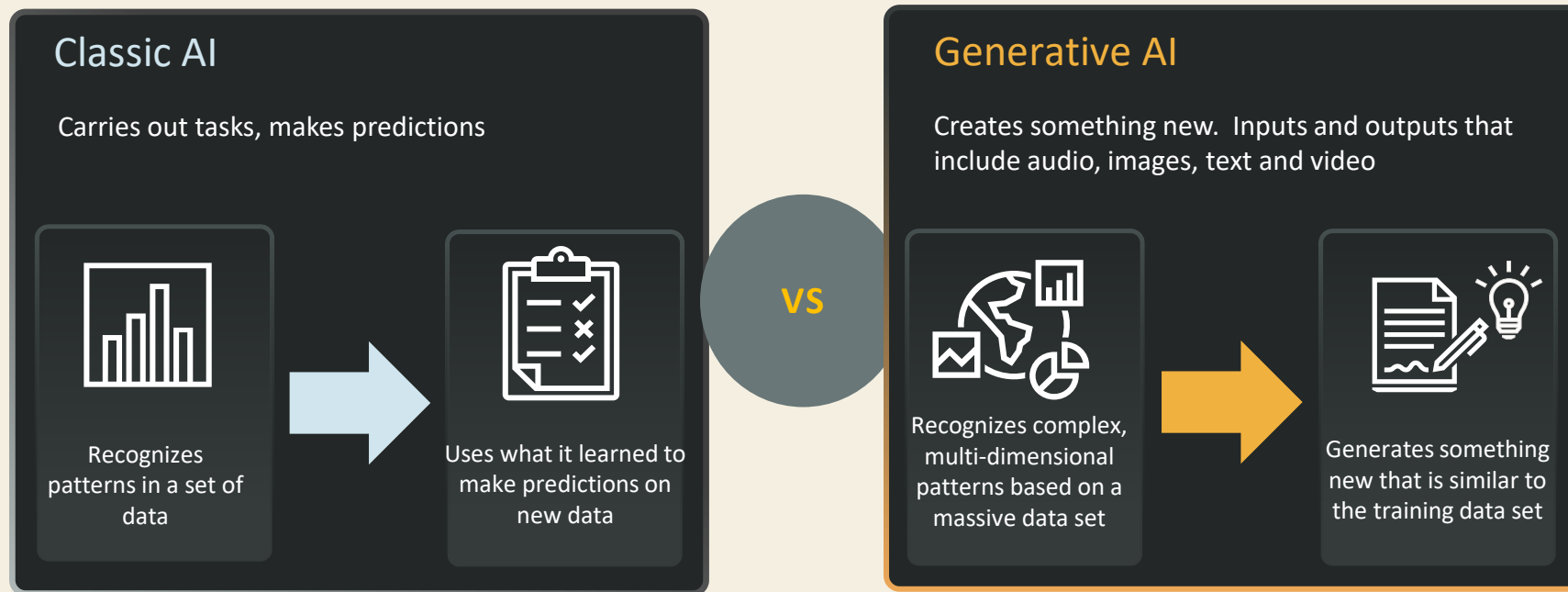


The image shows a screenshot of a Forbes article. At the top, the word "Forbes" is written in white on a black background. Below this, the article title "Larry Ellison's Shopping List Sparks Feverish Speculation In NVIDIA" is displayed in a large, black, serif font. Underneath the title, the author's name "Andrew Wilkinson" is listed as a "Contributor" with a small circular icon. To the right of the author's name is a blue button with the word "Follow" in white. At the bottom right of the article preview, there is a pink-bordered box containing the date and time "Sep 24, 2010, 02:30pm EDT".

# Artificial Intelligence (1:2)

Artificial Intelligence (AI) is newly popular jargon used to label collection of technologies that have existed, in some form, for decades.

What is new, is the vastly improved chip sets, servers, storage, and networks that make possible substantial improvements in scalability and performance.



## Artificial Intelligence (2:2)

Oracle is famously known for Larry Ellison naming the initial release of the Oracle Database "Version 2" because he knew customers prefer Version 2 to Version 1.

For most IT companies Artificial Intelligence is still Version 0.9 and they are scrambling with:

- New technology and architecture
- New code
- New integrations
- Inadequate testing
- Lack of support
- Lack of patches
- How, if they get to Version 1, they will update it
- A profound lack of technology to the IT maturity model
- Make it up as you go compliance with governance, regulatory requirements, and security

The foundations for AI have been in the Oracle Database for more than 20 years.

# The Foundations of AI At Oracle



# 1998: Non-Scalar Data Types

In version 8i, Oracle expanded CREATE TYPE to create non-scalar data types such as an object types named varying array and varray, and nested table type, and complete and incomplete object types.

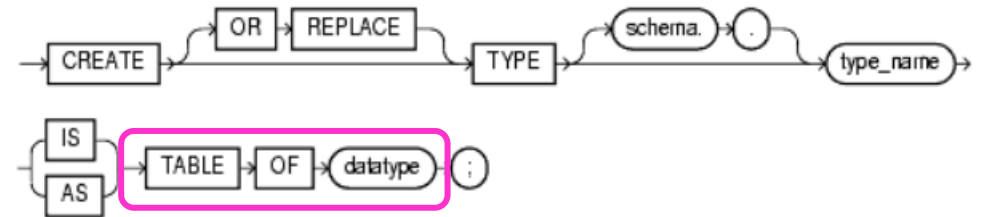
Oracle8i SQL Reference  
Release 3 (8.1.7)  
Part Number A85397-01



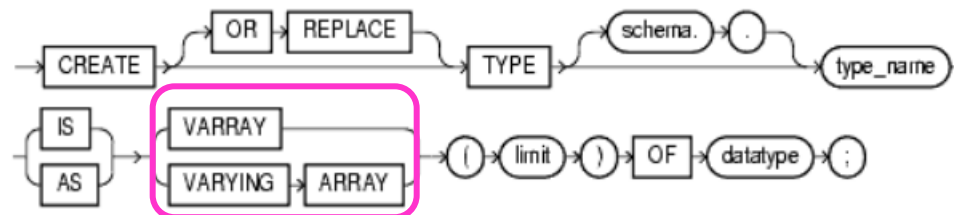
SQL Statements:  
CREATE SYNONYM to DROP ROLLBACK SEGMENT, 7 of 31

## CREATE TYPE

create\_nested\_table\_type::=



create\_varray\_type::=



Sam Altman was 13 years old.

# 1998: Oracle interMedia, Spatial and Visual Image Retrieval (1:2)

Getting to Know Oracle8i  
Release 2 (8.1.6)  
A76962-01



## 3

### Oracle 8i (8.1.5) New Features

This chapter describes new features, options, and enhancements available in Oracle8i release 8.1.5.

This chapter contains the following sections:

- [Content Management for the Internet](#)
- [Oracle interMedia, Spatial, Time Series, and Visual Image Retrieval](#)
- [Java](#)
- [Data Warehousing and Very Large Data Bases \(VLDB\)](#)
- [Database Features](#)

AI technology is based on complex data types like audio, images, text and video and the use of vectors to calculate and store metadata for their retrieval.

# 1998: Oracle interMedia, Spatial and Visual Image Retrieval (2:2)

Oracle Spatial and interMedia were first released in 1998.

The screenshot shows the Oracle Spatial Developer's Guide page for "1 Spatial Concepts". The page title is "Spatial Developer's Guide". The left sidebar contains a table of contents with "1 Spatial Concepts" highlighted. The main content area has the heading "1 Spatial Concepts" and a paragraph: "Oracle Spatial is an integrated set of functions, procedures, data types, and data models that support spatial analytics. The spatial features enable spatial data to be stored, accessed, and analyzed quickly and efficiently in an Oracle database." Below this is another paragraph: "Spatial data represents the essential location characteristics of real or conceptual objects as those objects relate to the real or conceptual space in which they exist." Under "Major topics:", there are two bullet points: "What Is Oracle Spatial?" and "Object-Relational Model". The "Object-Relational Model" bullet point contains the text: "Oracle Spatial supports the **object-relational model for representing geometries. This model stores an entire geometry in the Oracle native spatial data type for vector data, SDO\_GEOMETRY.**" The "On this page" sidebar on the right lists sections 1.1 through 1.5.4.

Not the "new" VECTOR data type ... but we learned how to write DML, DDL, PL/SQL, manage, patch, backup, restore, and replicate vector data.



# 2002: A Rich History Of Relevant Technologies

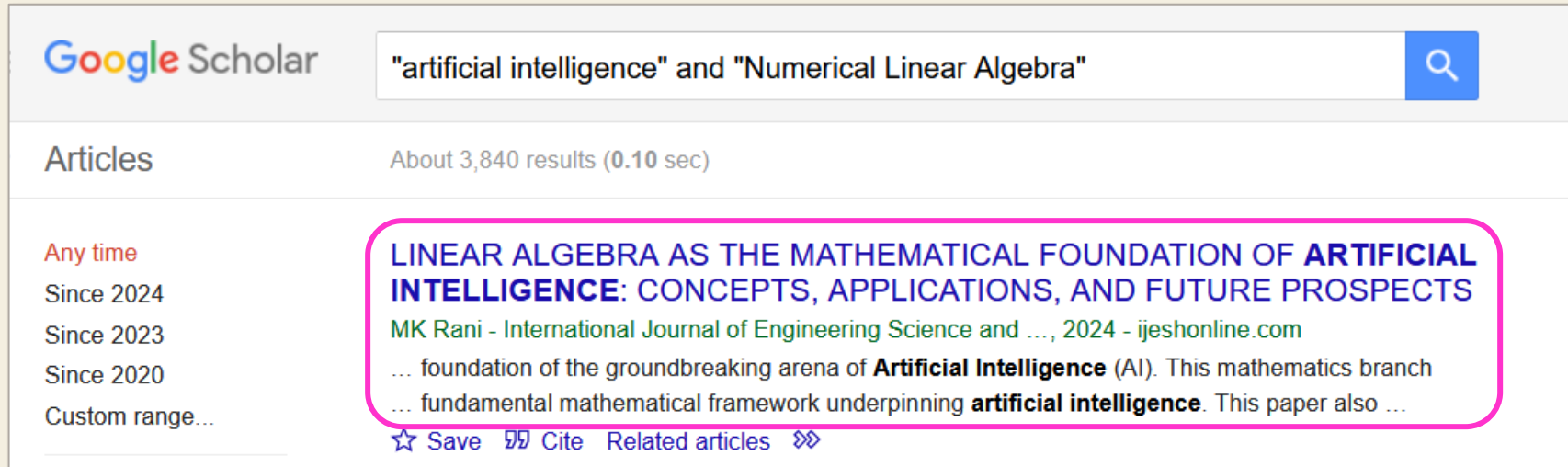
You may think of the Oracle Database as Relational, but that is more a statement about where we came from than it is about how we have evolved since 1978. Oracle version 8.0, released in 1997, had native support for objects with both user-defined data types, inheritance and polymorphism.

Oracle has supported non-relational indexes in schemas such as CTXSYS since version 9i in 2002. DBMS\_VECTOR\_CHAIN in 23ai is owned by CTXSYS: This is not a coincidence.

Sam Altman was 17 years old.



## 2003: Oracle Numerical Linear Algebra (1:2)



The screenshot shows a Google Scholar search interface. The search bar contains the query '"artificial intelligence" and "Numerical Linear Algebra"'. Below the search bar, it indicates 'Articles' and 'About 3,840 results (0.10 sec)'. On the left side, there are filters for 'Any time', 'Since 2024', 'Since 2023', 'Since 2020', and 'Custom range...'. The main search result is highlighted with a pink border and contains the following text:

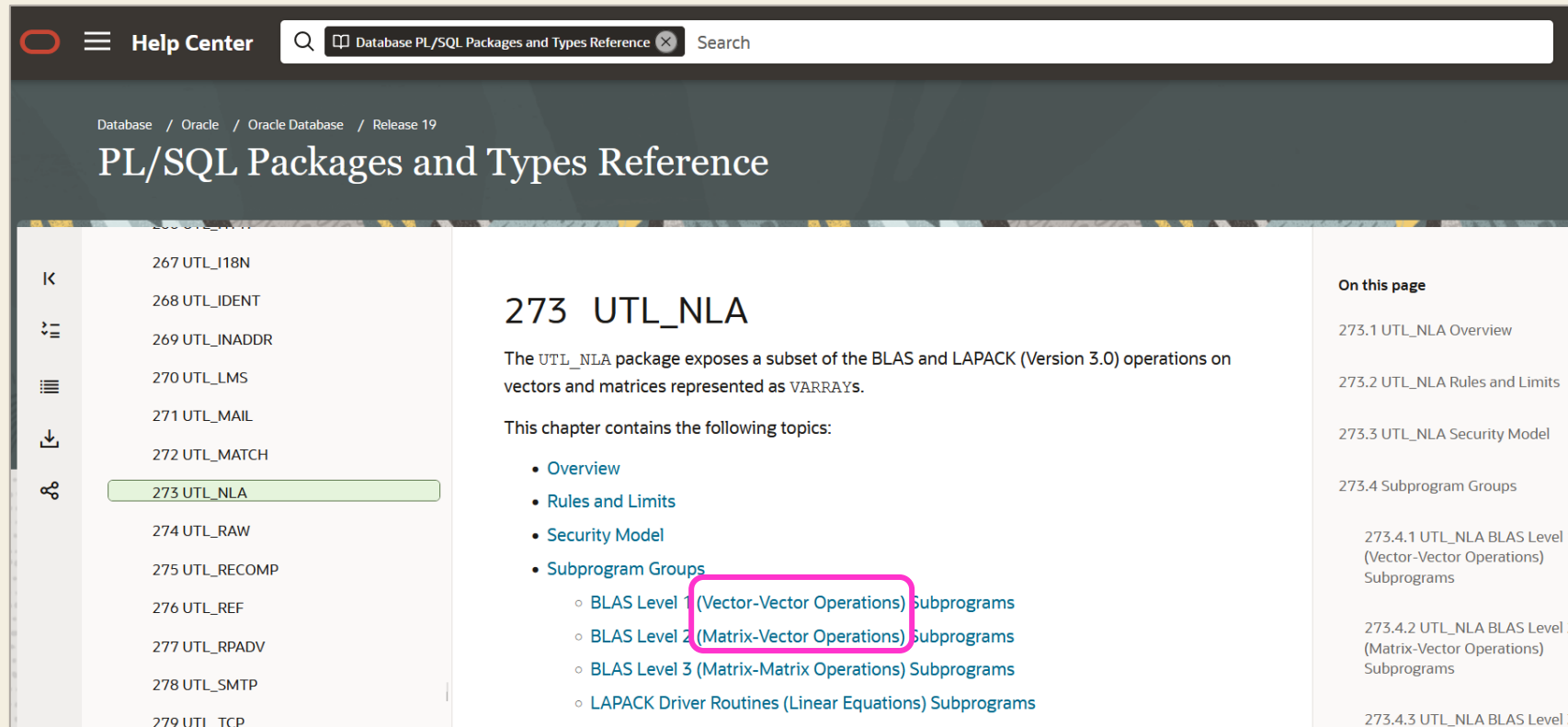
**LINEAR ALGEBRA AS THE MATHEMATICAL FOUNDATION OF ARTIFICIAL INTELLIGENCE: CONCEPTS, APPLICATIONS, AND FUTURE PROSPECTS**  
MK Rani - International Journal of Engineering Science and ..., 2024 - ijeshonline.com  
... foundation of the groundbreaking arena of **Artificial Intelligence** (AI). This mathematics branch  
... fundamental mathematical framework underpinning **artificial intelligence**. This paper also ...

Below the article title, there are icons for 'Save', 'Cite', 'Related articles', and a double arrow icon.

What is the connection between Numerical Linear Algebra and the Oracle Database?

# 2003: Oracle Numerical Linear Algebra (2:2)

The UTL\_NLA built-in PL/SQL package was created in December 2003 and was a component of version 10gR2 released the following year.

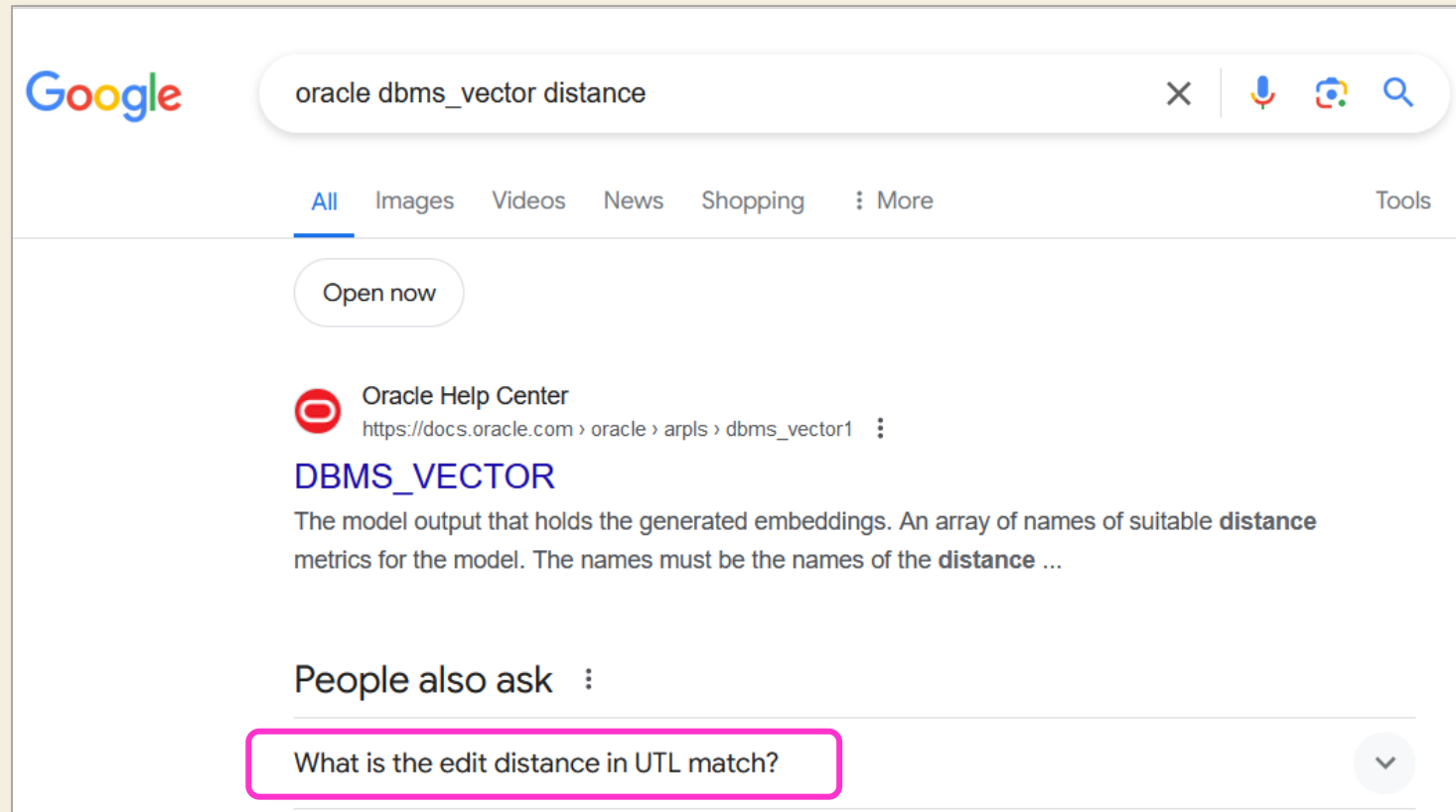


The screenshot shows the Oracle Help Center interface. The search bar at the top contains the text "Database PL/SQL Packages and Types Reference". The breadcrumb trail indicates the path: Database / Oracle / Oracle Database / Release 19. The main heading is "PL/SQL Packages and Types Reference". On the left, a navigation pane lists various packages, with "273 UTL\_NLA" highlighted in green. The main content area displays the title "273 UTL\_NLA" and a description: "The UTL\_NLA package exposes a subset of the BLAS and LAPACK (Version 3.0) operations on vectors and matrices represented as VARRAYs." Below this, it states "This chapter contains the following topics:" followed by a bulleted list: "Overview", "Rules and Limits", "Security Model", and "Subprogram Groups". Under "Subprogram Groups", there are four sub-items: "BLAS Level 1 (Vector-Vector Operations) Subprograms", "BLAS Level 2 (Matrix-Vector Operations) Subprograms", "BLAS Level 3 (Matrix-Matrix Operations) Subprograms", and "LAPACK Driver Routines (Linear Equations) Subprograms". The "BLAS Level 2 (Matrix-Vector Operations) Subprograms" item is circled in pink. On the right side, there is a section titled "On this page" with a list of links: "273.1 UTL\_NLA Overview", "273.2 UTL\_NLA Rules and Limits", "273.3 UTL\_NLA Security Model", "273.4 Subprogram Groups", "273.4.1 UTL\_NLA BLAS Level 1 (Vector-Vector Operations) Subprograms", "273.4.2 UTL\_NLA BLAS Level 2 (Matrix-Vector Operations) Subprograms", and "273.4.3 UTL\_NLA BLAS Level 3".

Numerical Linear Algebra has been available in the Oracle Database for 22 years.



# 2004: Oracle Distance Algorithms (1:2)

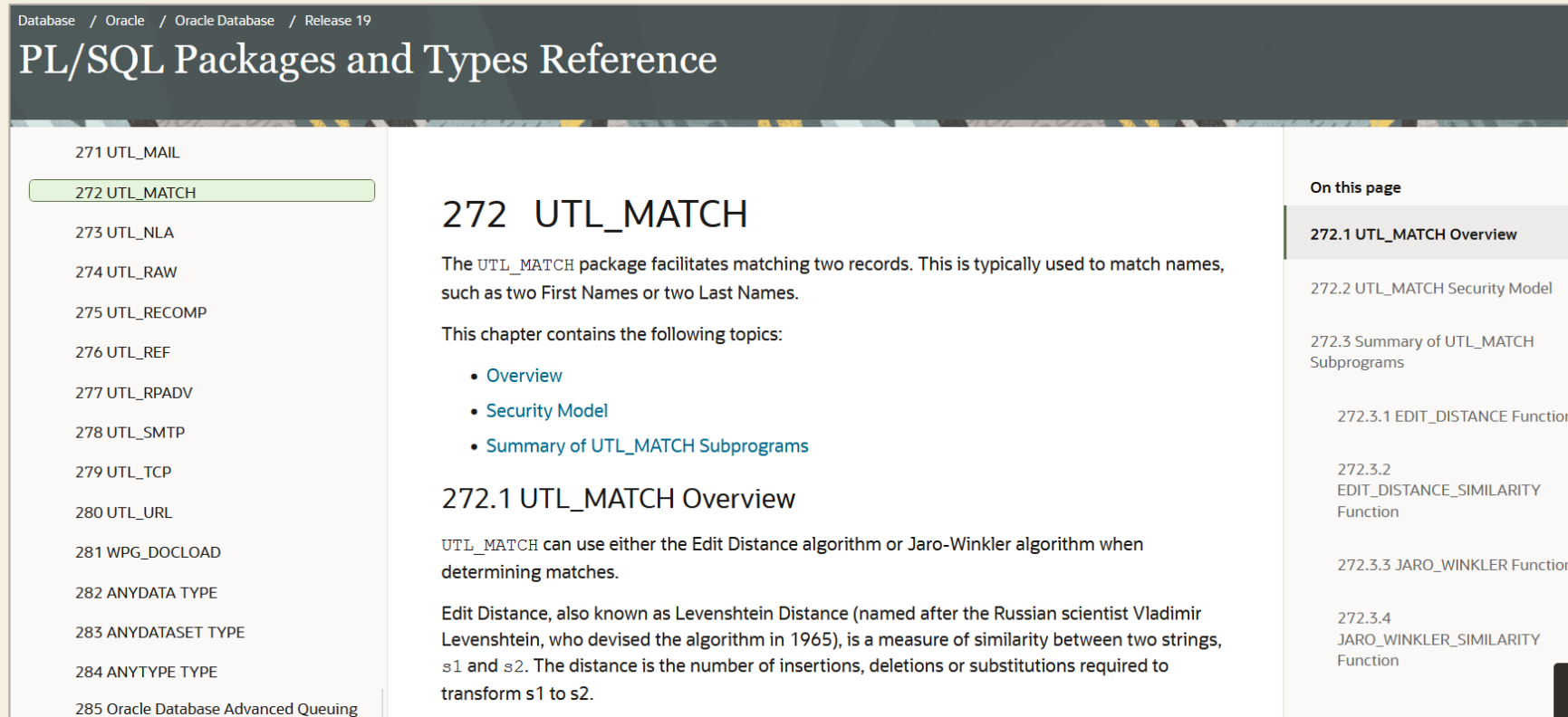


Why is the AI in a web search connecting the DBMS\_VECTOR AI package with something named UTL\_MATCH?



# 2004: Oracle Distance Algorithms (2:2)

The UTL\_MATCH built-in PL/SQL package was created in March 2004 and included in version 10gR2.



The screenshot shows the Oracle PL/SQL Packages and Types Reference page for UTL\_MATCH. The page is titled "PL/SQL Packages and Types Reference" and is for Oracle Database Release 19. The left sidebar lists various packages, with "272 UTL\_MATCH" highlighted. The main content area is titled "272 UTL\_MATCH" and contains the following text:

The `UTL_MATCH` package facilitates matching two records. This is typically used to match names, such as two First Names or two Last Names.

This chapter contains the following topics:

- [Overview](#)
- [Security Model](#)
- [Summary of UTL\\_MATCH Subprograms](#)

**272.1 UTL\_MATCH Overview**

`UTL_MATCH` can use either the Edit Distance algorithm or Jaro-Winkler algorithm when determining matches.

Edit Distance, also known as Levenshtein Distance (named after the Russian scientist Vladimir Levenshtein, who devised the algorithm in 1965), is a measure of similarity between two strings, `s1` and `s2`. The distance is the number of insertions, deletions or substitutions required to transform `s1` to `s2`.

The right sidebar, titled "On this page", lists the following sections:

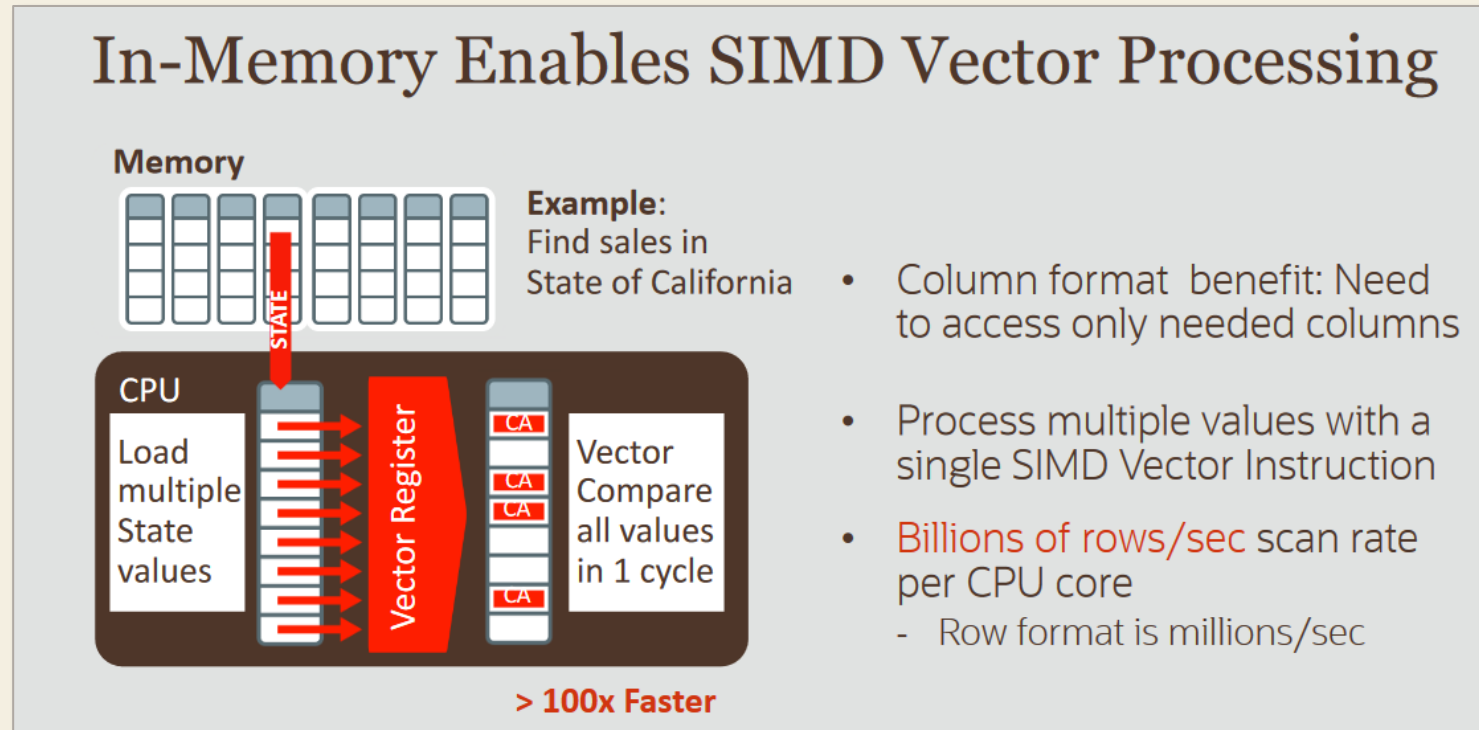
- 272.1 UTL\_MATCH Overview
- 272.2 UTL\_MATCH Security Model
- 272.3 Summary of UTL\_MATCH Subprograms
  - 272.3.1 EDIT\_DISTANCE Function
  - 272.3.2 EDIT\_DISTANCE\_SIMILARITY Function
  - 272.3.3 JARO\_WINKLER Function
  - 272.3.4 JARO\_WINKLER\_SIMILARITY Function

Technology that measures the "distance" between words and phrases.



# 2017: In-Memory Database

The incredible speed of AI in the Oracle Database is built upon a foundation of technologies like In-Memory Column Store (2017), In-Memory Expressions, Columnar Search, Columnar Compression, OSON (Oracle JSON), and SIMD\* Vector Processing.



\* SIMD is short for "Single Instruction Multiple Data" parallel processing.



# A Rich History Of Relevant Technologies

Others are scrambling to release new technology hoping that they can ride the wave ....



At Oracle we have been perfecting, deploying and managing the technologies that are the foundations of AI to hundreds-of-thousands of our customers: Backed them up with RMAN, exported them with Data Pump, made them Highly Available with RAC and Data Guard, sharded, partitioned, secured, queried, updated, patched, migrated, encrypted and compressed them with so little drama you may not have noticed.

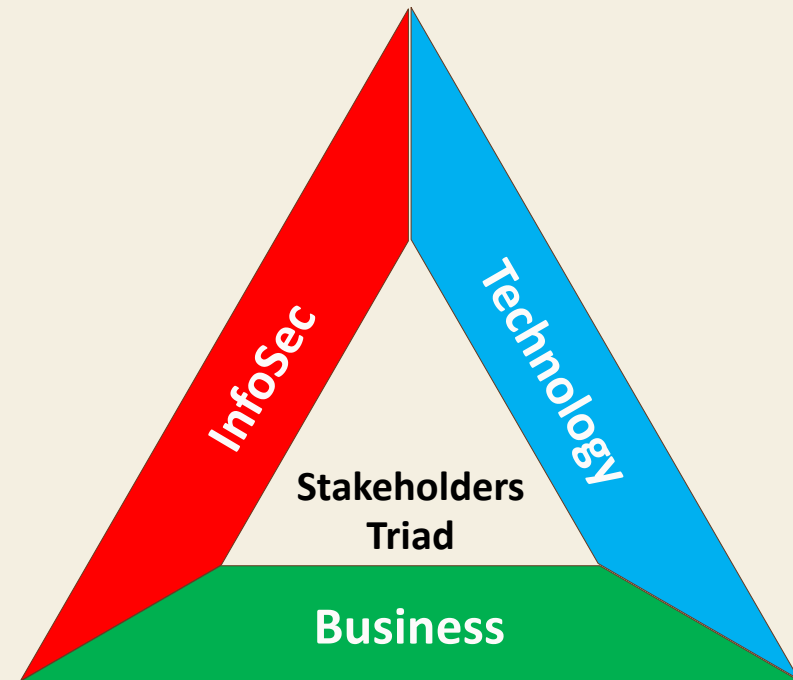
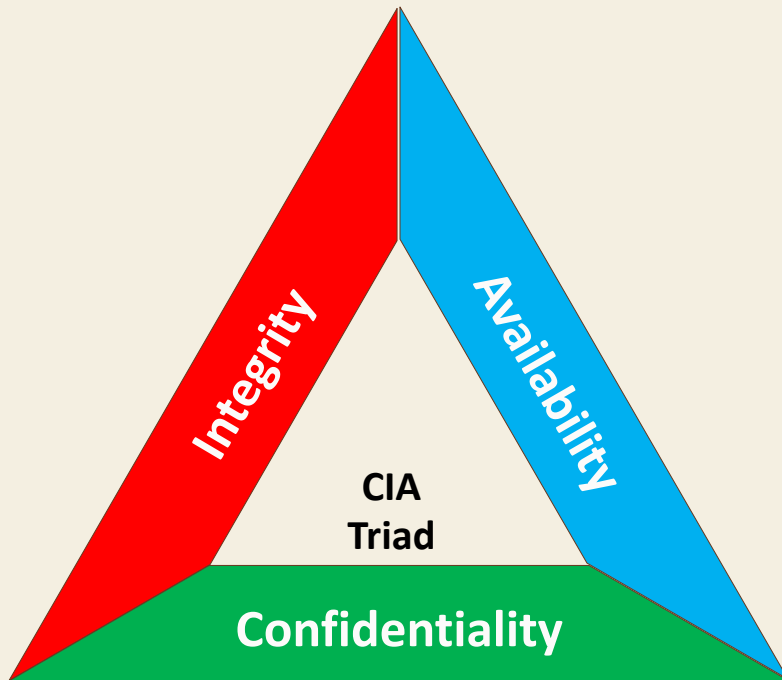
For 25+ years, our lack of drama is your assurance of maximum value, minimum risk.

# Business Case



# New Technology Is Cool, But ...

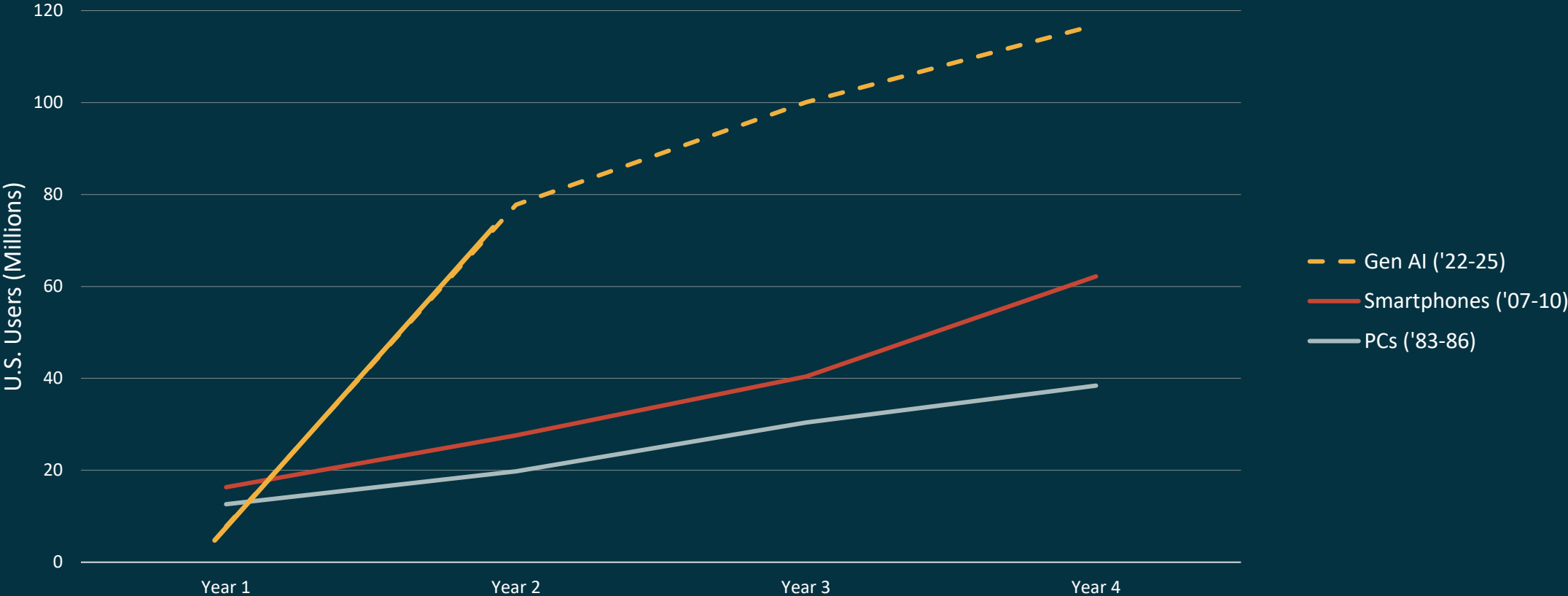
You have a business to run. If you are going to implement this new technology, you must address the real-world questions that will be asked by your Board, your C-Level, your employees, your customers, and your auditors and regulators.



# Its Great That AI Is Growing Rapidly



### Gen AI's Initial Adoption Curve vs Other Recent Technologies



Sources: Insider Intelligence and ITU



# It Is Great That Substantial Investments Are Being Made

**\$276B**

Global investment  
in AI in 2021

**37.3%**

AI's predicted CAGR  
from 2023 to 2030

**\$15.7T**

AI's expected impact on  
global GDP by 2030

Sources: PWC, Grand View Research

## But What Does That Mean To Your Organization Specifically?

You've heard that a substantial investment in infrastructure is required.

You've heard that the AI must be trained, who is going to do the training?

You have legitimate concerns about the risks of implementing another new technology.

What about hallucinations?

To discuss AI technology with your management you need to understand all of the issues related to integrating it your existing tech stack?

Your CEO, COO, CFO, and CISO are all going to have questions: You will need to have answers.

# Investment

The infrastructure demands of AI are computationally expensive and intensive and benefit from new server technologies such as NVIDIA A100 GPU's.

ORACLE

## Train ML models with up to 40% cost savings\*

Oracle Cloud Infrastructure (OCI) and NVIDIA have partnered to deliver infrastructure which can accelerate AI-workloads and drive widespread adoption of deep learning models.

Nodes powered by NVIDIA A100 GPUs and OCI cluster networking with less than 2 microseconds latency

1,600 Gb/sec

1,600 Gb/sec

1,600 Gb/sec

1,600 Gb/sec

The advertisement features a green background with white text. At the top left is the Oracle logo. The main headline is 'Train ML models with up to 40% cost savings\*'. Below this is a paragraph describing the partnership between Oracle Cloud Infrastructure (OCI) and NVIDIA. The central part of the ad shows a laptop screen displaying a diagram of a cluster of four server nodes. Each node is represented by a green square with a grid pattern and a vertical bar in the center, with '1,600 Gb/sec' written below it. Brackets group the nodes, and text indicates they are powered by NVIDIA A100 GPUs and OCI cluster networking with less than 2 microseconds latency.

Partner with Oracle and you can leverage our investment in OCI Cloud Infrastructure.



# Training Data

The most valuable data for Training your AI and maximizing the value of your investment is not petabytes of posts on Social Media or the contents of the Library of Congress: It's the data you already have in your ERP and other internal database systems: Proprietary data that is not available to your competitors and is specific to you and your organization.

To use most AI systems requires you to replicate your data to new systems: This will require:

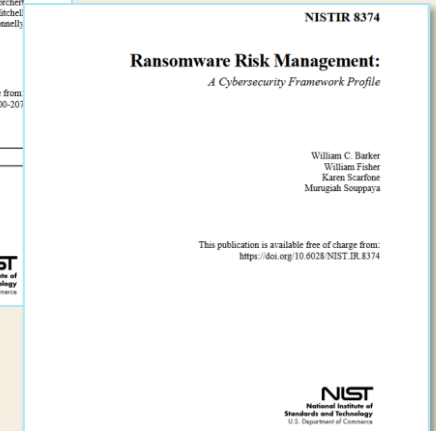
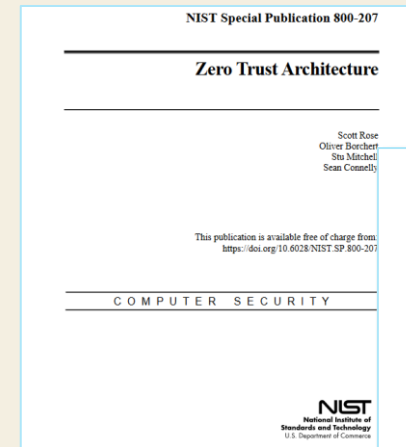
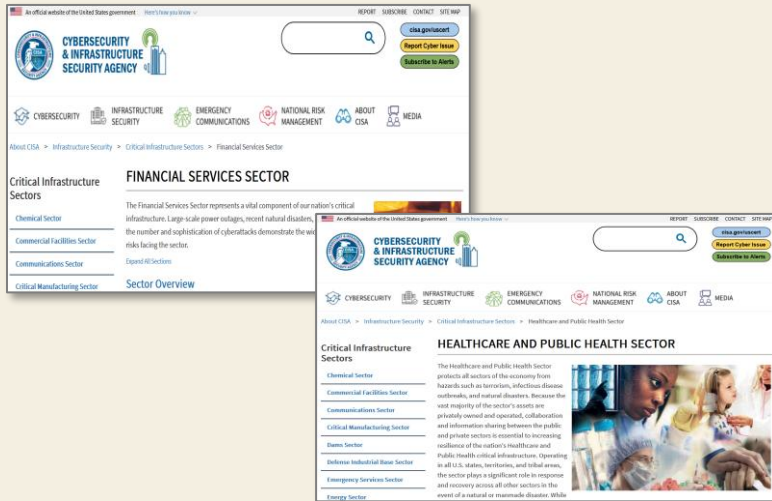
- Additional Data Center Space, Infrastructure, Insurance, and Personnel (FTEs)
- Auditing for governance and regulatory compliance

By expanding your footprint, this increases the attack surface and puts your valuable data at additional risk of exfiltration and exploitation.

Oracle Database allows you to bring AI to your current databases, with all of your current ability to audit, compress, encrypt, partition, and secure it: No data export or migration is required.

# Risk Mitigation

Your data, hosted in your current database systems or processed for use in AI is subject to the exact same governance and regulatory requirements. You don't get a pass if PII, PHI, PCI-DSS, GDPR, SOX, HIPAA, ITAR, DFARS, EAR, and other sensitive data are in an AI rather than in a relational database.



By bringing AI to your data, where it already passes audit, governance and regulatory requirements, you know you have full support for data sovereignty and non-repudiation and are mitigating both known and unknown risks.



# New Technologies

The term "AI" is new and cutting edge ...

But the foundational technologies that make AI work have been in the Oracle Database for more than two decades.

Oracle not only has the depth of experience working with vectors, with numerical linear algebra, partitioning, sharding, objects, non-scalar data types and in-memory ...

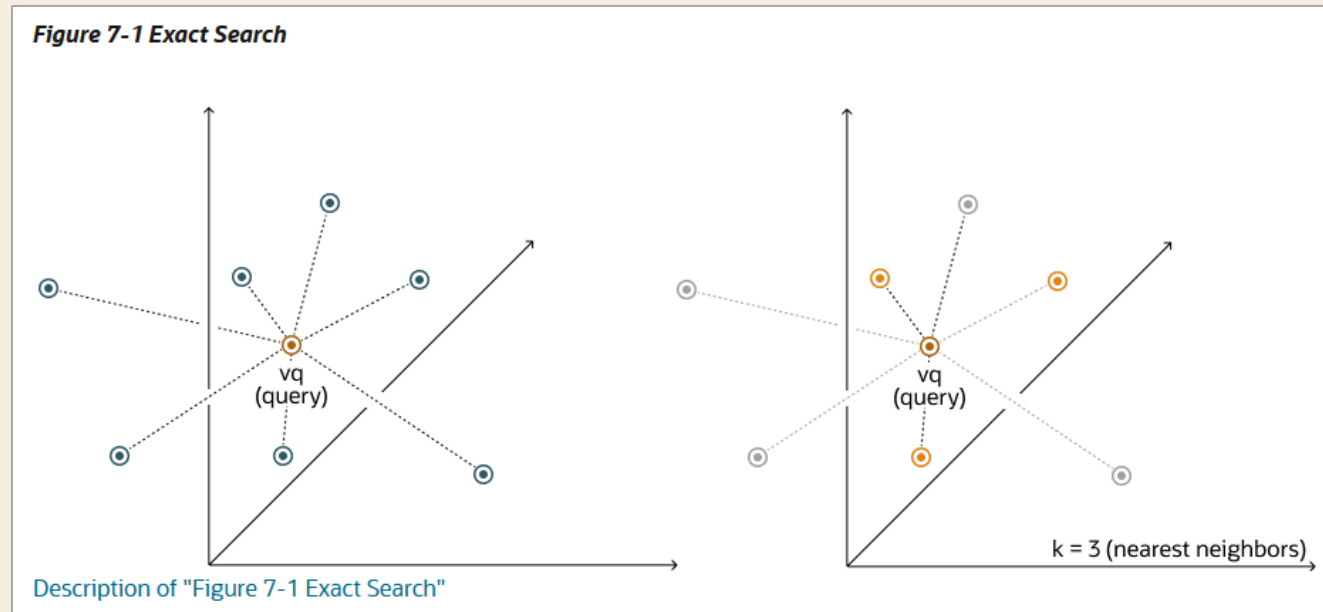
Oracle has decades of experiencing backing it up, securing it, patching it, upgrading it, migrating it and supporting it through the IT lifecycle.

Just because AI is "new" doesn't mean IT maturity models can be ignored.

# Hallucinations (1:2)

The popular media says: "Hallucinations are incorrect or misleading results that AI models generate."

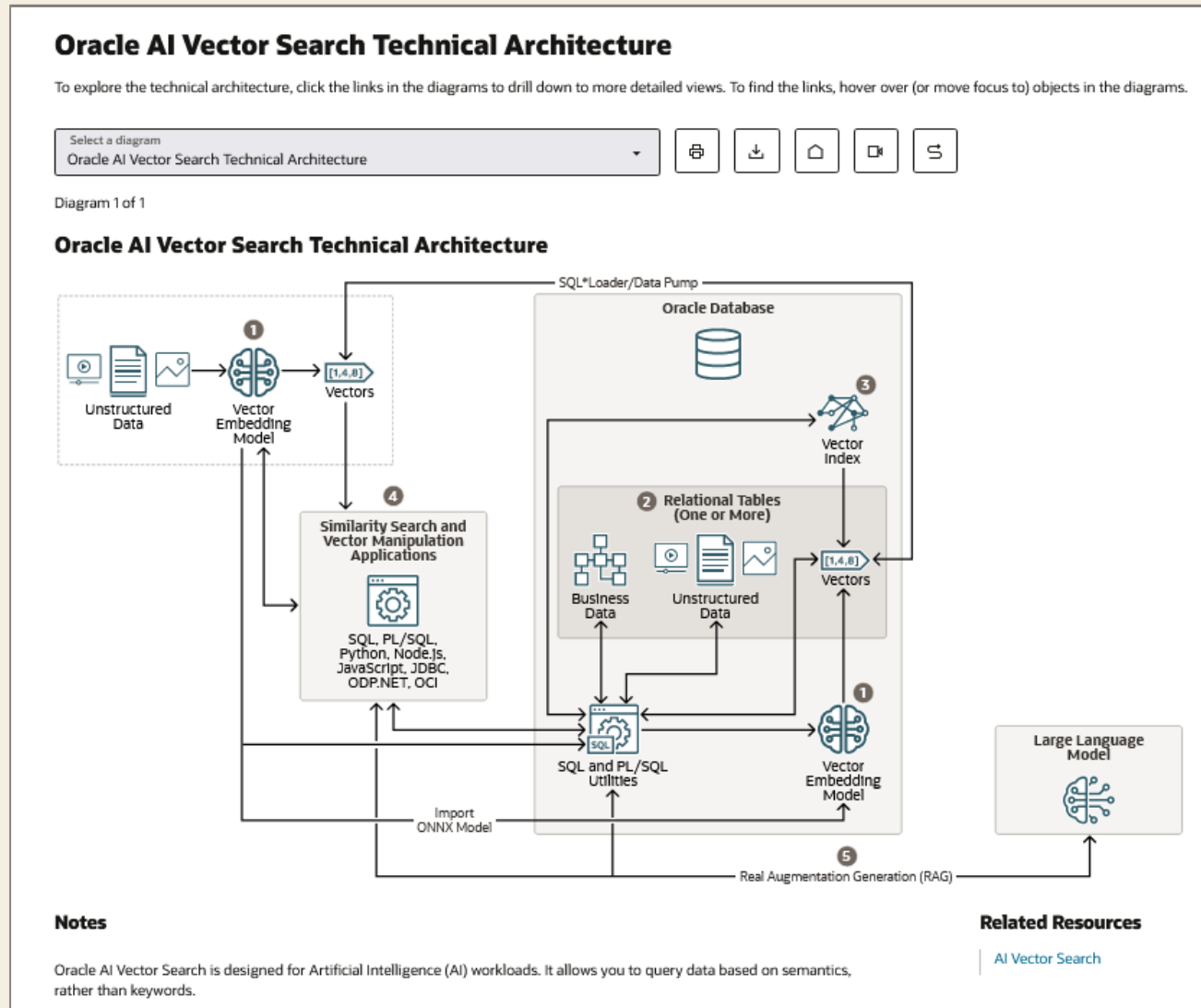
The accuracy of AI queries in Oracle is under the control of developers who have simultaneous access to structured relational data and unstructured CLOBS, BLOBS, JSON, SODA and XML.



[https://docs.oracle.com/en/database/oracle/oracle-database/23/vsiad/aivs\\_genarch.html](https://docs.oracle.com/en/database/oracle/oracle-database/23/vsiad/aivs_genarch.html)



# Hallucinations (2:2)



# AI for DBAs



## Data Types and DDL

The core AI technology in the Oracle Database is "Vector Search" which leverages the core technologies you already know and the skills and experience you already have.

Tablespaces are tablespaces, users are users, and creating tables for AI is in your wheelhouse.

```
oracle@asra23c:~  
SQL> CREATE TABLE ai_employee (  
2  employee_id    INTEGER,  
3  first_name     VARCHAR2(30),  
4  last_name      VARCHAR2(30),  
5  date_hired     DATE,  
6  resume         BLOB,  
7  id_photo       BLOB,  
8  vector_data    VECTOR)  
9  TABLESPACE users;  
  
Table created.
```



# Configuration Parameters

You already know how to query the data dictionary and make configuration changes.

```
oracle@asra23c:~  
SQL> SELECT name, value, isdefault, isses_modifiable, issys_modifiable, isinstance_modifiable  
2 FROM v$parameter  
3 WHERE name LIKE '%vector%'  
4 ORDER BY 1;
```

NAME	VALUE	ISDEFAULT	ISSES	ISSYS_MOD	ISINS
inmemory_deep_vectorization	TRUE	TRUE	TRUE	IMMEDIATE	TRUE
spatial_vector_acceleration	TRUE	FALSE	TRUE	IMMEDIATE	TRUE
vector_index_neighbor_graph_reload	OFF	TRUE	TRUE	IMMEDIATE	TRUE
vector_memory_size	0	TRUE	FALSE	IMMEDIATE	TRUE
vector_query_capture	ON	TRUE	TRUE	IMMEDIATE	TRUE

Your existing skills are immediately applicable to an Oracle Database with Vector Search.

```
oracle@asra23c:~  
SQL> ALTER SYSTEM SET vector_memory_size = 1024 SCOPE = SPFILE;  
  
System altered.
```



# Data Dictionary Objects

The Data Dictionary tables and views follow naming and querying rules you already know.

```
oracle@asra23c:~  
SQL> SELECT DISTINCT owner, object_name, object_type  
FROM dba_objects  
WHERE (object_name LIKE '%SEARCH%' OR object_name LIKE '%VECTOR%')  
AND object_name LIKE 'DBA%'  
ORDER BY 1,2 2 3;
```

OWNER	OBJECT_NAME	OBJECT_TYPE
SYS	DBA_VECTOR_INDEX_ACCURACY_REPORT	VIEW
PUBLIC	DBA_VECTOR_INDEX_ACCURACY_REPORT	SYNONYM
SYS	DBA_VECTOR_INDEX_ACCURACY_TASKS	VIEW
PUBLIC	DBA_VECTOR_INDEX_ACCURACY_TASKS	SYNONYM
CTXSYS	DBA_VECTOR_VOCAB	VIEW
PUBLIC	DBA_VECTOR_VOCAB	SYNONYM
CTXSYS	DBA_VECTOR_VOCAB_TOKENS	VIEW
PUBLIC	DBA_VECTOR_VOCAB_TOKENS	SYNONYM
CTXSYS	DBA_VECTOR_LANG	VIEW
PUBLIC	DBA_VECTOR_LANG	SYNONYM
CTXSYS	DBA_VECTOR_ABBREV_TOKENS	VIEW
PUBLIC	DBA_VECTOR_ABBREV_TOKENS	SYNONYM



# Functions

You've worked for years with functions like TO\_CHAR, TO\_CLOB, TO\_DATE, TO\_NUMBER ....

```
-- constructors
FUNCTION TO_VECTOR(vector_str VARCHAR2 CHARACTER SET any_cs, dimensions NATURALN, format NATURALN) RETURN vector;
  pragma BUILTIN('TO_VECTOR', 7, 12);

FUNCTION TO_VECTOR(vector_str CLOB CHARACTER SET ANY_CS, dimensions NATURALN, format NATURALN) RETURN vector;
  pragma BUILTIN('TO_VECTOR', 7, 13);

FUNCTION TO_VECTOR(vector_str BLOB, dimensions NATURALN, format NATURALN) RETURN vector;
  pragma BUILTIN('TO_VECTOR', 7, 18);

-- serializers
FUNCTION FROM_VECTOR(left vector) RETURN varchar2;
  pragma BUILTIN('FROM_VECTOR', 7, 14);

FUNCTION VECTOR_SERIALIZE(left vector) RETURN varchar2;
  pragma BUILTIN('VECTOR_SERIALIZE', 7, 14);
```

The new, TO\_VECTOR function, does precisely what you expect.

Read the docs, but don't expect a lot of surprises.

# Data Quality Operators: FUZZY\_MATCH

FUZZY\_MATCH uses a number of different algorithms to provide a metadata measurement of the "distance" between two different strings.

FUZZY\_MATCH evaluates strings character-by-character whereas UTL\_MATCH evaluates byte-by-byte: The difference is FUZZY\_MATCH handles multi-by character sets and UTL\_MATCH only works with single-byte characters.

PHRASE1	PHRASE2	BG	DLEV	JW	UMJW	UMJWS	LEV	LCS	TRI	WWM
The Quick Brown Fox	The quick brown fox	12	3	.93	.9368	93	3	5	8	1
The Quick Brown Fox	The Quick Brown Fox	18	0	1.00	1.0000	100	0	19	17	4
The quick grey wolf	The quick timber wolf	13	5	.92	.9291	92	5	10	11	3
The glacially slow turtle	the glacially slow turtle	23	1	.97	.9733	97	1	24	22	3
Espresso	Four score and seven years ago	1	25	.46	.4639	46	25	2	0	0



# Data Quality Operators: PHONIC\_ENCODE

PHONIC\_ENCODE takes the algorithm to be used as the first argument, the string to be processed as the second argument, and an optional max\_code\_len argument that controls the length of the desired output which must be an integer between 1 and 12 and returns an encoded representation of the word or phrase.

PHRASE	C1	C2	MCL1	MCL4	MCL8	MCL12
The quick brown fox	OKKP	TKKP	T	TKKP	TKKPRNFK	TKKPRNFKS
The Quick Brown Fox	OKKP	TKKP	T	TKKP	TKKPRNFK	TKKPRNFKS
The quick timber wolf	OKKK	TKLX	T	TKKT	TKKTMPRL	TKKMPRLF
the glacially slow turtle	OKLS	TKLX	T	TKLX	TKLXLSLT	TKLXLSLTRTL
Four score and seven years ago	AKSP	FRSK	F	FRSK	FRSKRNTS	FRSKRNTSFNRS



# Built-in Packages: DBMS\_SEARCH

Operators

- K
- 2 Oracle Text Indexing Elements
- 3 Oracle Text CONTAINS Query Operators
- 4 Special Characters in Oracle Text Queries
- 5 CTX\_ADM Package
- 6 CTX\_ANL Package
- 7 CTX\_CLS Package
- 8 CTX\_DDL Package
- 9 CTX\_DOC Package
- 10 CTX\_ENTITY Package
- 11 CTX\_OUTPUT Package
- 12 CTX\_QUERY Package
- 13 CTX\_REPORT Package
- 14 CTX\_THES Package
- 15 CTX\_ULEXER Package
- 16 DBMS\_SEARCH Package
- 17 Oracle Text Utilities
- 18 Oracle Text Alternative Spelling
- A Oracle Text Result Tables
- B Oracle Text Supported Document Formats
- C Text Loading Examples for Oracle Text
- D Oracle Text Multilingual Features
- E The Oracle Text Scoring Algorithm
- F Oracle Text Views
- G Stopword Transformations in Oracle Text

## 16 DBMS\_SEARCH Package

The `DBMS_SEARCH` PL/SQL package provides procedures and functions to create, manage, and query search indexes for a textual and range-based ubiquitous database search.

Name	Description
<a href="#">CREATE_INDEX</a>	Creates a ubiquitous search index. You can add a set of tables and views as data sources to this index.
<a href="#">ADD_SOURCE</a>	Adds a table or view to the index as a data source.
<a href="#">REMOVE_SOURCE</a>	Removes a table or view and all its associated data from the index.
<a href="#">DROP_INDEX</a>	Removes a search index and all its associated data from the database.
<a href="#">GET_DOCUMENT</a>	Returns a virtual indexed JSON document for the specified source metadata.
<a href="#">FIND</a>	Retrieves a hitlist, and facets an aggregations of JSON documents based on the specified filter conditions.

### 16.1 CREATE\_INDEX

The `DBMS_SEARCH.CREATE_INDEX` procedure creates a ubiquitous search index for a full-text and range-based search across multiple schema objects.

**Notes**

- When run, the `DBMS_SEARCH.CREATE_INDEX` procedure creates a JSON search index with predefined set of preferences and settings, enabled for performing full text search on multiple columns, tables, and views. An index table named `INDEX_NAME` is created with `DATA` and `METADATA` columns. This table is partitioned by `OWNER` and `SOURCE`, where `OWNER` specifies the table owner name and the `SOURCE` specifies the table or view name from which the data is indexed.
- You can define which tables or views should be indexed by adding them as data sources into your index. All the columns of those tables or views are indexed. Use the `DBMS_SEARCH.ADD_SOURCE` and `DBMS_SEARCH.REMOVE_SOURCE` procedures to manage data sources.
- The `DBMS_SEARCH` index is created with the following default indexing preferences:

Preference	Description
<code>BASIC_WORDLIST</code>	Enables wildcard indexing for a fast wildcard search.
<code>SEARCH_ON</code>	Allows both the full-text and range-search queries for a specific data type. The supported data types are <code>NUMBER</code> (for indexing numeric values) and <code>TIMESTAMP</code> (for indexing date-time values).
<code>SYNC</code> and <code>OPTIMIZE</code>	Automatically synchronizes and optimizes the <code>DBMS_SEARCH</code> index in the background at predefined intervals. You do not need to run the <code>SYNC_INDEX</code> and <code>OPTIMIZE_INDEX</code> operations on this index.



# Built-in Packages: DBMS\_VECTOR

The screenshot shows the Oracle Help Center interface. At the top, there is a navigation bar with the Oracle logo, a hamburger menu, and the text "Help Center". A search bar contains the text "Database PL/SQL Packages and Types Reference" and a search icon. Below the navigation bar, the breadcrumb path is "Database / Oracle / Oracle Database / Release 23". The main heading is "PL/SQL Packages and Types Reference". On the left side, there is a vertical list of package names, with "205 DBMS\_VECTOR" highlighted in green. The main content area displays the title "205 DBMS\_VECTOR" and a description: "The DBMS\_VECTOR package simplifies common operations with Oracle AI Vector Search, such as chunking and embedding data, generating text for prompts, or creating vector indexes." It also states that "These functions accept their respective input parameters in JSON format." Below this, there is a "Related Topics" section with a link to "Oracle Database AI Vector Search User's Guide". The next section is "Summary of DBMS\_VECTOR Subprograms", which includes a table listing subprograms and their descriptions. The table is titled "Table 205-1 DBMS\_VECTOR Package Subprograms" and has two columns: "Subprogram" and "Description".

Database / Oracle / Oracle Database / Release 23

## PL/SQL Packages and Types Reference

- 194 DBMS\_TRANSACTION
- 195 DBMS\_TRANSFORM
- 196 DBMS\_TSDP\_MANAGE
- 197 DBMS\_TSDP\_PROTECT
- 198 DBMS\_TTS
- 199 DBMS\_TYPES
- 200 DBMS\_UMF
- ▶ 201 DBMS\_UNDO\_ADV
- 202 DBMS\_USER\_CERTS
- 203 DBMS\_USERDIAG
- 204 DBMS\_UTILITY
- 205 DBMS\_VECTOR**
- 206 DBMS\_VECTOR\_CHAIN
- 207 DBMS\_WARNING
- 208 DBMS\_WM
- 209 DBMS\_WORKLOAD\_CAPTURE
- 210 DBMS\_WORKLOAD\_REPLAY
- 211 DBMS\_WORKLOAD\_REPOSITORY
- 212 DBMS\_XDB
- 213 DBMS\_XA
- 214 DBMS\_XDB\_ADMIN
- 215 DBMS\_XDB\_CONFIG
- 216 DBMS\_XDB\_CONSTANTS

### 205 DBMS\_VECTOR

The `DBMS_VECTOR` package simplifies common operations with Oracle AI Vector Search, such as chunking and embedding data, generating text for prompts, or creating vector indexes.

These functions accept their respective input parameters in JSON format.

#### Related Topics

- [Oracle Database AI Vector Search User's Guide](#)

#### Summary of DBMS\_VECTOR Subprograms

This table lists the `DBMS_VECTOR` subprograms and briefly describes them.

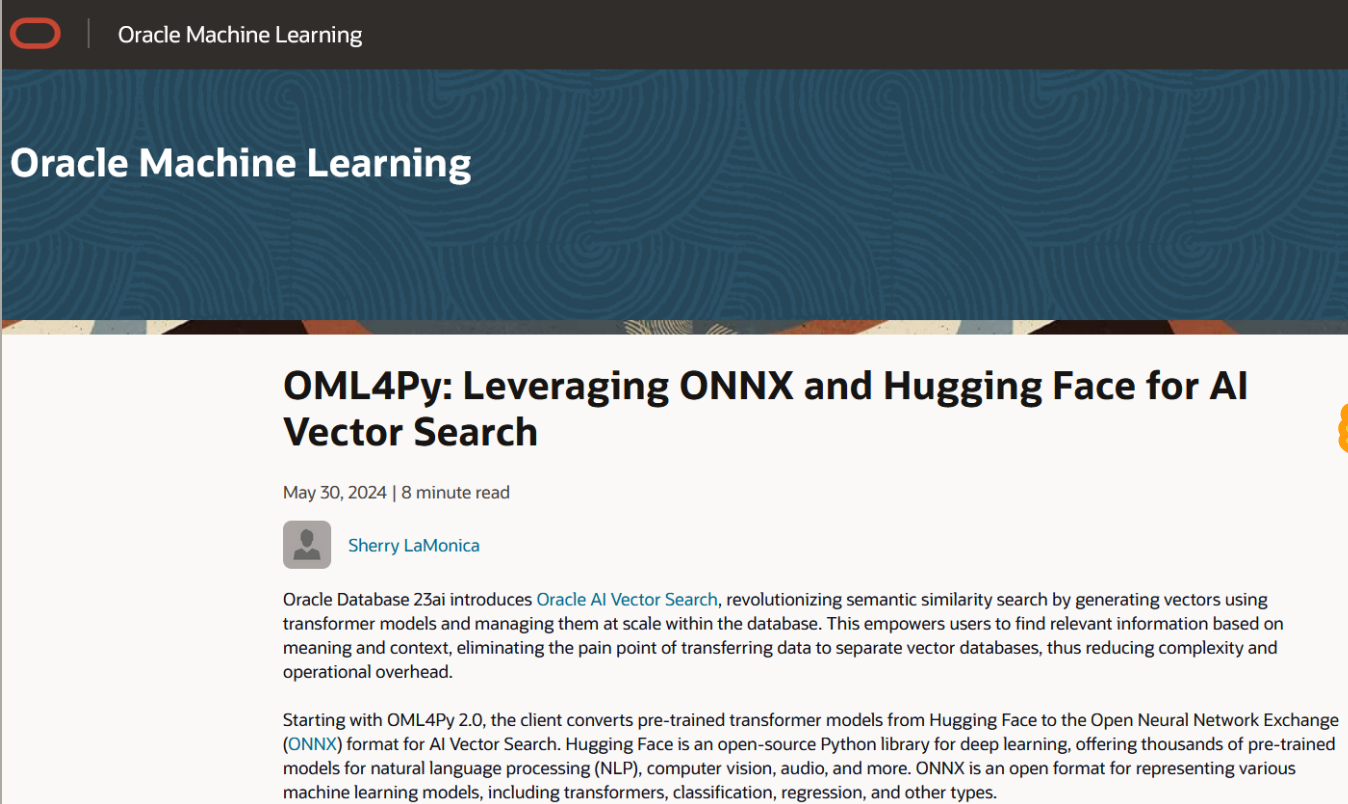
**Table 205-1 DBMS\_VECTOR Package Subprograms**

Subprogram	Description
<b>ONNX Model Related Procedures:</b>	
These procedures enable you to load an ONNX model into Oracle Database and drop the ONNX model.	
<a href="#">LOAD_ONNX_MODEL</a>	Loads an ONNX model into the database
<a href="#">DROP_ONNX_MODEL Procedure</a>	Drops the ONNX model
<b>Chainable Utility (UTL) Functions:</b>	
These functions are a set of modular and flexible functions within vector utility PL/SQL packages. You can chain these together to automate end-to-end data transformation and similarity search operations.	
<a href="#">UTL_TO_CHUNKS</a>	Splits data into smaller pieces or chunks
<a href="#">UTL_TO_EMBEDDING</a> and <a href="#">UTL_TO_EMBEDDINGS</a>	Converts data to one or more vector embeddings
<a href="#">UTL_TO_GENERATE_TEXT</a>	Generates text for a prompt or input string



# ONNX and Hugging Face


Do you remember UNDO replacing Rollback Segments? Read the docs: AI is not all legacy.



The screenshot shows a web page for Oracle Machine Learning. At the top, there is a dark header with the Oracle logo and the text "Oracle Machine Learning". Below this is a blue banner with the text "Oracle Machine Learning". The main content area has a white background and features the following text:

## OML4Py: Leveraging ONNX and Hugging Face for AI Vector Search

May 30, 2024 | 8 minute read

 Sherry LaMonica

Oracle Database 23ai introduces [Oracle AI Vector Search](#), revolutionizing semantic similarity search by generating vectors using transformer models and managing them at scale within the database. This empowers users to find relevant information based on meaning and context, eliminating the pain point of transferring data to separate vector databases, thus reducing complexity and operational overhead.

Starting with OML4Py 2.0, the client converts pre-trained transformer models from Hugging Face to the Open Neural Network Exchange (ONNX) format for AI Vector Search. Hugging Face is an open-source Python library for deep learning, offering thousands of pre-trained models for natural language processing (NLP), computer vision, audio, and more. ONNX is an open format for representing various machine learning models, including transformers, classification, regression, and other types.



And consider the value of being able to say "Hugging Face" to your manager and co-workers.

<https://blogs.oracle.com/machinelearning/post/oml4py-leveraging-onnx-and-hugging-face-for-advanced-ai-vector-search>

# AI Security



## And, Speaking Of Hugging Face (which I like to do as often as I can)



There are very serious security issues related to AI generated code and SQL injection.

The screenshot shows the Hugging Face website interface. At the top left is the Hugging Face logo and name. A search bar contains the text 'Search models, datasets, users...'. Navigation links for 'Models', 'Datasets', 'Spaces', 'Posts', 'Docs', and 'Solutions' are visible. Below the navigation is a breadcrumb link '← Back to Articles'. The main heading of the article is 'CyberSecEval 2 - A Comprehensive Evaluation Framework for Cybersecurity Risks and Capabilities of Large Language Models'. To the right of the heading is an 'Upvote 20' button and a row of user avatars. Below the heading, the publication date 'Published May 24, 2024' is displayed.

<https://huggingface.co/blog/leaderboard-llamaguard>

# Using AI To Write Code Is Far More Dangerous Than Most People Can Think (1:3)

AI coding facilitates Substitution attacks: An issue for most AI but not for Oracle

The image shows a screenshot of an Ars Technica article. At the top left is the 'ars TECHNICA' logo. On the right side of the header are navigation links: 'SECTIONS', 'FORUM', 'SUBSCRIBE', a search icon, and 'SIGN IN'. The article title is 'AI-generated code could be a disaster for the software supply chain. Here's why.' in large white text on a dark background. Above the title is a small green skull icon and the text 'ELECTRIC KOOL-AID LLM TEST'. Below the title is a subtitle: 'LLM-produced code could make us much more vulnerable to supply-chain attacks.' At the bottom left of the article preview is the author's name 'DAN GOODIN' and the date 'APR 29, 2025 6:15 AM' with a comment count of '98'. On the right side of the article preview is an image of a robot sitting at a desk with multiple computer monitors displaying various data visualizations and code. Below the image is a caption: '→ Abstract images of a robot working on many computer screens.'

<https://arstechnica.com/security/2025/04/ai-generated-code-could-be-a-disaster-for-the-software-supply-chain-heres-why>

# Using AI To Write Code Is Far More Dangerous Than Most People Can Think (2:3)

AI coding facilitates Substitution attacks: An issue for most AI but not for Oracle

AI-generated computer code is rife with references to non-existent third-party libraries, creating a golden opportunity for supply-chain attacks that poison legitimate programs with malicious packages that can steal data, plant backdoors, and carry out other nefarious actions, newly published research shows.

The study, which used 16 of the most widely used large language models to generate 576,000 code samples, found that 440,000 of the package dependencies they contained were “hallucinated,” meaning they were non-existent. Open source models hallucinated the most, with 21 percent of the dependencies linking to non-existent libraries. A dependency is an essential code component that a separate piece of code requires to work properly. Dependencies save developers the hassle of rewriting code and are an essential part of the modern software supply chain.

<https://arstechnica.com/security/2025/04/ai-generated-code-could-be-a-disaster-for-the-software-supply-chain-heres-why>

# Using AI To Write Code Is Far More Dangerous Than Most People Can Think (3:3)

AI coding facilitates Substitution attacks: An issue for most AI but not for Oracle

## Package hallucination flashbacks

These non-existent dependencies represent a threat to the software supply chain by exacerbating so-called dependency confusion attacks. These attacks work by causing a software package to access the wrong component dependency, for instance by publishing a malicious package and giving it the same name as the legitimate one but with a later version stamp. Software that depends on the package will, in some cases, choose the malicious version rather than the legitimate one because the former appears to be more recent.

<https://arstechnica.com/security/2025/04/ai-generated-code-could-be-a-disaster-for-the-software-supply-chain-heres-why>

# Final Thoughts



## Please Call Me If ....

OpenAI, or Google, Microsoft or Apple, or anyone else demonstrates they can:

- Perform a point-in-time recovery
- Upgrade a customer from version 1 to version 2
- Provide regular quarterly security patches
- Enable an 800 number for 7x24x365 support
- Comply with Center for Internet Security (CIS)
- Comply with DFARS
- Comply with HIPAA
- Comply with ITAR
- Comply with NIST 800-53
- Comply with PCI-DSS 4.0
- Comply with Sarbanes Oxley
- Comply with Department of Defense Secure Technical Implementation Guidelines

Oracle can and does!



## The Road Ahead

As you start your AI journey to "when, where, and how to incorporate AI into your existing systems," you are right to be concerned about finding the value **and avoiding the mines.**



**Oracle 12c had been released for 2 years when OpenAI was founded on Dec. 11, 2015.**

# Clearing A Path Through The Minefield

The most important aspect of hyperbole self-defense is working with your procurement and policy groups to make sure a few important considerations are not missing from their vendor/product evaluation.



# Questions That Need To Be Asked and Answered Before You Buy

- What is the current version number?
- What tools are certified for backup, restoration and recovery?
- What tools are certified for export and import?
- What tools are certified for replication?
- What encryption standards are supported?
- What compression algorithms are supported?
- What are your support policies and what is the process for opening a Service Request?
- What diagnostic tools are certified for monitoring performance metrics?
- How do we integrate product with (name any software tool you currently require)?
- Is the upgrade process an upgrade in place or is a net new installation required?
- Can you demonstrate to us how your transaction auditing works?
- Can you provide us with a link to your support documentation?
- Can you provide us with a copy of your most recent security patch documentation?
- Can you provide us with a link to your Point-in-Time Recovery documentation?



# Recursive Recursion

Today, almost all of the content available to be scraped and used for Large Language Model (LLM) training was created by humans.

Every minute, of every day, the percentage of AI generated content increases.

There is no way to accurately identify content generated by humans as opposed to content generated by Generative AI and which AI and which LLM were the source of what was generated.

It does not take a lot of imagination to see where this is leading.

Content originally created by Generative AI is used to train Generative AI models and, at some finite point in time, the majority of Generative AI training materials were not human generated: The result is, essentially, a closed loop.

The extremely good news: This is not likely to be an issue for Oracle's AI implementations because we are driving the input from the customer's data such as an ERP application.

# Would You Want Your Surgeon To Practice 1980's Medicine?



# Oracle: Superior Technology Following Our Decades Of Experience



Choose your AI partner with care  
Only **ORACLE** has the decades of  
enterprise experience required for success



# The AI Technologies You Want, The Experienced Partner You Need





**ERROR at line 1:**  
**ORA-00028: your session has been killed**

# Questions?



```
SELECT time_to_learn_new_stuff_again
FROM new_technologies
WHERE keyword IN ('VECTOR', 'AI', 'FUZZY',
                  'ONNX', 'GENERATIVE',
                  'DISTANCE', 'EMBEDDING')
ORDER BY 1;
```

---

Daniel Morgan  
Technical Director Database and Cloud Security  
Oracle Consulting Services

# Next Steps

---

Daniel Morgan  
Technical Director Database and Cloud Security  
Oracle Consulting Services



ORACLE