

Rich Niemiec

Chief Innovation Officer
at Viscosity North America

Understand & Where to use
AI & Machine Learning 101

Conf42 DevSecOps 2023

Thursday • November 30th • 5PM GMT



CONF42

Where to use AI & Machine Learning 101

You've Just Entered the Fifth Dimension with ChatGPT

Conf42 DevSecOps 2023

Thursday • November 30th • 5PM GMT

Machine Learning



"how"
learning algorithms

Deep Learning



Features
architectures

Foundation Models

Functionalities
models

AI



Rich Niemiec

Oracle ACE Director | CIO

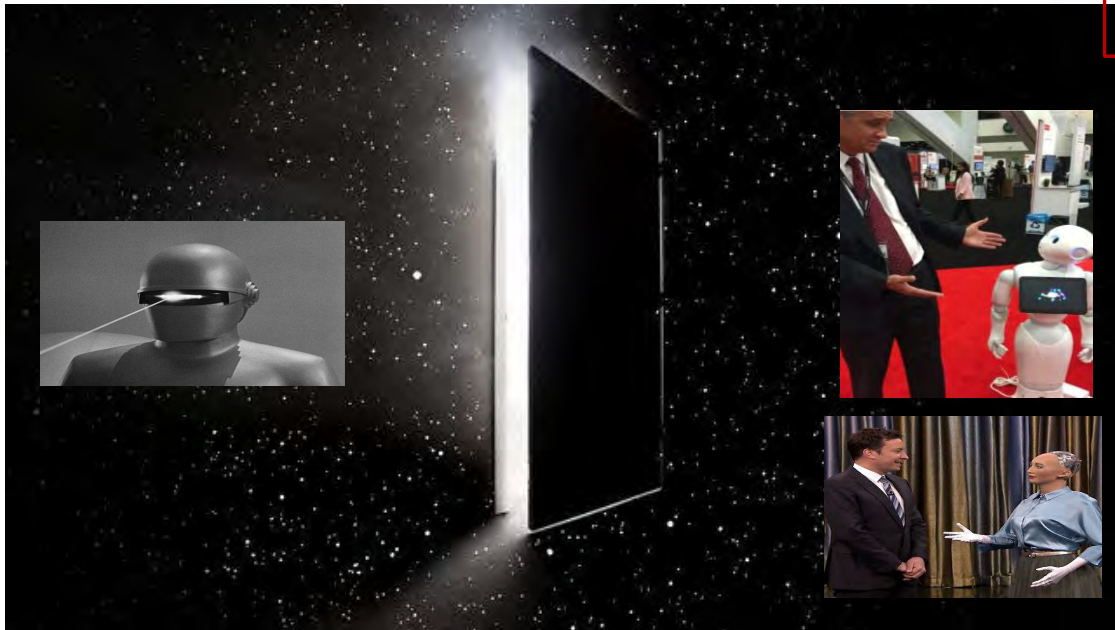
[@richniemiec](#)

[linkedin.com/in/richard-niemiec-b65828/](https://www.linkedin.com/in/richard-niemiec-b65828/)

rich.niemiec@viscosityna.com



ORACLE
ACE Director





Oracle License Management

Get the most out of your Oracle investment



ZERO DOWNTIME
Migrations



Apps

Oracle APEX
SaaS/PaaS
Custom Development



Professional Services

Where you need it most



Performance Health Checks

How's it running?



Staff Aug

Workforce Capacity
on Demand



DBA Services

Remote and On-site



On-Call Support
Managed Services

Viscosity Pillars and Delivery Models

We Enable Business Transformation at a Time When Companies Must Change to Survive!



DATA

Oracle & SQL Server Postgres
Performance Tuning
Data Replication
Data Warehousing Analytics
Data Integration
ERP Blue Prints
Database Upgrades



APPS

APEX
EBS
Web/Mobile Apps
.Net and C#
E-Business Suite
SAAS/PAAS



CLOUD

Azure Gold Partner
Cloud Migrations
Engineered Systems
Oracle Cloud Partner
Google Partner
AWS Partner Hybrid Cloud

Workshops

Assessments

Proof of
Concepts

Training

Turnkey
Projects

Managed
Services



viscosityna.com



[@ViscosityNA](https://twitter.com/ViscosityNA)

Viscosity's Oracle ACEs

The Oracle ACE Program



The Oracle ACE Program recognizes and rewards individuals for their contributions to the Oracle community.



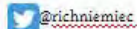
Charles Kim
CEO | Co-Founder



ACE Director



Rich Niemiec
Chief Innovation Officer



ACE Director



Craig Shallahamer
Applied AI Scientist



ACE Director



Sean Scott
Principal Consultant



ACE Director

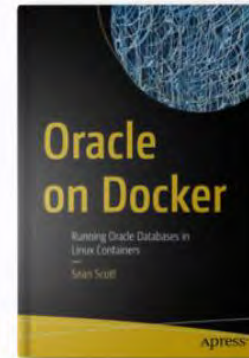
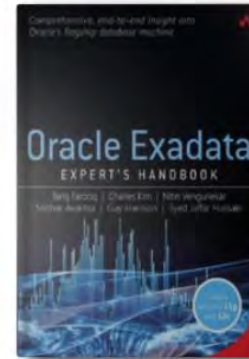


Gary Gordhamer
Principal Consultant



ACE Pro

We've written the Books!



Quick FREE notes

Send email to (for slides):

hello@viscosityna.com

(richniemiec@gmail.com)

@richniemiec - twitter



CONF42



Agenda –Apply Technology & W.I.N.

SUBMITTED
FOR YOUR
APPROVAL

- ❑ Economic Potential of GenAI
- ❑ Economic Impact of Robots
- ❑ ML & Oracle - Overview & ADB
- ❑ Oracle Generative AI Coming
- ❑ Transformers, ChatGPT & Foundation Models
- ❑ Oracle Vector DB, GenAI, & Cohere
- ❑ Machine Learning Future, Robots & VR



Quantum Entanglement

“Spooky action-at-a-distance”
(A. Einstein)



FYI

The economic potential of GenAI*

- McKinsey & Co.¹

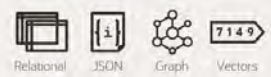
The very nature of digital data presents a challenge

Digital

Can be copied perfectly, infinitely

Data

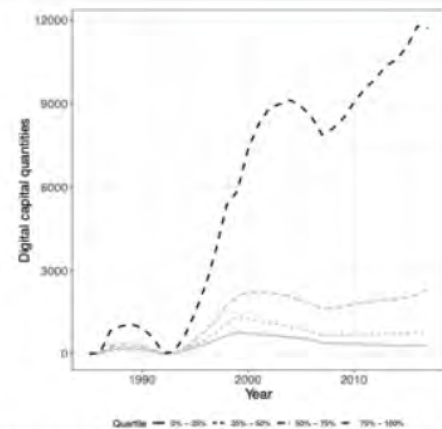
Shape of origin is rarely shape of use



Digital Capital and Superstar Firms¹

25%

of a firm's value is digital capital



Digital capital quantities by market value

Digital capital is more concentrated than any other asset class

Most value concentrated in top decile of firms by market value

Digital capital accumulation predicts firm productivity 3 years out



Requires a whole system to create, repurpose, and protect data products



Requires shapeshifting to achieve maximum value

*Data-driven Transformation, Oracle

Symbiotic coming Relationship; You & Robots



Deepen your digital skills with our NEW [Executive Certificate in Digital Business](#).



COURSES FOR INDIVIDUALS

CERTIFICATE PROGRAMS

ONLINE COURSES

SOLUTIONS FOR ORGANIZATIONS

WEBINARS, BLOG & MORE

Courses for Individuals



Download Brochure

Enroll Now!

Select Format

Select Format

Accelerating Digital Transformation with Algorithmic Business Thinking

favorite takeaways, to effectively harness advancements in technologies, organizations need to engineer the human centered genes into their organizations and "take the robot out of the human".

—Martha Anderson Sr. Director, Digital Transformation at Walmart

Duration	Time Commitment	Price
2 days	8 hours/day	\$4,300

Topics

- Digital Business & IT
- Organizations & Leadership
- Strategy & Innovation

\$2,800 - \$4,300

Add to Cart

Why generative AI needs a creative human touch



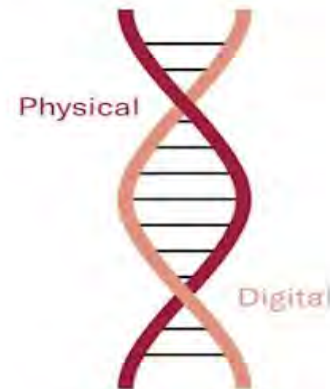
by Beth Stackpole | Jun 5, 2023

Why It Matters

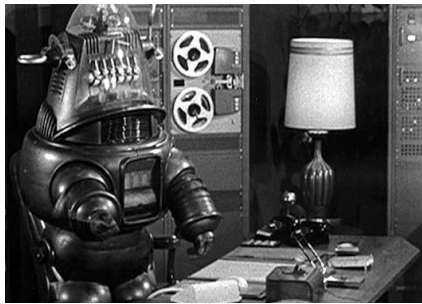
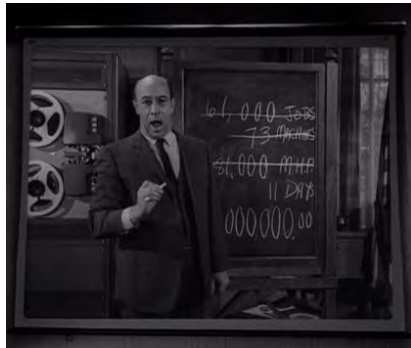


{ Human × Machine Capability_

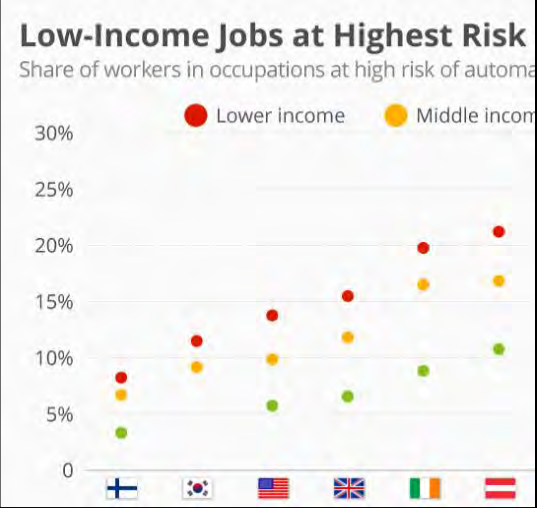
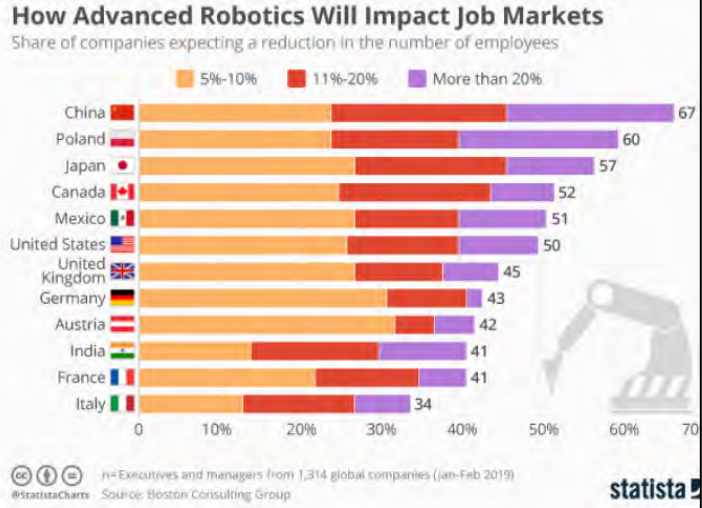
- /* Today we live, work, and operate, in 2 worlds: physical, and digital
- /* Combining human and machine capabilities is a multiplier
- /* Our organizations will evolve in partnership with technology
- /* By exploring and experimenting, we will evolve our organizations



The Brain Center at Whipple's & ChatGPT

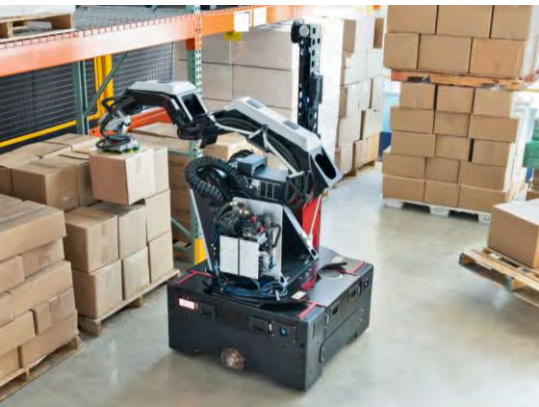


Robotics/Automation Impact to Jobs



The New York Times

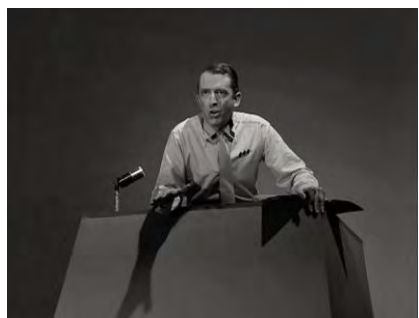
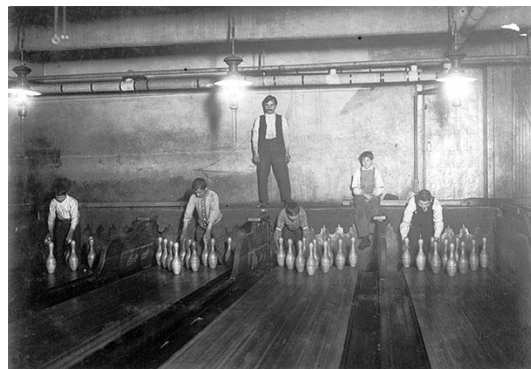
Uber's Self-Driving Trucks Hit the Highway, but Not Local Roads



Leverage – DB, GPS & Robotics! Impact to Jobs?



The Obsolete Man



Autonomous Database – Replacing the DBA?



Rich Niemiec @RichNiemiec · Oct 2
I'll ask #Pepper if she can tune my #database in 11 AM session on
#innovation #iot #robotics #cloud at #oow17 #ioug #viscosityna #oraclece



Are DBAs Obsolete?

Posted on February 20, 2014

Before we go any further, let me briefly answer the question
“No Way!” OK ... with that out of the way, let me

Every so often, some industry pundit gets h
“Database administrators are obsolete” or t
hear this, it makes me shake my head sadly
can be



OOH HOME | OCCUPATION FINDER | OOH FAQ | OOH GLOSSARY | A-Z INDEX | OOH SITE
OCCUPATIONAL OUTLOOK HANDBOOK

Occupational Outlook Handbook > Computer and Information Technology >

Database Administrators

Summary What They Do Work Environment How to Become One Pay Job Outlook

Summary

Quick Facts: Database Administrators

2017 Median Pay	\$87,020 per year \$41.84 per hour
Typical Entry-Level Education	Bachelor's degree
Work Experience in a Related Occupation	None
On-the-job Training	None
Number of Jobs, 2016	119,500
Job Outlook, 2016-26	11% (Faster than average)
Employment Change, 2016-26	13,700

What Database Administrators Do

Database administrators (DBAs) use specialized software to store and organize data, such as information and customer shipping records. They make sure that data are available to users

**Good News: DBA +11% Increase
(2016-2026)**

Biju Thomas - *Emerging Jobs (Developer/DBA)*

#8 Data Engineer

LinkedIn

2020
Emerging
Jobs Report

- 33% annual growth
- Skills
 - Apache Spark
 - Hadoop
 - Python
 - ETL
 - AWS
- Industries
 - Computer Software
 - Information Technology
 - Financial Services
 - Healthcare & Hospitals
- Primary job responsibilities involve preparing data for analytical or operational uses.
- Works as part of an analytics team, providing data in a ready-to-use form to data scientists.
- Commonly deal with both structured and unstructured data sets

Characteristics of Big Data - The Five V's

Volume

Big data comes in one size: large. Enterprises are awash with data, easily amassing terabytes and even petabytes of information.

TB, Records, Transactions, Tables, Files

Velocity

Often time-sensitive, big data is as it is streaming in to the system to maximize its value to the user.

Batch, Near time, Real time

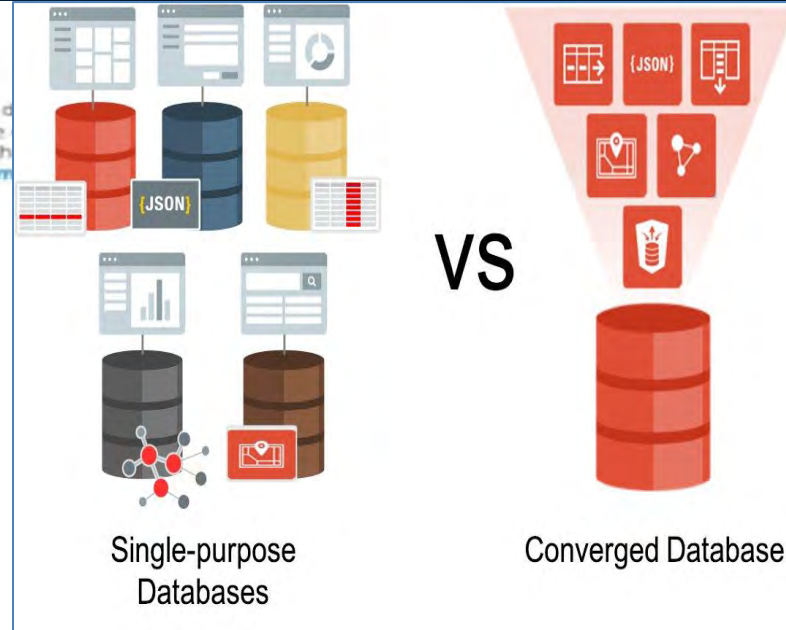
Value

Business value of Big Data

Variety

Big data extends beyond structured data, including semi-structured and unstructured data of all varieties: text, audio, video, click streams, log files and more.

Structured, Unstructured, Semistructured



Big Data Themes

- HW & SW technologies for large data volumes
- Focus on Web 2.0 technologies
- Database Scale-out
- Relational & Distributed Data Analytics
- Real Time Analytics

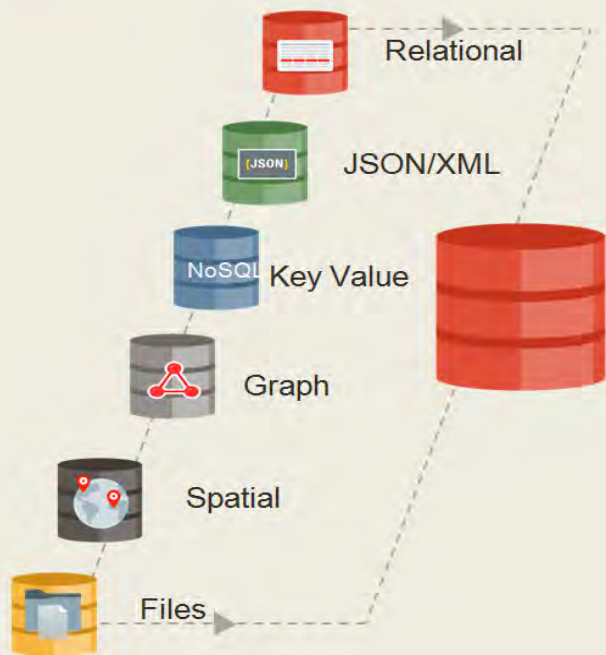
Big Data Domains

- Digital Marketing Optimization
- Data Exploration & Discovery
- Fraud Detection & Prevention
- Social Network & Relationship Analysis
- Machine-generated Data Analytics

Converged Database - Oracle Multi-Model Database*

- Benefits of Oracle's Converged Database are broad

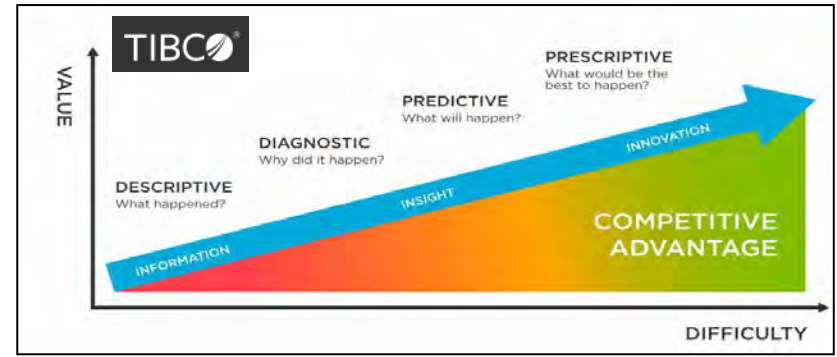
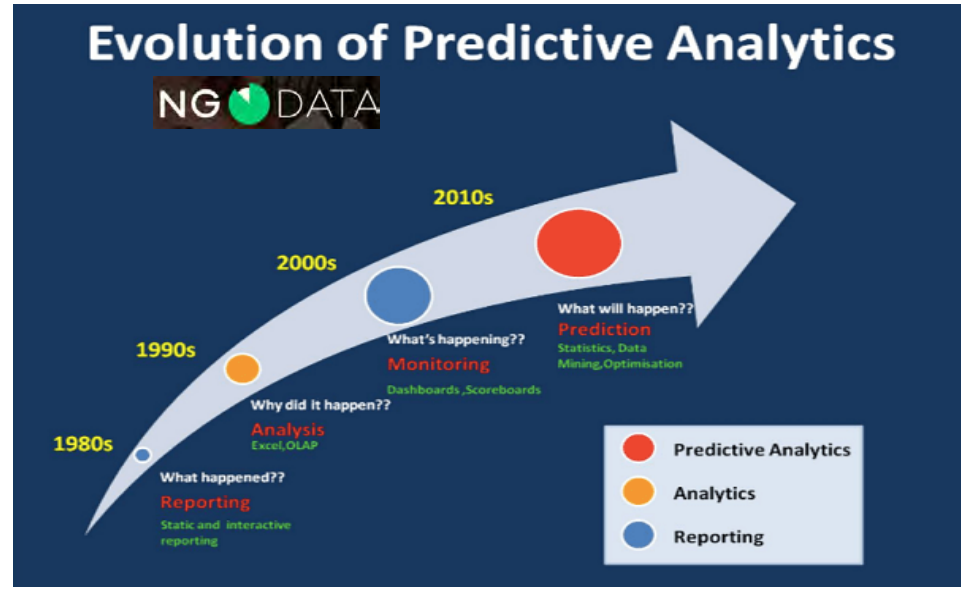
*Slide from Oracle's,
Nitin Vengurlekar



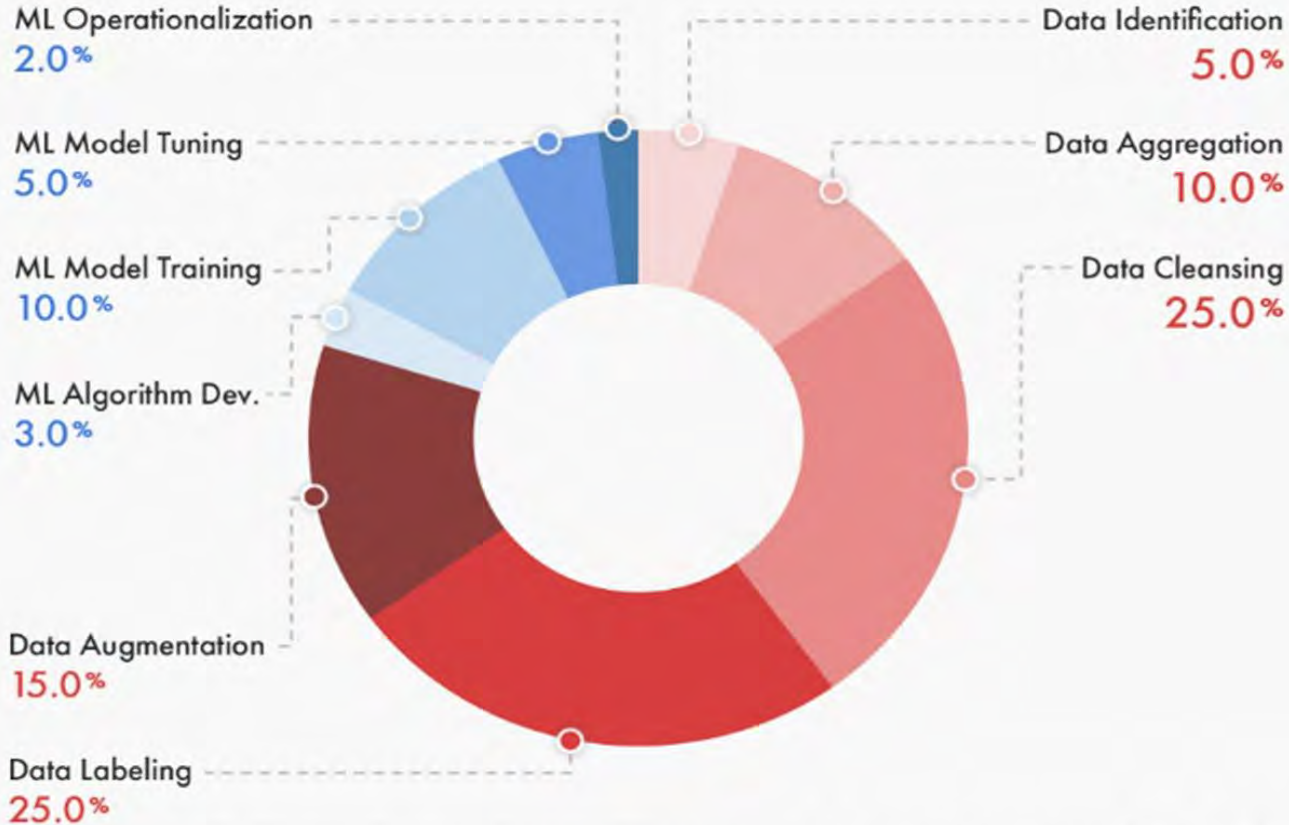
- Integrated development tools (Apex, SQL Dev, Spatial Studio) ✓
- 3rd party and Open Source development tools ✓
- Machine Learning ✓
- Node.js, Python, many others ✓
- In-memory database ✓
- Spatial, Graph support ✓
- NoSQL (JSON, key-value, wide column, XML) ✓
- Containers, microservices, virtualization (Docker, MT) ✓
- Integrated Security ✓
- Deployment choice (on-prem, cloud, hybrid) ✓
- Integrated High Availability and Disaster Recovery ✓



What You Need; Nick of Time (Predictive Analytics)

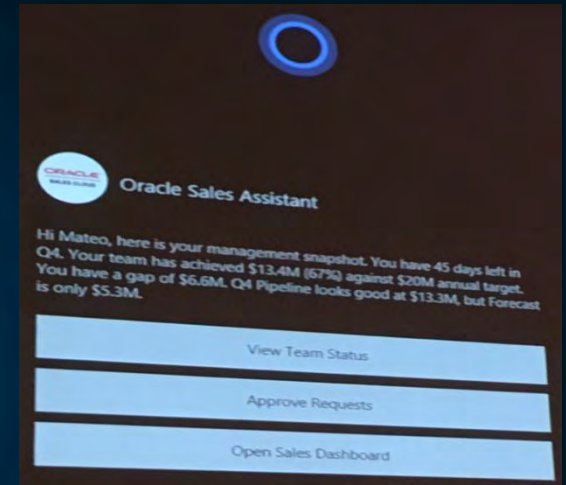


Percentage of Time Allocated to Machine Learning Project Tasks



**The DBA is
most
important
part of the
Machine
Learning
Process**

Niti Sharma An Introduction to Data Labeling in Artificial Intelligence *Data Wrangling consumes over 80% of the time in AI projects.*



A Robot may not look like one!

Oh Yeah... they never complain, always happy to do more, work anywhere, get smarter as time goes on, leverages AI & ML, works 24x7, doesn't ask for a raise, no union (yet).

A large red banner is stretched across the facade of a modern glass skyscraper. The banner features the Oracle logo and the text 'Database Cloud' on the left, and 'The World's First Autonomous Database' on the right. The building's glass reflects the surrounding city skyline.

ORACLE®
Database Cloud

The World's First Autonomous Database

A Robot may not look one!

*Robots that Manage a Database (ADW)!

*Robots that secure a system and use ML & AI

Autonomous DB : Future DBA & Robot DB

- Self-Managing (Driving)
- Self-Securing
- Self-Tuning
- Self-Recovering
- Self-Scaling Administration



Fully automated patching, upgrades, backups, & availability architecture

**Oracle Unveils World's First
Autonomous Database Cloud**



Oracle Focus: Cloud, Data, Apps, Security, Self-Service, ML, & AI!
Innovation Components for Business!



ORACLE
OPENWORLD 2013
SAN FRANCISCO



"I admire risk takers. I like leaders – people who do things before they become fashionable or popular. I find that kind of integrity inspirational."

LAWRENCE J. ELLISON | *Chairman & Chief Executive Officer, 2002*



Oracle Machine Learning: **Brief Highlights Only**



READY-TO-GO

ORACLE

SaaS APPS w/EMBEDDED ML

Pre-built, packaged ML and data-driven SaaS applications

READY-TO-WORK

ORACLE

AUTONOMOUS DATABASE

Embedded AI in Databases to simplify enterprise data management

READY-TO-BUILD

ORACLE

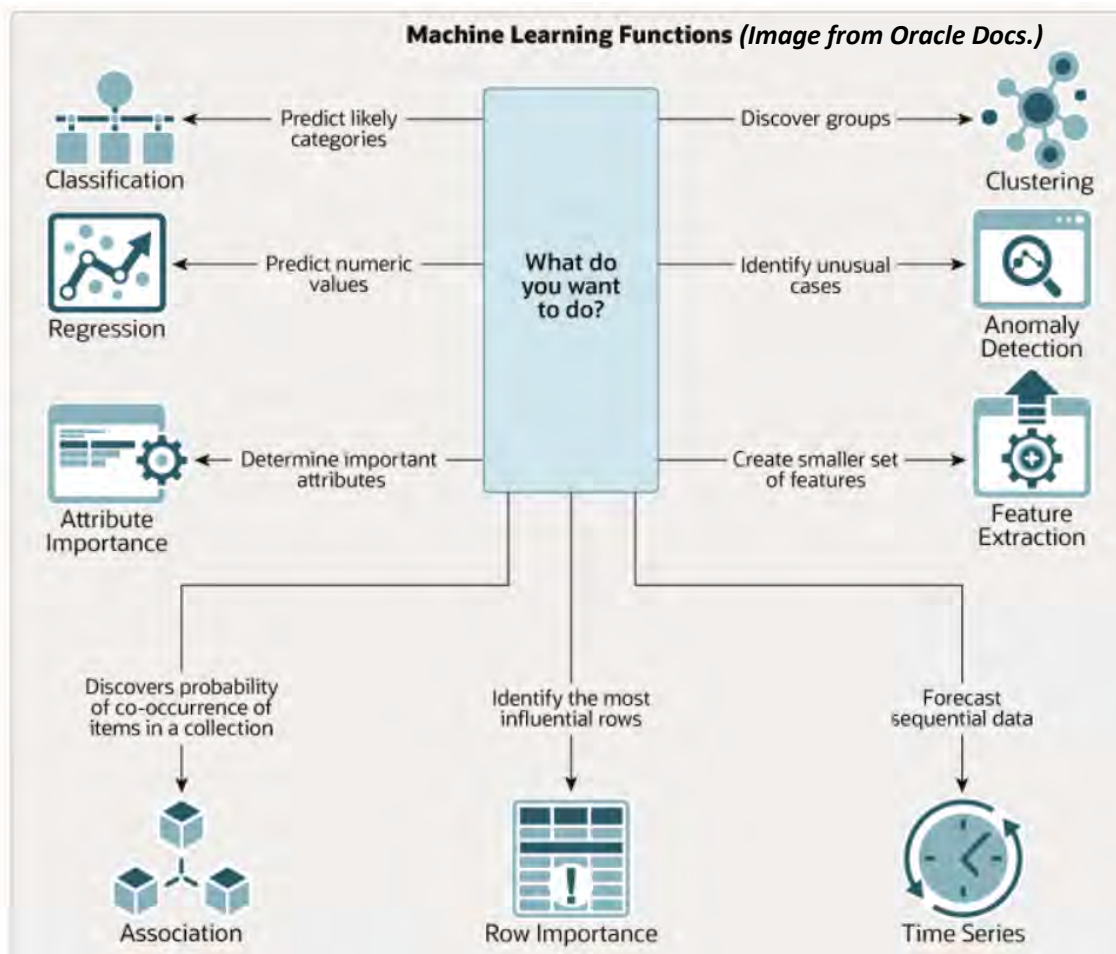
DATA SCIENCE PLATFORM

Complete platform to build and support ML-powered applications

Machine Learning Process (Supervised Learning)

- **First:** Clear *Business Problem to Solve*
 - **Second:** *Function to Perform*
 - **Third:** *Algorithms to use*
-
- **First:** *Build/Train the Model:* When you build it use about 60% of your data.
 - **Second:** *Test/Score Model* for accuracy/precision using about 40% of data.

Compare Algorithms!



Business Understanding

Be Extremely Specific in Problem Statement:

Poorly Defined	Better	Data Mining Technique
Predict employees that leave	<ul style="list-style-type: none"> Based on past employees that voluntarily left: <ul style="list-style-type: none"> Create New Attribute EmplTurnover → 0/1 	
Predict customers that churn	<ul style="list-style-type: none"> Based on past customers that left (churn): <ul style="list-style-type: none"> Create New Attribute Churn → YES/NO 	
Target “best” customers	<ul style="list-style-type: none"> Recency, Frequency Monetary (RFM) Analysis Specific Dollar Amount over Time Window: <ul style="list-style-type: none"> Who has spent \$500+ in most recent 18 months 	
How can I make more \$\$?	<ul style="list-style-type: none"> What helps me sell soft drinks & coffee? 	
Which customers are likely to buy?	<ul style="list-style-type: none"> How much is each customer likely to spend? 	
Who are my “best customers”?	<ul style="list-style-type: none"> What descriptive “rules” describe “best customers”? 	
How can I combat fraud?	<ul style="list-style-type: none"> Which transactions are the most anomalous? <ul style="list-style-type: none"> Then roll-up to physician, claimant, employee, etc. 	

OAA Model Build and Real-time SQL Apply

Simple SQL Syntax - Attribute Importance

ML Model Build (PL/SQL)

```
BEGIN
  DBMS_DATA_MINING.CREATE_MODEL (
    model_name          => 'BUY_INSURANCE_AI',
    mining_function     => DBMS_DATA_MINING.ATTRIBUTE_IMPORTANCE,
    data_table_name     => 'CUST_INSUR_LTV',
    case_id_column_name => 'cust_id',
    target_column_name  => 'BUY_INSURANCE',
    settings_table_name => 'CUST_INSUR_LTV_SET');
END;
/
```

Additional Detail:

```
drop table CUST_INSUR_LTV_SET;
exec
dbms_data_mining.drop_model('BUY_INSURANCE_AI');
```

```
create table CUST_INSUR_LTV_SET (setting_name
varchar2(30), setting_value varchar2(4000));
```

```
insert into CUST_INSUR_LTV_SET values
('ALGO_NAME','ALGO_SUPPORT_VECTOR_MACHINES');
insert into CUST_INSUR_LTV values ('PREP_AUTO','ON');
commit;
```

Model Results (SQL query)

```
SELECT attribute_name, explanatory_value, rank
FROM BUY_INSURANCE_AI
ORDER BY rank, attribute_name;
```

<u>ATTRIBUTE NAME</u>	<u>RANK</u>	<u>ATTRIBUTE VALUE</u>
BANK_FUNDS	1	0.2161
MONEY_MONTHLY_OVERDRAWN	2	0.1489
N_TRANS_ATM	3	0.1463
N_TRANS_TELLER	4	0.1156
T_AMOUNT_AUTOM_PAYMENTS	5	0.1095

OAA Model Build and Real-time SQL Apply Prediction

Oracle Advanced Analytics (OAA) Simple SQL - Classification

ML Model Build (PL/SQL)

```
BEGIN
  DBMS_DATA_MINING.CREATE_MODEL(
    model_name          => 'BUY_INSUR1',
    mining_function     => dbms_data_mining.classification,
    data_table_name     => 'CUST_INSUR_LTV',
    case_id_column_name => 'CUST_ID',
    target_column_name  => 'BUY_INSURANCE',
    settings_table_name => 'CUST_INSUR_LTV_SET');
END;
/
```

Additional Detail:

```
drop table CUST_INSUR_LTV_SET;
exec dbms_data_mining.drop_model('BUY_INSUR1');
create table CUST_INSUR_LTV_SET
(setting_name varchar2(30), setting_value varchar2(4000));

insert into CUST_INSUR_LTV_SET values
('ALGO_NAME','ALGO_SUPPORT_VECTOR_MACHINE');
insert into CUST_INSUR_LTV_SET values ('PREP_AUTO','ON');
commit;
```

Model Apply (SQL query)

```
Select prediction_probability(BUY_INSUR1, 'Yes'
  USING 3500 as bank_funds, 825 as checking_amount, 400 as credit_balance, 22 as age,
  'Married' as marital_status, 93 as MONEY_MONTHLY_OVERDRAWN, 1 as house_ownership)
from dual;
```

SQL All Rows Fetched: 1 in 0.043 seconds	
PREDICTION_PROBABILITY(BUY_INSUR1,'YES'USING3500ASBANK_FUNDS,825ASCHECKING_AMOUNT,400ASCREDIT_BALANCE	
1	0.9276956709910801

DBMS_DATA_MINING Oracle Algorithms*

FYI

Algorithm	Abbreviation	Function	*Oracle Database PL/SQL Packages and Types Referen
Apriori	AR	Association	
CUR Matrix Decomposition	CUR	Attribute Importance	
Decision Tree	DT	Classification	
Expectation Maximization	EM	Clustering	
Explicit Semantic Analysis	ESA	Feature Extraction, Classification	
Exponential Smoothing	ESM	Time Series	
Generalized Linear Model	GLM	Classification, Regression	
<i>k</i> -Means	KM	Clustering	
Minimum Descriptor Length	MDL	Attribute Importance	
Naive Bayes	NB	Classification	
Neural Networks	NN	Classification, Regression	
Non-Negative Matrix Factorization	NMF	Feature Extraction	
Orthogonal Partitioning Clustering	O-Cluster	Clustering	
Random Forest	RF	Classification	
Singular Value Decomposition and Principal Component Analysis	SVD and PCA	Feature Extraction	
Support Vector Machine	SVM	Classification, Regression, Anomaly Detection	

DBMS_DATA_MINING Oracle Algorithms *

FYI

ALGO_NAME Value	Description	Mining Function
ALGO_AI_MDL	Minimum Description Length	Attribute Importance
ALGO_APRIORI_ASSOCIATION_RULES	Apriori	Association Rules
ALGO_CUR_DECOMPOSITION	CUR Decomposition	Attribute Importance
ALGO_DECISION_TREE	Decision Tree	Classification
ALGO_EXPECTATION_MAXIMIZATION	Expectation Maximization	Clustering
ALGO_EXPLICIT_SEMANTIC_ANALYSIS	Explicit Semantic Analysis	Feature Extraction Classification
ALGO_EXPONENTIAL_SMOOTHING	Exponential Smoothing	Time Series
ALGO_EXTENSIBLE_LANG	Language used for extensible algorithm	All mining functions supported
ALGO_GENERALIZED_LINEAR_MODEL	Generalized Linear Model	Classification, Regression; also Feature Selection and Generation
ALGO_KMEANS	Enhanced k -Means	Clustering
ALGO_NAIVE_BAYES	Naive Bayes	Classification
ALGO_NEURAL_NETWORK	Neural Network	Classification
ALGO_NONNEGATIVE_MATRIX_FACTORIZATION	Non-Negative Matrix Factorization	Feature Extraction
ALGO_O_CLUSTER	O-Cluster	Clustering
ALGO_RANDOM_FOREST	Random Forest	Classification
ALGO_SINGULAR_VALUE_DECOMP	Singular Value Decomposition	Feature Extraction
ALGO_SUPPORT_VECTOR_MACHINES	Support Vector Machine	Classification and Regression

These are the **actual algorithm names** that need to be use.

These are **values for ALGO_NAME** setting

There are many additional settings that go with each Mining Function (i.e. **Confidence, Length, Rules, Aggregates, Cost, Number of Clusters...etc.**)

Machine Learning in ADW/ATP - www.oracle.com/cloud/free/



DB-20220510121752

Database Actions DB Connection Performance Hub Service Console More Actions

Autonomous Database Information Tools Tags

Autonomous Database Provision - 2 minutes Start - 30 seconds

Development SQL SQL Execute query create database LIQUIBASE Track schema SCHEDULED Schedule, manage tasks APEX

ORACLE Machine Learning

Example Templates

+ Create Notebook

Template Name	Description	Author	Date Added	Tags	Likes	Views	Comments
OML4R Clustering EM	This notebook builds and applies a clustering mod...	Oracle	9/30/21, 8:40 AM	'Clustering' 'Expectation Maximization' 'SH.S...	0	11	1
OML4R Clustering KM	This notebook builds and applies a clustering mod...	Oracle	11/26/20, 5:41 AM	'Clustering' 'K-Means' 'SH.SUPPLEMENTARY_...	0	107	5
OML4R Clustering OC	This notebook builds and applies a clustering mod...	Oracle	2/17/21, 5:18 PM	'Python' 'Data Cleaning' 'Duplicates Removal'...	0	76	2
OML4Py Classification SVM	This notebook builds and applies a Support Vector...	Oracle	9/30/21, 8:40 AM	'Create View' 'SH.SUPPLEMENTARY_DEMOGR...	0	16	1
OML4Py Clustering EM	This notebook builds and applies a clustering mod...	Oracle	9/30/21, 8:40 AM	'Clustering' 'Expectation Maximization' 'SH.S...	0	11	1
OML4Py Clustering KM	This notebook builds and applies a clustering mod...	Oracle	11/26/20, 5:41 AM	'Clustering' 'K-Means' 'SH.SUPPLEMENTARY_...	0	107	5
OML4Py Data Cleaning Duplicates Re...	This notebook illustrates removing duplicate rows ...	Oracle	2/17/21, 5:18 PM	'Python' 'Data Cleaning' 'Duplicates Removal'...	0	76	2
OML4SQL Classification DT	This notebook builds and applies a Decision Tree Classifica...	Oracle	11/26/20 5:40 AM	'SQL' 'Classification' 'Decision Tree' 'sql' 'Create View' ...	0	175	27
OML4SQL Classification GLM	This notebook builds and applies a Generalized Linear Clas...	Oracle	9/30/21 8:40 AM	'sql' 'Create View' 'SH.SUPPLEMENTARY_DEMOGRAP...	0	13	1
OML4SQL Classification NB	This notebook builds and applies a Naïve Bayes Classificati...	Oracle	9/30/21 8:40 AM	'sql' 'Create View' 'SH.SUPPLEMENTARY_DEMOGRAP...	0	4	1
OML4SQL Classification NN	This notebook builds and applies a Neural Network Classifi...	Oracle	9/30/21 8:40 AM	'sql' 'Create View' 'SH.SUPPLEMENTARY_DEMOGRAP...	0	28	2

Machine Learning connection to ADW/ATP

Build Anomaly Detection model (1-Class Support Vector Machine)

```
%script
--Build Anomaly Detection Model (1-Class SVM) on CUSTOMERS360 data

DECLARE
v_sql varchar2(100);

BEGIN

--Create a Build Setting table for Model Build

EXECUTE IMMEDIATE 'CREATE TABLE CUSTOMERS360_SET (setting_name VARCHAR2(30),setting_value VARCHAR2(400))';
EXECUTE IMMEDIATE 'INSERT INTO CUSTOMERS360_SET (setting_name, setting_value) VALUES (''ALGO_NAME'', ''ALGO_SUPPORT_VECTOR_MACHINES'')';
EXECUTE IMMEDIATE 'INSERT INTO CUSTOMERS360_SET (setting_name, setting_value) VALUES (''PREP_AUTO'', ''ON'')';
DBMS_OUTPUT.PUT_LINE ('Created model build settings table: CUSTOMERS360_SET ');

--Build the 1-Class SVM model.

EXECUTE IMMEDIATE 'CALL DBMS_DATA_MINING.CREATE_MODEL(''CUSTOMERS360MODEL'', ''CLASSIFICATION'', ''CUSTOMERS360'', ''CUST_ID'', null, ''CUSTOMERS360_SET'')';
DBMS_OUTPUT.PUT_LINE ('Created model: CUSTOMERS360_MODEL ');

END;
```

Created model build settings table: CUS
Created model: CUSTOMERS360_MODEL
PL/SQL procedure successfully completed

Graph all Customers's PROBABILITY_ANOMALOUS vs. YRS_RESIDENCE grouped by CUST_MARITAL_STATUS



Machine Learning connection to ADW/ATP

Display the top 15 most anomalous customers

FINISHED

```
%sql
--Display the Top 15 Most Anomalous Customers

SELECT *
FROM (SELECT CUST_ID, round(probability_anomalous,2) probability_anomalous, HOUSEHOLD_SIZE, YRS_RESIDENCE,
      CUST_GENDER, CUST_MARITAL_STATUS, rank() over (ORDER BY probability_anomalous DESC) rnk
      FROM (SELECT CUST_ID, HOUSEHOLD_SIZE, YRS_RESIDENCE, CUST_GENDER,
      CUST_MARITAL_STATUS, prediction_probability(CUSTOMERS360MODEL, '0' using *) probability_anomalous
      FROM CUSTOMERS360))
WHERE rnk <= 15
ORDER BY probability_anomalous DESC;
```

View Prediction_Details that explain why the record was selected as anomalous

FINISHED

```
%sql
-- Select customers with OCCUPATION of 'TechSup' and more than 55% probability of being anomalous

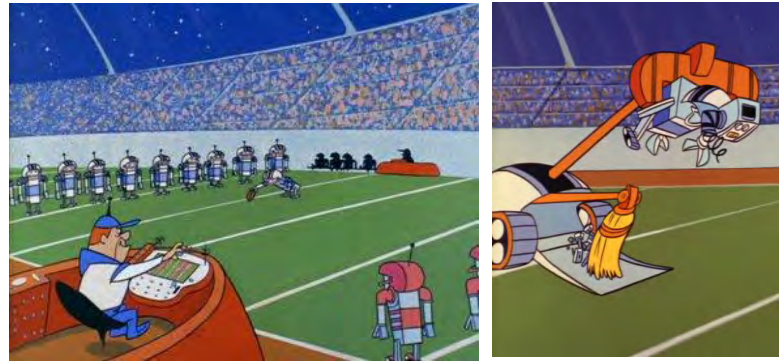
SELECT CUST_ID,
      RTRIM(TRIM(SUBSTR(OUTPRED."Attribute1",17,100)), 'rank="1"/>') FIRST_ATTRIBUTE,
      RTRIM(TRIM(SUBSTR(OUTPRED."Attribute2",17,100)), 'rank="2"/>') SECOND_ATTRIBUTE,
      RTRIM(TRIM(SUBSTR(OUTPRED."Attribute3",17,100)), 'rank="3"/>') THIRD_ATTRIBUTE,
      RTRIM(TRIM(SUBSTR(OUTPRED."Attribute4",17,100)), 'rank="4"/>') FOURTH_ATTRIBUTE,
      RTRIM(TRIM(SUBSTR(OUTPRED."Attribute5",17,100)), 'rank="5"/>') FIFTH_ATTRIBUTE
FROM (SELECT CUST_ID, PREDICTION_DETAILS(CUSTOMERS360MODEL, '0' USING *) PREDICTION_DETAILS FROM CUSTOMERS360
WHERE PREDICTION_PROBABILITY(CUSTOMERS360MODEL, '0' USING *) > 0.50
AND OCCUPATION = 'TechSup'
ORDER BY CUST_ID) OUT,
```

CUST_ID	PR
103154	0.7
100199	0.7
102948	0.7
101137	0.6
104359	0

CUST_ID	FIRST_ATTRIBUTE	SECOND_ATTRIBUTE	THIRD_ATTRIBUTE
100646	"CUST_MARITAL_STATUS" actualValue="Widowed" weight=".226"	"CUST_YEAR_OF_BIRTH" actualValue="1941" weight=".118"	"CUST_CREDIT_LIMIT" actualValue="1500" weight="."
102922	"CUST_MARITAL_STATUS" actualValue="Widowed" weight=".222"	"CUST_YEAR_OF_BIRTH" actualValue="1931" weight=".169"	"CUST_CREDIT_LIMIT" actualValue="1500" weight="."
103441	"CUST_MARITAL_STATUS" actualValue="Widowed" weight=".222"	"CUST_YEAR_OF_BIRTH" actualValue="1941" weight=".117"	"EDUCATION" actualValue="Bach." weight=".076"
104286	"EDUCATION" actualValue="9th" weight=".165"	"HOUSEHOLD_SIZE" actualValue="4-5" weight=".146"	"CUST_CREDIT_LIMIT" actualValue="1500" weight="."

A Game of Pool (*Talent / Luck / Work / Nerve*)

(Your Current Team Plays Robots built on the Best Past Teams or ML)



Gambling on virtual reality: the online casinos of the future

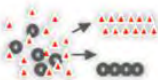


Image Credit: SlotsMillion

Oracle Machine Learning Algorithms and Analytics in Oracle Database

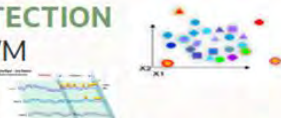
CLASSIFICATION

- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**



ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES

- Forecasting - Exponential Smoothing
- Includes popular models e.g. Holt-Winters with trends, seasonality, irregular time series



[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION

- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**



ATTRIBUTE IMPORTANCE

- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI



ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

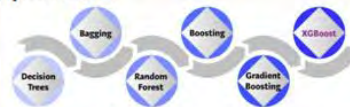
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



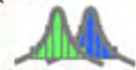
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



Includes support for partitioned models, integrated text mining, automated data preparation

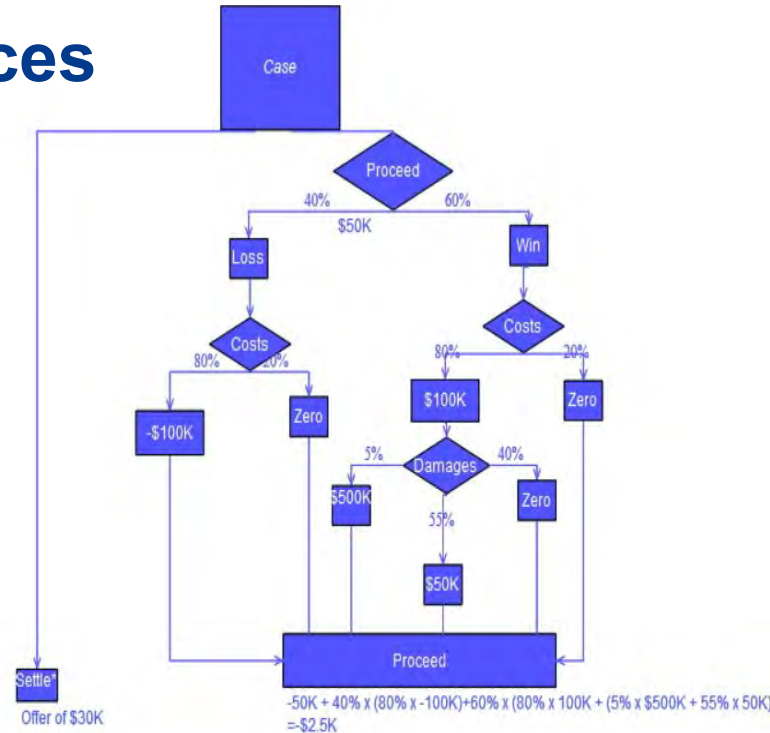
** New in 21c*



Decision Tree Algorithm (ML Classifier)

- Tree model for decisions
- Decisions made to reach goal
- **Calculate odds of different choices**
- Can have chance outcomes
- **Calculate costs & utility (value)**
- Could be used for chatbots (to answer questions)

Flow Chart calculates when to settle a case or not based on costs/probabilities



OML (OAA) Oracle Data Mining SQL Sample (PARTIAL)

dmdtdemo.sql → DBMS_DATA_MINING package – Decision Tree

```
-- Given demographic data about a set of customers, predict the  
-- customer response to an affinity card program using a classifier  
-- based on Decision Trees algorithm.
```

```
INSERT INTO dt_sh_sample_settings VALUES  
(dbms_data_mining.algo_name, dbms_data_mining.algo_decision_tree);  
--  
-- CREATE A NEW MODEL  
-- Build a DT model
```

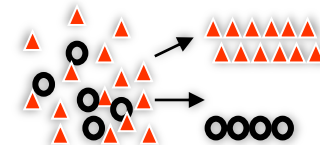
```
BEGIN  
DBMS_DATA_MINING.CREATE_MODEL(  
  model_name      => 'DT_SH_Class_sample',  
  mining_function => dbms_data_mining.classification,  
  data_table_name => 'mining_data_build_v',  
  case_id_column_name => 'cust_id',  
  target_column_name => 'affinity_card',  
  settings_table_name => 'dt_sh_sample_settings');  
END;  
/
```



```
SELECT T.cust_id, S.prediction, S.probability, S.cost  
FROM (SELECT cust_id,  
  PREDICTION_SET(dt_sh_class_sample COST MODEL USING *) pset  
FROM mining_data_apply_v  
WHERE cust_id < 100011) T,  
TABLE(T.pset) S  
ORDER BY cust_id, S.prediction;
```

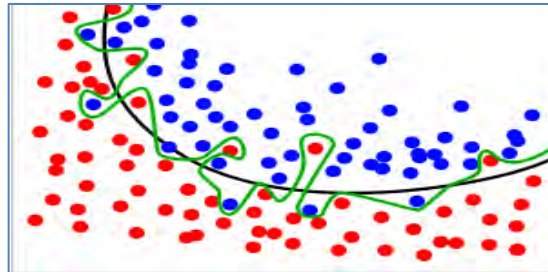
Script Output x
Task completed in 3.075 seconds

CUST_ID	PREDICTION	PROBABILITY	COST
100001	0	.966183575	.270531401
100001	1	.033816425	.966183575
100002	0	.740384615	2.076923077
100002	1	.259615385	.740384615
100003	0	.909090909	.727272727
100003	1	.090909091	.909090909
100004	0	.909090909	.727272727
100004	1	.090909091	.909090909
100005	0	.272357724	5.821138211
100005	1	.727642276	.272357724
100006	0	1.000000000	.000000000
100006	1	.000000000	1.000000000
100007	0	.909090909	.727272727
100007	1	.090909091	.909090909
100008	0	.909090909	.727272727
100008	1	.090909091	.909090909
100009	0	.272357724	5.821138211
100009	1	.727642276	.272357724
100010	0	.675965665	2.592274678
100010	1	.324034335	.675965665

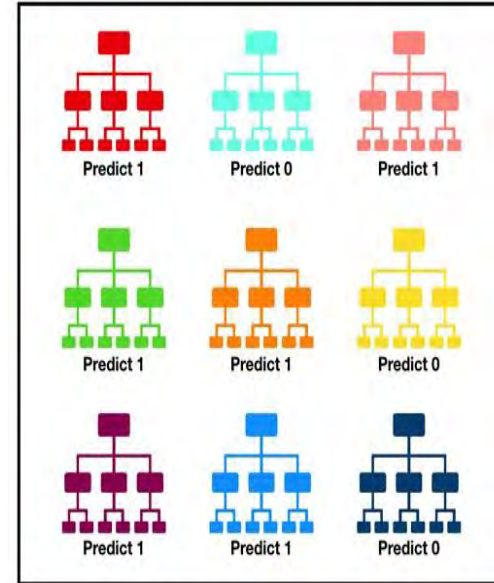


Random Forest (ML Classifier)

- Multitude of Decision Trees (picks the mode)
- **Output: Class of decision tree most likely**
- *Mode* of the classes is mean or *most likely*
- Mode is found during training time
- **Separate the Real Data from the Noise!**
(divide and conquer)
- **Fix decision tree “overfitting” or “too close of fit”**
with supervised learning training set.
- Miss outliers if overfitting occurs
- *Bagging* smooths things out



TowardsDataScience.com



Tally: Six 1s and Three 0s
Prediction: 1

Neural Network (ML Classifier)

- Artificial Neural Networks (ANN) make ML/AI fast
- Deep Learning (Google's **Tensor Flow** leverages this)
- Image Classification with Tensor Flow
- Classify patterns & sequences, regression testing...
- Data Mining, Spam Filtering, Robotics
- **Supervised or Unsupervised Learning**
- **Autonomous Vehicles**
- Use for Predictive Modeling
- Used with **Speech Recognition**



Neural Network (ML Classifier): Pixels

Machine Learning

"how"

learning algorithms

Deep Learning

features

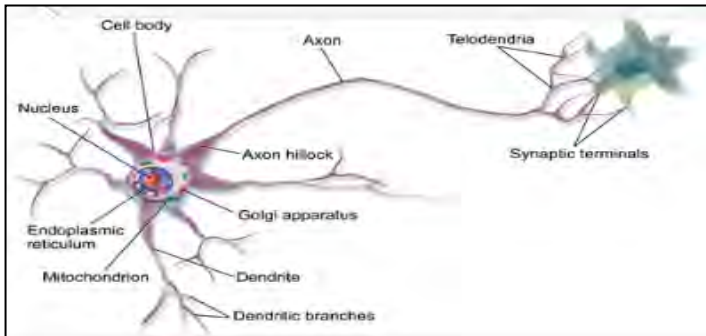
architectures

Foundation Models

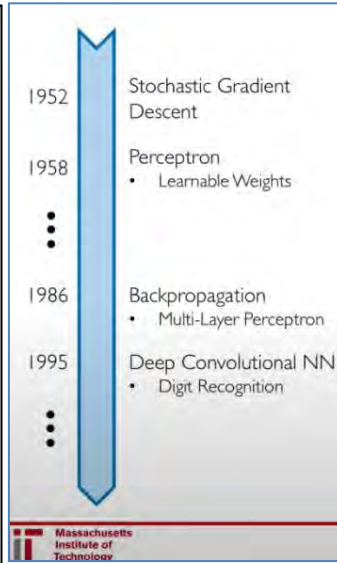
functionalities

models

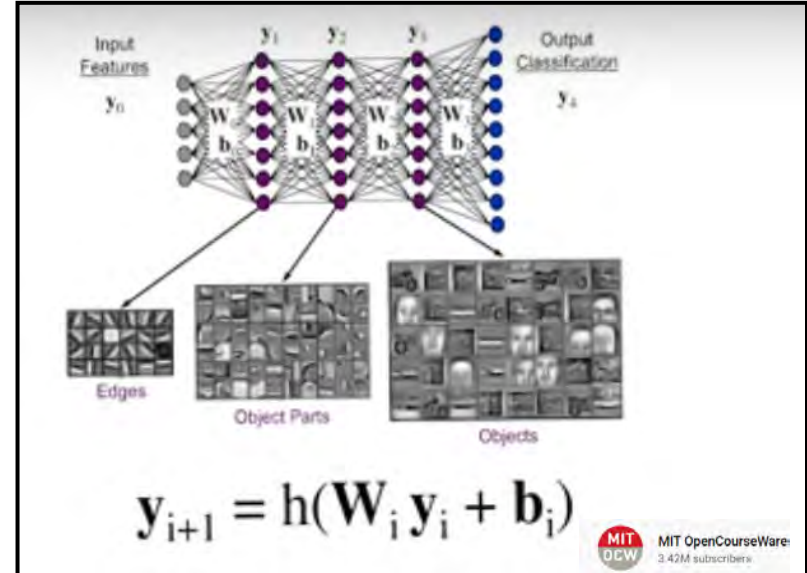
Biological Neural Network



A **neural circuit** is a population of neurons interconnected by synapses to carry out a specific function when activated. Neural circuits interconnect to one another to form large scale brain networks. Biological neural networks have inspired the design of artificial neural networks, but artificial neural networks

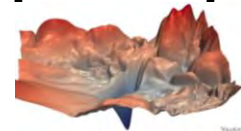


AI Neural Network



y = vector/image
W=weight (each line weight/strength)
b = thresholds/vector or bias
y times W is just Vector x Matrix
h = rectified linear unit (>0)
Set: #layers / W's / b's / #neurons / features

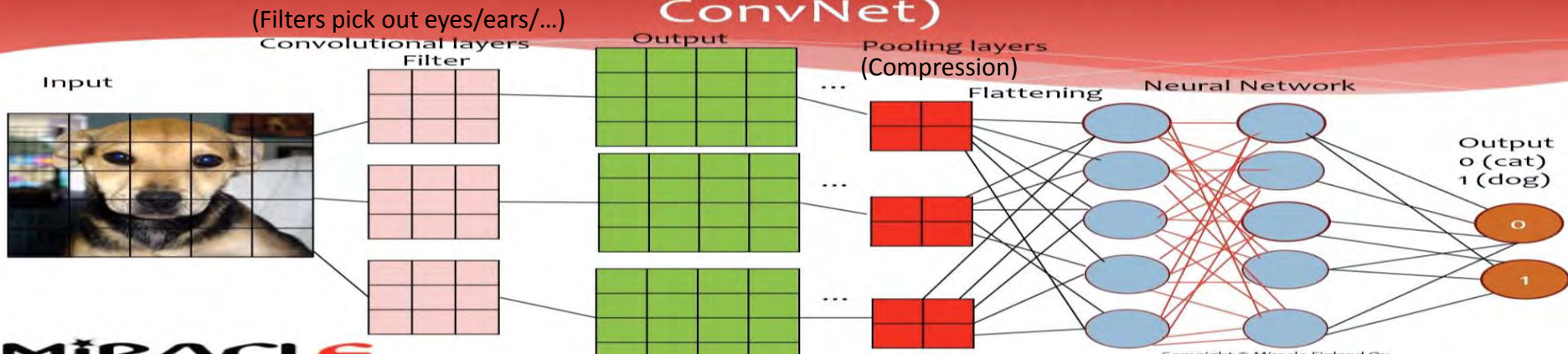
Use BP or Backpropagation to update weights to minimize loss; use gradient descent or stochastic gradient descent [differential].
(PyTorch autograd / Dropout so you don't overfit)



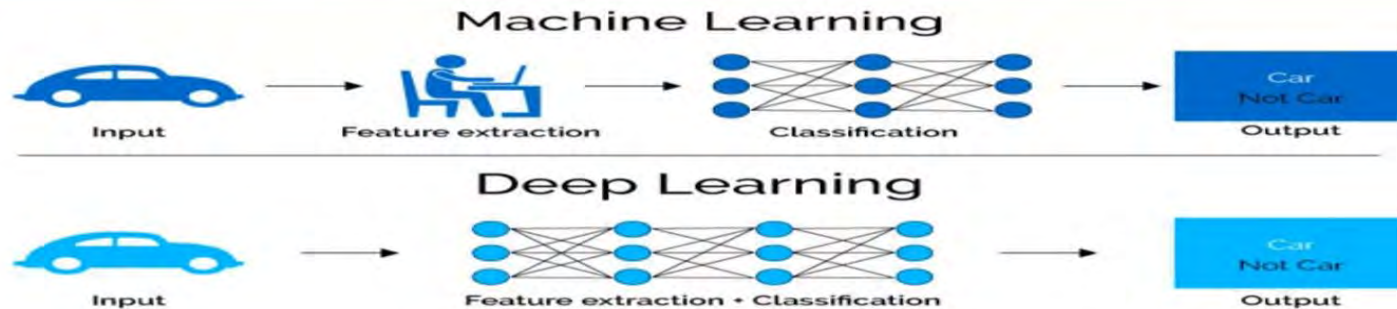
NN's also called ANN's /Artificial Neural Networks

Neural Network - CNN (ML Classifier)* *HeliFromFinland

Convolutional Neural Networks (CNN, or ConvNet)



What is deep learning?



Oracle Machine Learning Algorithms and Analytics in Oracle Database

CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models
e.g. Holt-Winters with trends,
seasonality, irregular time series

[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE



- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

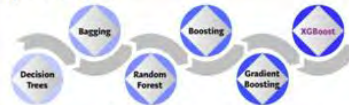
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



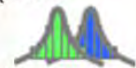
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test,
Pearson's, Chi-Sq, ANOVA, etc.



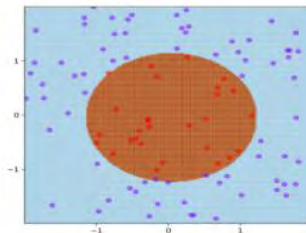
*Includes support for partitioned models,
integrated text mining,
automated data preparation*

** New in 21c*



One-Class SVM (ML Anomaly Detection)

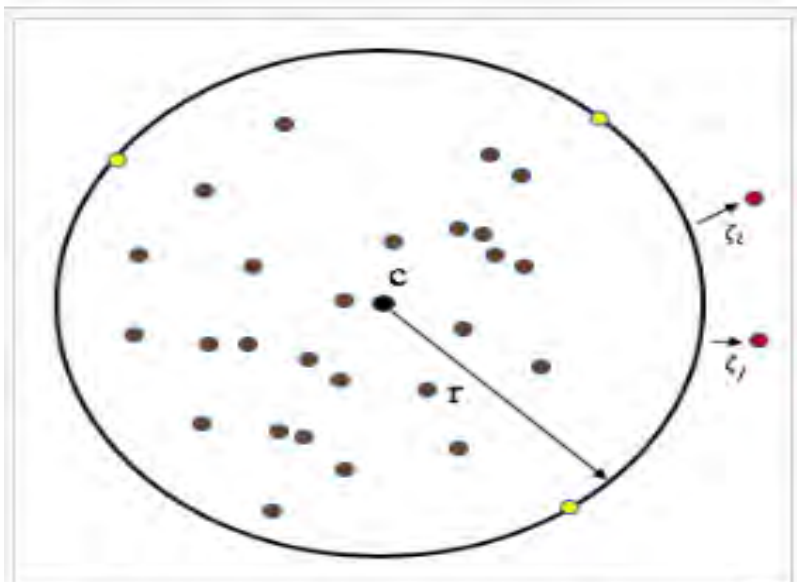
- Support Vector Machine (**SVM**)
- **One-Class** Classification (**OCC**)
- Used to Classify Data
- **Supervised Learning** using **Algorithms**
- **Identify specific objects of a class** based on supervised learning based on objects of that class.
- SVM maps training data into separated areas
- **Unsupervised Learning** does data **clustering**
- They analyze the data used for **regression**
- SVM as binary non-probabilistic linear classifier
- **Classify images, text, & even hand written text.**



Fraud detection
Use SVM OCC

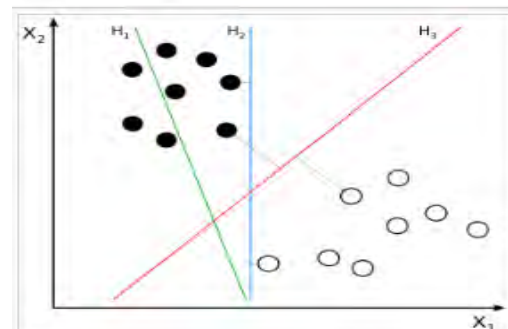
One-Class SVM (ML Anomaly Detection)*

FYI



The hypersphere containing the target data having center a and radius R . Objects on the boundary are support vectors, and two objects lie outside the boundary having slack greater than 0.

Support Vector Data Description (SVDD):
Find the smallest hypersphere containing all data points (use supervised training to get it)

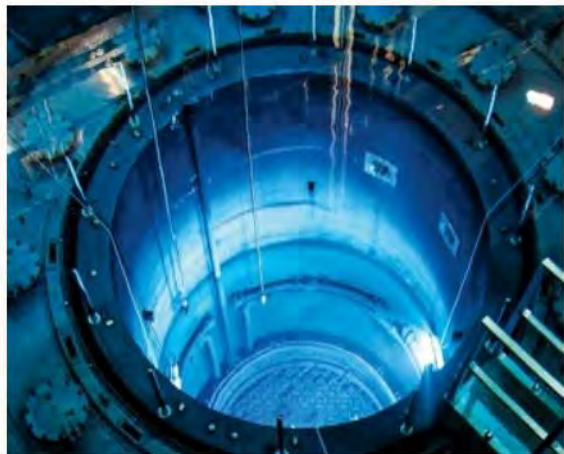


H_1 does not separate the classes.
 H_2 does, but only with a small margin.
 H_3 separates them with the maximal margin.

Linear SVM

Smart Anomaly Detection with MSET2 (SAM)

Created by nuclear scientists to analyze physical systems



All 90+ US nuclear reactors and most of the 450 reactors worldwide use MSET for prognostic analytics in safety systems



Oracle Exadata machines and Oracle Cloud use MSET for predictive maintenance, helping to deliver highest cloud uptime in the industry



Major airlines use MSET in jet engines for predictive health monitoring, benefiting uptime, maintenance activities, and supply chain costs

Oracle Machine Learning Algorithms and Analytics in Oracle Database

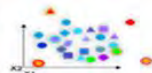
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models
e.g. Holt-Winters with trends,
seasonality, irregular time series

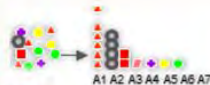
[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE



- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

ASSOCIATION RULES



- A priori

SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



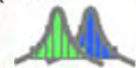
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



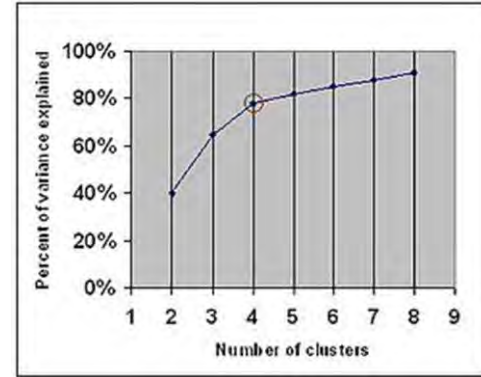
*Includes support for partitioned models,
integrated text mining,
automated data preparation*

** New in 21c*



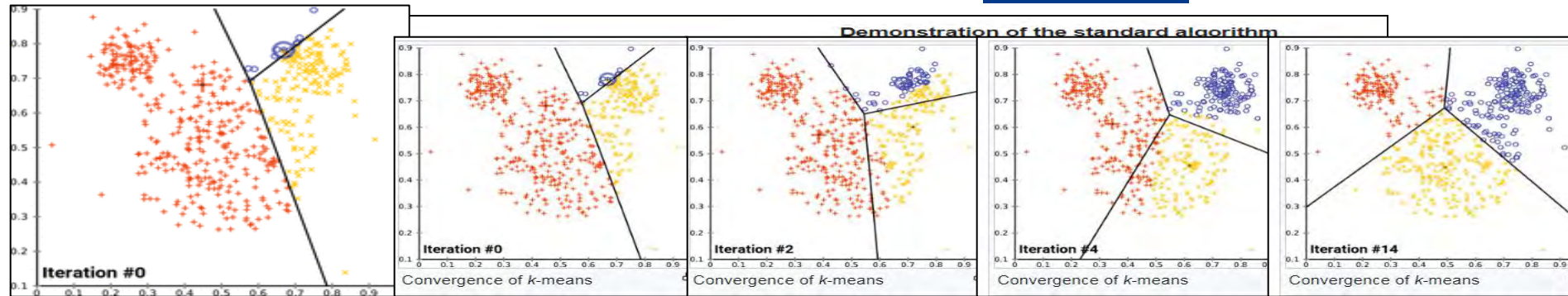
Hierarchical K-Means (ML Cluster)

- Use the **K-means (Lloyd's) Algorithm**
- Use **observations** of the Data Set
- Clusters **generated based on each observation with nearest mean & hierarchies of clusters.**



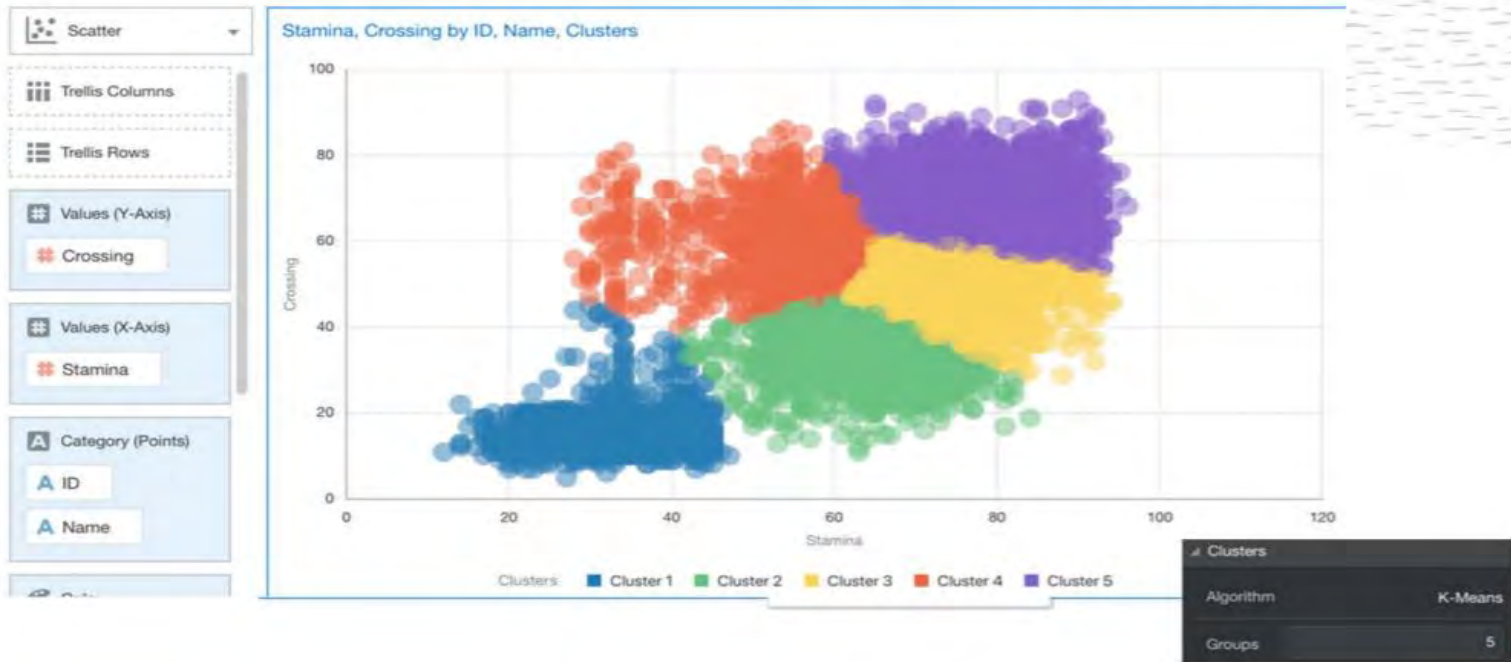
How Many? Elbow Method (4 in image)!

- **Centroid cluster is new mean - Repeat**



Oracle Analytics Cloud (OAC) to Cluster Data

Easy Models

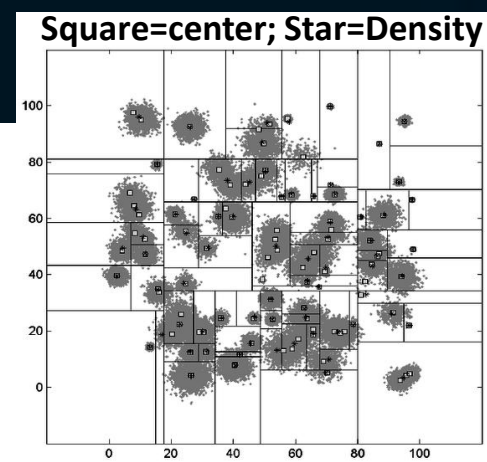


Oracle Analytic Summit 2020



Hierarchical O-Cluster (ML Cluster)

- **Orthogonal Partitioning Clustering**
- **Oracle-Proprietary clustering algorithm**
- **Algorithm operates recursively**
- **Creates dense areas in attribute space**
- **Hierarchical grid-based with Axis-parallel (orthogonal) partitions**
- *Sensitivity defines baseline density level (to separate peaks/valleys)*
- **Areas with peak density above baseline are clusters**
- **Clusters used to generate Bayesian probability**
- **Outliers can cause issue; pre-clip outliers**
- **K-Means is Distance Based; O-Cluster is Density Based!**



Oracle Machine Learning Algorithms and Analytics in Oracle Database

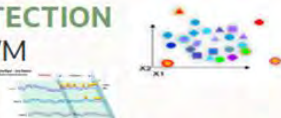
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES

- Forecasting - Exponential Smoothing
- Includes popular models e.g. Holt-Winters with trends, seasonality, irregular time series



[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE

- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI



ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

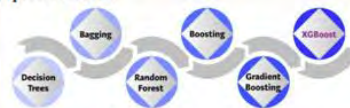
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



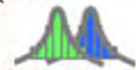
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



Includes support for partitioned models, integrated text mining, automated data preparation

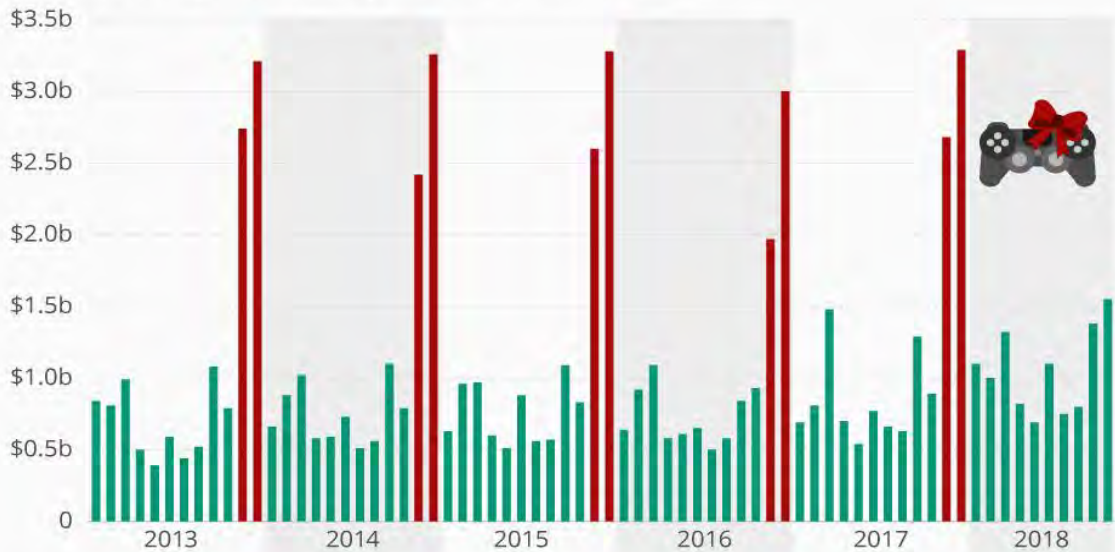
** New in 21c*



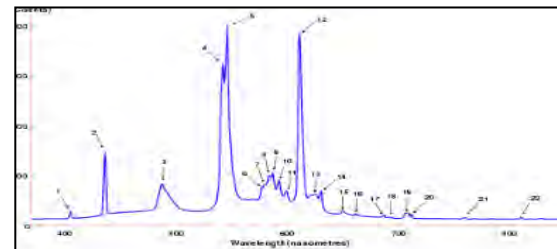
Seasonal, Irregular & Missing Data: Time Series Algorithm

Video Game Sales Are Extremely Seasonal

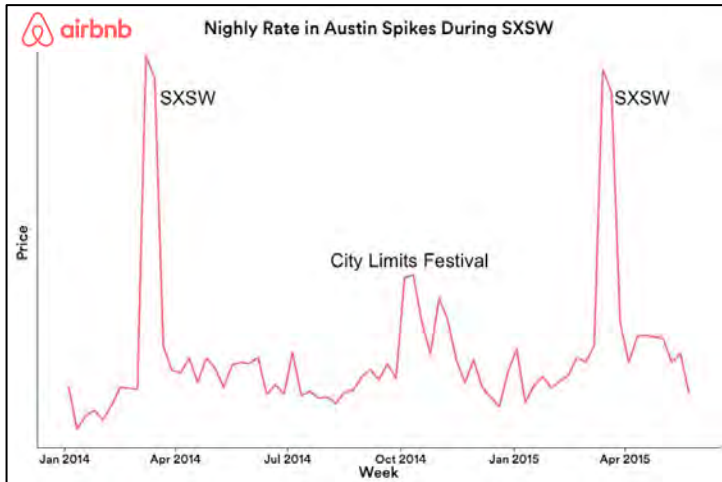
Monthly video games industry sales in the United States*



* incl. video games hardware, PC & video games software, accessories and game cards
@StatistaCharts Source: NPD Group

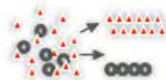


The power spectrum $S_{x,x}(f)$ of a time series $x(t)$ describes the distribution of power into frequency components composing that signal. According to Fourier analysis, any physical signal can be decomposed into a number of discrete frequencies, or a spectrum of frequencies over a continuous range.



Oracle Machine Learning Algorithms and Analytics in Oracle Database

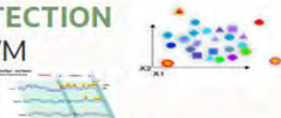
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models
e.g. Holt-Winters with trends,
seasonality, irregular time series

[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE



- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

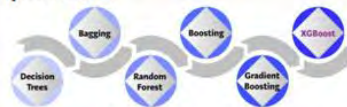
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



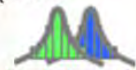
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



*Includes support for partitioned models,
integrated text mining,
automated data preparation*

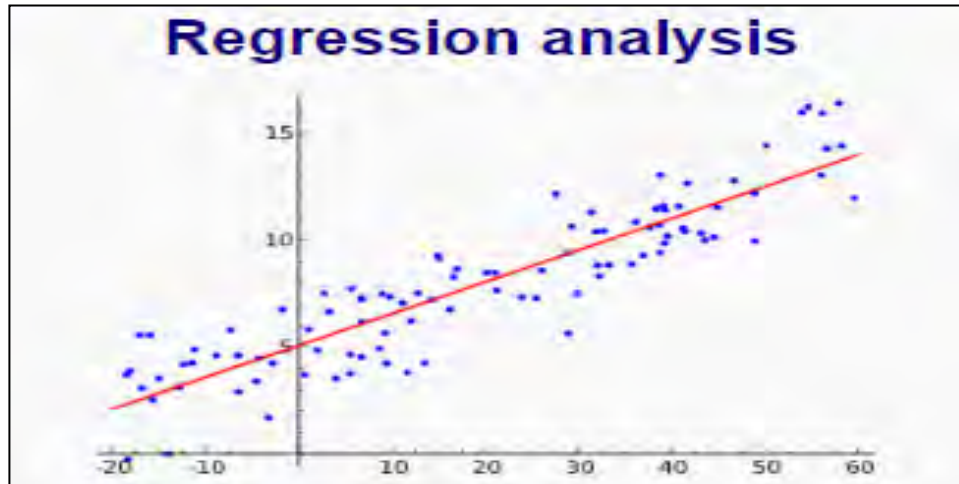
** New in 21c*



Linear Model (Regression)

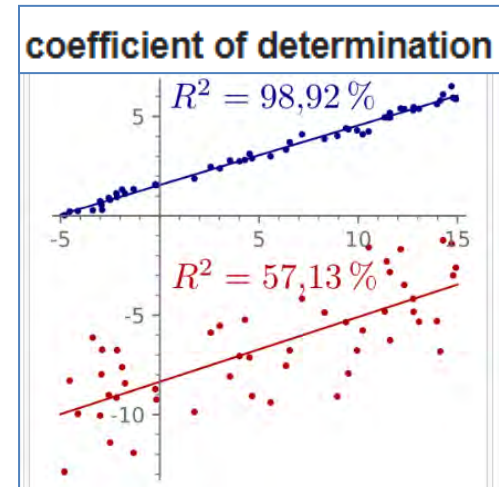
Linear Regression (below):

Using simple **Linear Regression** we find the line between the points. There is one independent and one dependent variable. This helps us get to predictive analytics.



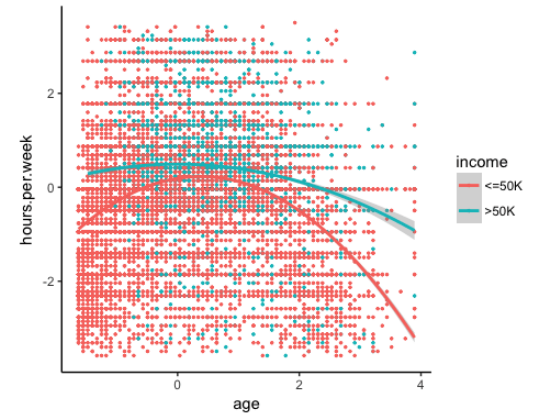
Linear Regression (above):

Unemployment goes down & GDP goes up



Generalized Linear Model (GLM)

- GLM: *Flexible* generalization of linear regression
- **GLM** allows output that **doesn't match normal linear regression** & *generalizes* it by a *link function* based on variance from predicted value.
- GLM for varying situations without Normal Distributions
- GLM assumes observations uncorrelated
- GLM assumes **random vs. normal effects**
- GLM mixed models allow random effects that can be correlated through some probability.



Oracle Machine Learning Algorithms and Analytics in Oracle Database

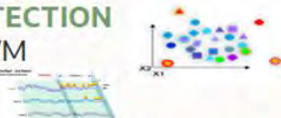
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models e.g. Holt-Winters with trends, seasonality, irregular time series

[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE

- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI



ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



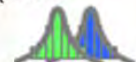
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



Includes support for partitioned models, integrated text mining, automated data preparation

** New in 21c*



Principal Component Analysis (Attribute Importance)

- PCA: A dimension-reduction tool **taking large set of attributes to small set (reduction)** with most important information .
- Use Feature Selection of Most Important Attributes
- PCA speeds up your ML - reducing dimension input
- “The **eigenvectors** and **eigenvalues** of a covariance (or correlation) matrix are the “core” of **PCA** and...
- The **eigenvectors** (principal components) determine the directions of the new feature space, and the **eigenvalues** determine their magnitude.”
- PCA is the simplest eigenvector based analysis (multivariate)

(see next slide for example & later slides for example on PCA Feature Extraction)

Living Doll

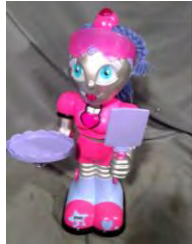


These dolls are spying on your kids, consumer groups say



By **Doug Criss**, CNN

Published 3:46 PM EST, Thu December 8, 2016



"Privacy advocates try to keep 'creepy,' 'eavesdropping' Hello Barbie from hitting shelves," Sarah Halzack, [Washington Post](#)

Oracle Machine Learning Algorithms and Analytics in Oracle Database

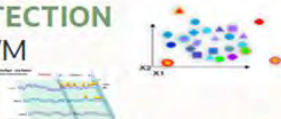
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models
e.g. Holt-Winters with trends,
seasonality, irregular time series

[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE



- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



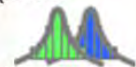
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



*Includes support for partitioned models,
integrated text mining,
automated data preparation*

** New in 21c*

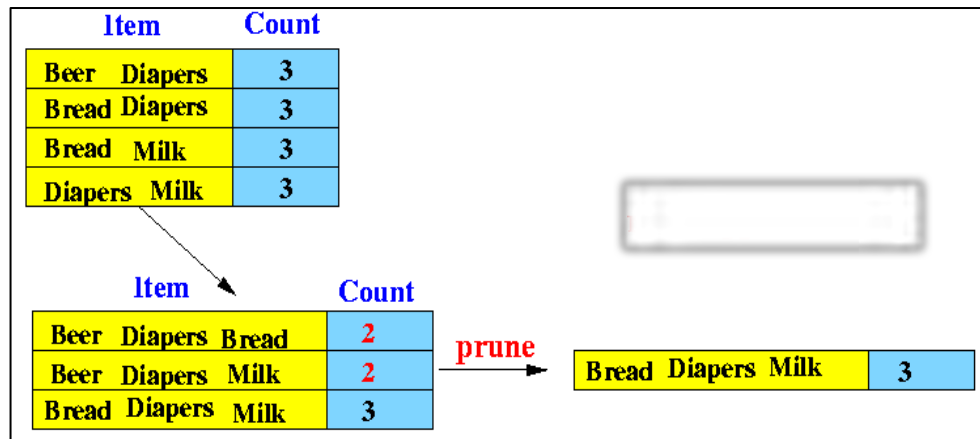
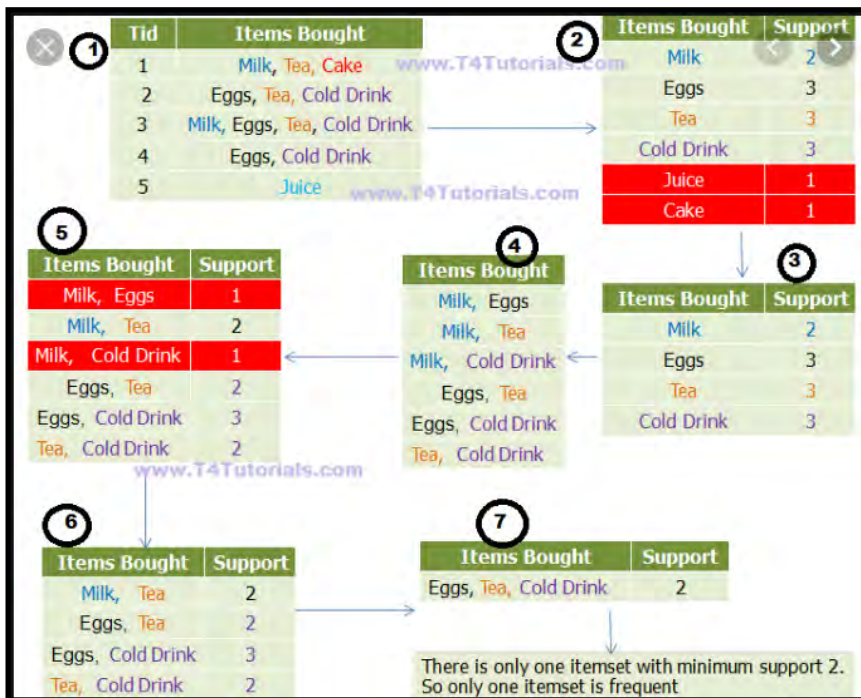


Apriori / Market Based (Association Rules)

- **Apriori Algorithm** is used for association rule learning databases. **Oracle Excels here!**
- Identifies frequently associated database items
- Frequent items determine Association Rules (trends)
- Helps with Market Basket analysis (purchase behavior of customers (retail) by using information to discount/promote items
- Also shows items bought together (cheese/crackers)
- The algorithm may scan database many times (downside)
- Finds too many matches (must limit this at times)



A Priori / Market Based (Association Rules)



T4Tutorials: Finding best Items Frequently Bought together.

Emory Univ. CS: Maybe we don't need the beer that close to the Bread, Diapers & Milk

Oracle Machine Learning Algorithms and Analytics in Oracle Database

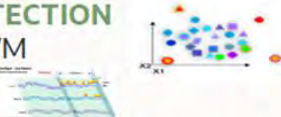
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models e.g. Holt-Winters with trends, seasonality, irregular time series

[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE



- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



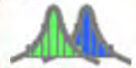
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.

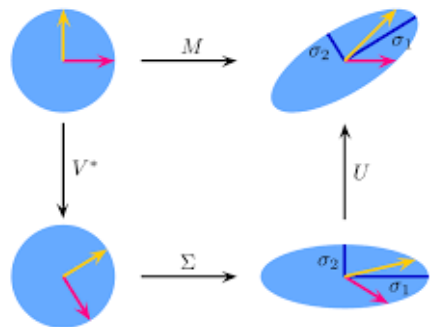


Includes support for partitioned models, integrated text mining, automated data preparation

** New in 21c*



Singular Value Decomposition (Feature Extraction)



$$M = U \cdot \Sigma \cdot V^*$$

Item x subject matrix (ISM)

	S1	S2	S3	S4	S5
dog	1	1	1	1	1
cat	1	1	0	1	0
cow	0	0	1	0	1
lion	0	0	1	1	0
tiger	1	1	0	0	1

I2tutorials.com

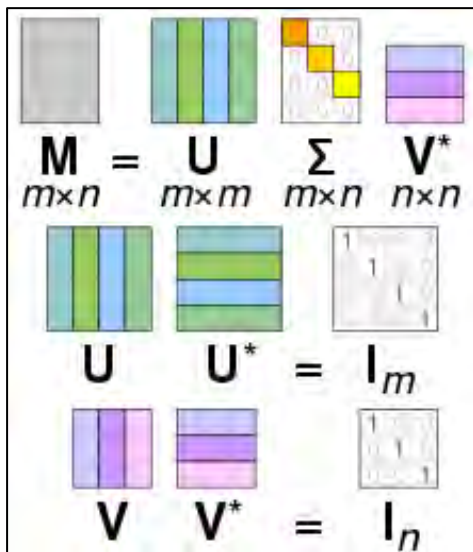
Singular decomposition analysis (SVD)

$$C_{m \times n} = U_{m \times r} \times \Sigma_{r \times r} \times V'_{r \times n}$$

Item vectors Singular values Subject vectors

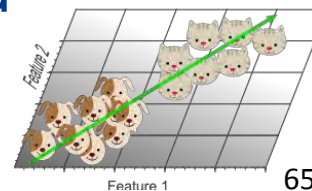
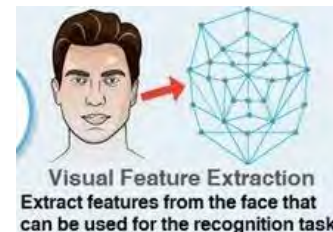
Reducing dimensions from r to k

$$\tilde{C}_{m \times n} = U_{m \times k} \times \Sigma_{k \times k} \times V'_{k \times n}$$



Principal Component Analysis (Feature Extraction)

- **PCA Feature extraction** reduces large data sets, by **building new set of attributes** making it faster to process
- **Present** the same information with Fewer Variables
- Ensure you **preserve the structure (variance)**
- Use Feature Extraction vs. **Feature Selection**
- PCA is a dimension-reduction tool **taking large set of variables to small set** with most important information.
- First component has largest variance and succeeding variables are orthogonal & highest variance.
- PCA uses orthogonal transformation to convert possibly correlated variables (numeric) into linear uncorrelated variables (principle components)
- **New Attributes** a Combination of Old Attributes



In His Image (Available Now - 100K+ Waiting List)



ETER9

The most daring
experience ever made
on the Internet.

Let your counterpart work for you. The very
first-ever intelligent virtual self of you.

DOUBLE YOURSELF

Currently, we have over 100,000 registered users who are on the waiting list for ETER9
brainchild: [dduplicata](#). Be a part of this remarkable phase too!



CBS AOL. BBC The Telegraph theguardian *Stm*



Artificial Intelligence

ETER9 is a social network that relies on Artificial Intelligence as a central element, and it's currently in the BETA stage. Even in your absence, the virtual beings will publish, comment and interact with you intelligently.



Counterpart

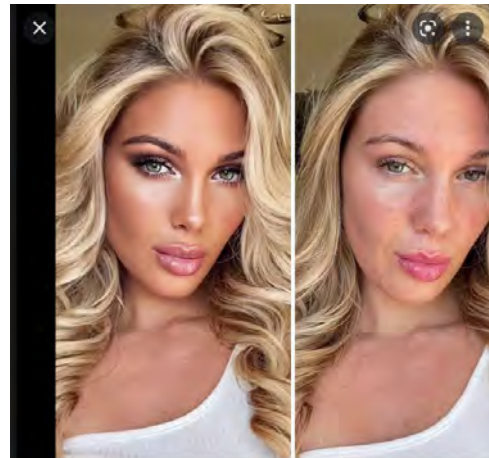
The Counterpart is your Virtual Self that will stay in the system and interact with the world just like you would if you were present. Your Counterpart will learn more with each action you take. The more you interact in the new social network, the more your Counterpart will learn!



Cyber Eternity

Eternizing is a way of keeping your thoughts and posts for all time. Are you curious? Come meet your Counterpart and become eternal. Challenge the impossible.

Number 12 Looks Like You (2020)



Bored Panda
TikTok Goes Viral By Exposing The Truth Behind Heavily-Edited Pictures

TikTok guidelines said not to promote posts by 'ugly' and poor users



Camouflage Makeup - Amazon.com



hide or disguise the presence of



The Dangers of social media filters Forbes

BEFORE

AFTER



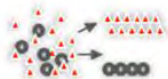
BUILD YOUR OWN BIRCHBOX STARTING AT \$15

#AgeChallenge



Oracle Machine Learning Algorithms and Analytics in Oracle Database

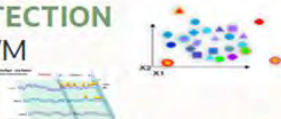
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models e.g. Holt-Winters with trends, seasonality, irregular time series

[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

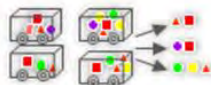
ATTRIBUTE IMPORTANCE



- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

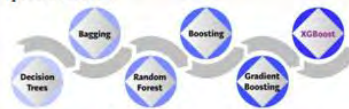
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



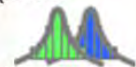
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



Includes support for partitioned models, integrated text mining, automated data preparation

** New in 21c*

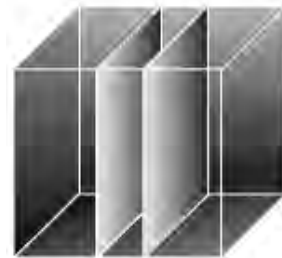


SQL Analytics (Windows / Patterns / Aggregates)*

*Oracle Docs



Product Mgr. View



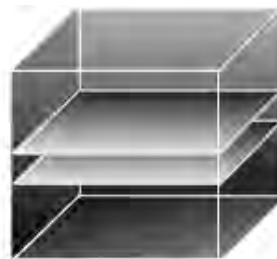
Financial Mgr. View



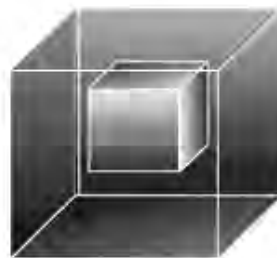
Market

SALES

Time



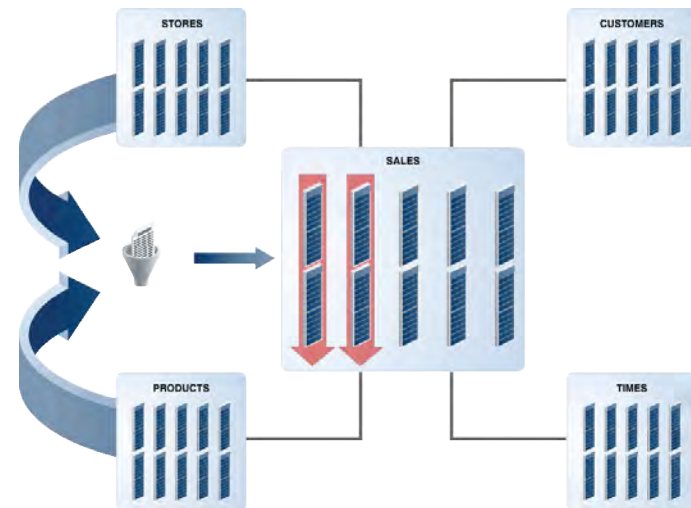
Regional Mgr. View



Ad Hoc View

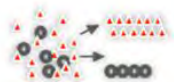
Cubes show many dimensions for various users of the system.

In-Memory will use Vector Group By Aggregation



Oracle Machine Learning Algorithms and Analytics in Oracle Database

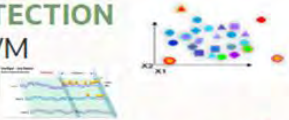
CLASSIFICATION



- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- *XGBoost**

ANOMALY DETECTION

- One-Class SVM
- *MSET-SPRT**



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



TIME SERIES



- Forecasting - Exponential Smoothing
- Includes popular models e.g. Holt-Winters with trends, seasonality, irregular time series

[OML Algorithm Cheat Sheet](#)
[Algorithm Documentation](#)

REGRESSION



- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- *XGBoost**

ATTRIBUTE IMPORTANCE



- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

ASSOCIATION RULES

- A priori



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates



FEATURE EXTRACTION

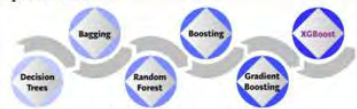
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

ROW IMPORTANCE

- CUR Decomposition

RANKING

- *XGBoost**



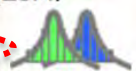
TEXT MINING SUPPORT

- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)



STATISTICAL FUNCTIONS

- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



Includes support for partitioned models, integrated text mining, automated data preparation

** New in 21c*



Statistical Functions in Oracle (Partial List)*

FYI

Numeric Functions

ABS
ACOS
ASIN
ATAN
ATAN2
BITAND
BITAND
CEIL
COS
COSH
EXP
FLOOR
LN
LOG
MOD
NANVL
POWER
REMAINDER
ROUND (number)
SIGN
SIN
SINH
SQRT
TAN
TANH
TRUNC (number)
WIDTH_BUCKET

Analytic Functions

AVG *
CLUSTER_DETAILS
CLUSTER_DISTANCE
CLUSTER_ID
CLUSTER_PROBABILITY
CLUSTER_SET
CORR *
COUNT *
COVAR_POP *
COVAR_SAMP *
CUME_DIST
DENSE_RANK
FEATURE_DETAILS
FEATURE_ID
FEATURE_SET
FEATURE_VALUE
FIRST
FIRST_VALUE *
LAG
LAST
LAST_VALUE *
LEAD
LISTAGG
MAX *
MIN *
NTH_VALUE *
NTILE
PERCENT_RANK
PERCENTILE_CONT
PERCENTILE_DISC
PREDICTION
PREDICTION_COST
PREDICTION_DETAILS
PREDICTION_PROBABILITY
PREDICTION_SET
RANK
RATIO_TO_REPORT
REGR_ (Linear Regression)
ROW_NUMBER
STDDEV *
STDDEV_POP *
STDDEV_SAMP *
SUM *
VAR_POP *
VAR_SAMP *
VARIANCE *

Aggregate Functions

APPROX_COUNT
APPROX_COUNT_DISTINCT
APPROX_COUNT_DISTINCT_AGG
APPROX_COUNT_DISTINCT_DETAIL
APPROX_MEDIAN
APPROX_PERCENTILE
APPROX_PERCENTILE_AGG
APPROX_PERCENTILE_DETAIL
APPROX_RANK
APPROX_SUM
AVG
COLLECT
CORR
CORR *
COUNT
COVAR_POP
COVAR_SAMP
CUME_DIST
DENSE_RANK
FIRST
GROUP_ID
GROUPING
GROUPING_ID
JSON_ARRAYAGG
JSON_OBJECTAGG
LAST
LISTAGG
MAX
MEDIAN
MIN
PERCENT_RANK
PERCENTILE_CONT
PERCENTILE_DISC
RANK
REGR_ (Linear Regression) Functions
STATS_BINOMIAL_TEST
STATS_CROSSTAB
STATS_F_TEST
STATS_KS_TEST
STATS_MODE
STATS_MW_TEST
STATS_ONE_WAY_ANOVA
STATS_T_TEST_*
STATS_WSR_TEST
STDDEV
STDDEV_POP
STDDEV_SAMP
SUM
SYS_OP_ZONE_ID
SYS_XMLAGG
TO_APPROX_COUNT_DISTINCT
TO_APPROX_PERCENTILE
VAR_POP
VAR_SAMP
VARIANCE
XMLAGG

Data Mining Functions

CLUSTER_DETAILS
CLUSTER_DISTANCE
CLUSTER_ID
CLUSTER_PROBABILITY
CLUSTER_SET
FEATURE_COMPARE
FEATURE_DETAILS
FEATURE_ID
FEATURE_SET
FEATURE_VALUE
ORA_DM_PARTITION_NAME
PREDICTION
PREDICTION_BOUNDS
PREDICTION_COST
PREDICTION_DETAILS
PREDICTION_PROBABILITY
PREDICTION_SET

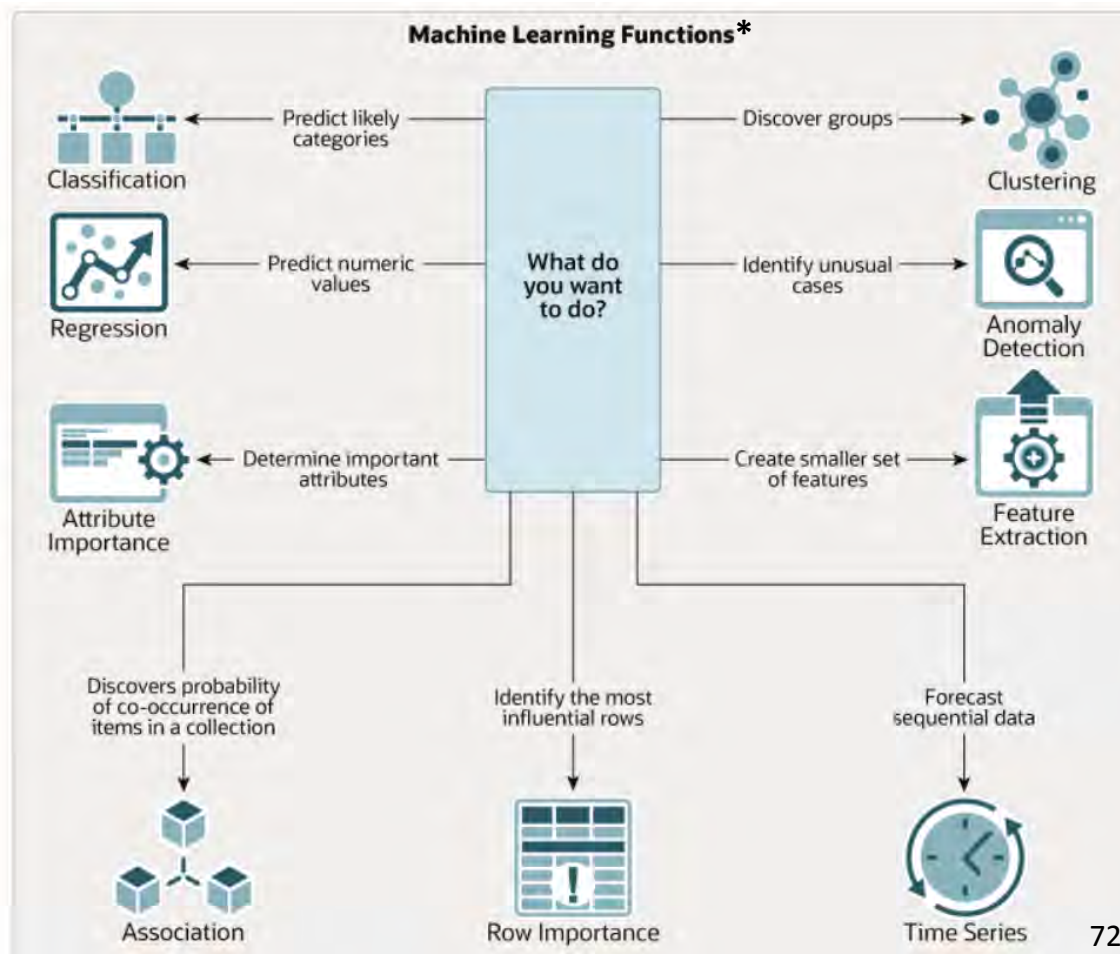
Also: T-Test, F-Test, Chi-Squared, ANOVA
(Analysis of Variance), Pearson's Test...

* Oracle Database SQL Language Reference

Machine Learning Functions* - Oracle Docs.

- **First:** Clear *Business Problem* to Solve
 - **Second:** *Function* to Perform
 - **Third:** *Algorithm* to use
-
- **First:** Build/*Train the Model* when you build it using about 60% of data.
 - **Second:** *Test/Score Model* for accuracy/precision using about 40% of data.


Compare Algorithms!



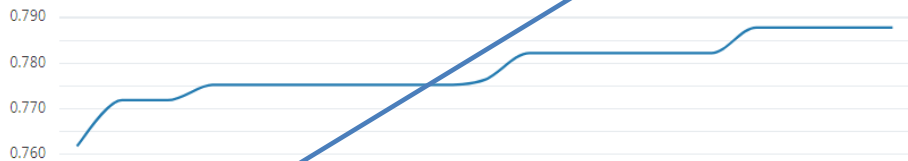
AutoML is here for Autonomous Database

<- Experiments

AutoML_Rich

Experiment Settings  Edit

Accuracy



Leader Board

Deploy	Create Notebook	Metrics
Algorithm	Model Name	Accuracy
Support Vector Machine (Gaussian)	svmg_09cd02ee4e	0.7878
Random Forest	rf_df53fb31a8	0.7821
Generalized Linear Model	glm_74cbb2b7ac	0.7821
Generalized Linear Model (Ridge Regres...	glmr_df24288679	0.7821
Support Vector Machine (Linear)	svml_a031bb0ae1	0.7431


Leader Board

Deploy	Create Notebook	Metrics
Algorithm	Model Name	Accuracy
Support Vector Machine (Gaussian)	svmg_09cd02ee4e	0.7878
Random Forest	rf_df53fb31a8	0.7821
Generalized Linear Model	glm_74cbb2b7ac	0.7821
Generalized Linear Model (Ridge Regres...	glmr_df24288679	0.7821
Support Vector Machine (Linear)	svml_a031bb0ae1	0.7431

Create Notebook

Create a notebook based on selected model and this experiment's settings. Use a generated notebook to further tune your approach using Python.

Notebook Name:



AutoML_Classify



Build Data

```
%python
import oml

columns = 'CUST_ID', 'CUST_CREDIT_LIMIT', 'CUST_GENDER', 'CUST_INCOME_LEVEL', 'CUST_MARITAL_STATUS', 'CUST_YEAR_OF_BIRTH', 'EDUCATION', 'HOUSEHOLD_SIZE', 'OCCUPATION', 'YRS_RESIDENCE',
         'Y_BOX_GAMES', 'AFFINITY_CARD'
schema='ML_USER'
table='CUSTOMERS360'

column = ','.join(columns)
query = 'SELECT ' + column + ' FROM ' + schema + '.' + table

data_build = oml.sync(query=query)
% show(data_build)
```

Create Train Data

READY

```
%python
import oml

X_train = data_build[:,['CUST_ID', 'CUST_CREDIT_LIMIT', 'CUST_GENDER', 'CUST_INCOME_LEVEL', 'CUST_MARITAL_STATUS', 'CUST_YEAR_OF_BIRTH', 'EDUCATION', 'HOUSEHOLD_SIZE', 'OCCUPATION', 'YRS_RESIDENCE',
                        'Y_BOX_GAMES']]
v_train = data_build[:, 'AFFINITY_CARD']
```

Build 'SUPPORT_VECTOR_MACHINES' Model

READY

```
%python
import oml

svm_settings = {
    'ODMS_SAMPLE_SIZE': '1294', 'SVMS_COMPLEXITY_FACTOR': '66.70000000000002', 'SVMS_KERNEL_FUNCTION': 'SVMS_GAUSSIAN', 'SVMS_STD_DEV': '2.23606797749979', 'CLAS_WEIGHTS_BALANCED': 'OFF',
    'SVMS_NUM_PIVOTS': '200', 'ODMS_DETAILS': 'ODMS_DISABLE', 'ODMS_SAMPLING': 'ODMS_SAMPLING_ENABLE', 'ALGO_NAME': 'ALGO_SUPPORT_VECTOR_MACHINES'
}

svm_mod = oml.svm(**svm_settings)

svm_mod = svm_mod.fit(X_train, v_train)
```

Quick Actions



AutoML

Create and run AutoML Experiments



Scratchpad

Run Scratchpad



Notebooks

The place for data discovery and analytics



Jobs

Schedule notebooks to run at certain times



Examples

Check out some examples

Recent Activities

today



Rich Niemiec created AutoML_Classify notebook in ML_USER Project [ML_USER Workspace]

3/19/21 4:27 PM



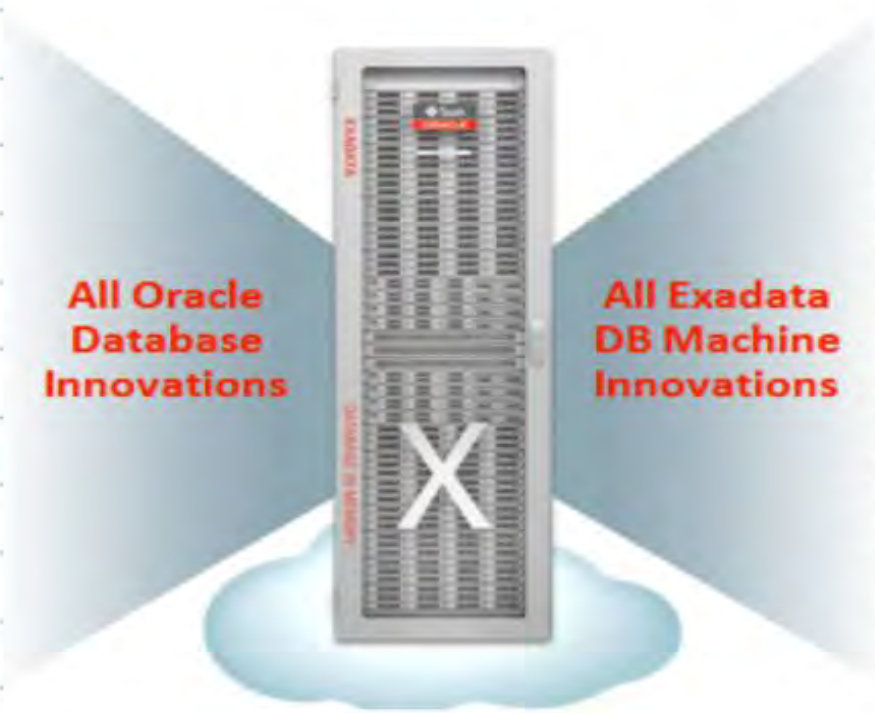
Time Enough at Last for ML with AutoML












Exadata Cloud Machine: ALL Features (X10M is Out - 1T/sec!) (Oracle's Juan Loaiza presentation on Exadata Cloud; X10M has > 2P Flash)

Exadata Cloud: Compatible, Scalable, Available, Secure
 Decades of Database Innovation Proven at Millions of Mission-Critical Deployments

	Multitenant
	In-Memory DB
	Real Application Clusters
	Active Data Guard
	Partitioning
	Advanced Compression
	Advanced Security, Label Security, DB Vault
	Real Application Testing
	Advanced Analytics, Spatial and Graph
	Management Packs for Oracle Database



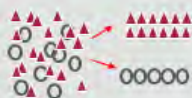
Offload SQL to Storage	
InfiniBand Fabric	
Smart Flash Cache, Log	PCI Flash
Storage Indexes	
Columnar Flash Cache	
Hybrid Columnar Compression	
I/O Resource Management	
Network Resource Management	
In-Memory Fault Tolerance	
Exafusion Direct-to-Wire Protocol	

Oracle Machine Learning Algorithms



CLASSIFICATION

- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine
- Explicit Semantic Analysis
- XGBoost*



CLUSTERING

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)



ANOMALY DETECTION

- One-Class SVM
- MSET-SPRT*



TIME SERIES

- Forecasting - Exponential Smoothing
- Includes popular models
e.g. Holt-Winters with trends,
seasonality, irregularity, missing data



Includes support for Partitioned Models, Transactional data and aggregations, Unstructured data, Geo-spatial data, Graph data. etc,

REGRESSION

- Linear Model
- Generalized Linear Model
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- XGBoost*



ATTRIBUTE IMPORTANCE

- Minimum Description Length
- Principal Comp Analysis (PCA)
- Unsupervised Pair-wise KL Div
- CUR decomposition for row & AI



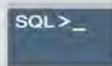
ASSOCIATION RULES

- A priori/ market basket



PREDICTIVE QUERIES

- Predict, cluster, detect, features



SQL ANALYTICS

- SQL Windows
- SQL Patterns
- SQL Aggregates

FEATURE EXTRACTION

- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

TEXT MINING SUPPORT

- Algorithms support text
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA) for document similarity



STATISTICAL FUNCTIONS

- Basic statistics: min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.



R & PYTHON ** Coming soon*

- Third-party R & Python Packages through Embedded Execution
- Spark MLlib algorithm integration



MODEL DEPLOYMENT & MONITORING

- SQL—1st Class Objects
- Oracle RESTful API (ORDS)
- OML Web Services (for Apps)



** New in 21^c*



#1 AI / ML Specialist (Engineer)

LinkedIn

2020
Emerging
Jobs Report

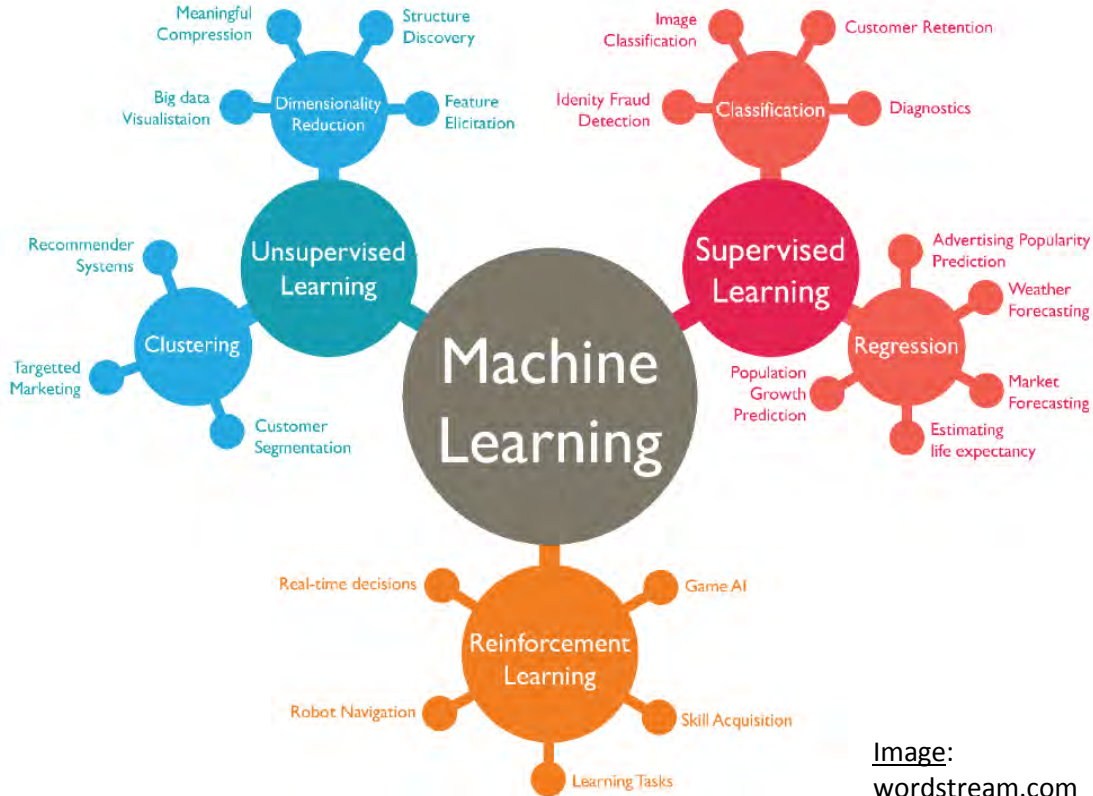
- 74% annual growth
- Skills
 - Machine Learning
 - Deep Learning
 - TensorFlow
 - Python
 - Natural Language Processing
- Industries
 - Computer Software
 - Information Technology
 - Higher Education
 - Consumer Electronics
- Take the data scientists' code and making it more effective and scalable.
- Introduce various programming rules and good practices.
- Join the results from potentially unrelated tasks to enhance the models performance even more.
- Glue all the "data science" and "software" parts together.

ML & Business Apps

ORACLE®
SALES CLOUD

ORACLE®
RETAIL CLOUD

ORACLE®
CONFIGURE, PRICE,
AND QUOTE CLOUD



ORACLE®
E-BUSINESS SUITE

ORACLE®
IDENTITY MANAGEMENT

ORACLE®
FINANCIAL SERVICES

ORACLE®
HUMAN CAPITAL
MANAGEMENT
CLOUD

Image:
wordstream.com

Applications - AI Powered; Analytics & ML

E-Business Analytics Accelerator: Pre-built solutions to speed insights



Oracle
E-Business
Suite



App and Data
Integration



Autonomous
Data Warehouse



Analytics &
Visualization

Generate
continuous
business outcomes
with your data



100s of Pre-built Dashboards

Outcomes	
Financials	<ul style="list-style-type: none">• Profitability drivers• Improve use of working capital• Control business expenditures
Supply Chain	<ul style="list-style-type: none">• Improve efficiency• Reduce costs• Ensure customer satisfaction
HCM	<ul style="list-style-type: none">• Complete view of employees• Identify best resources• Identify attrition factors and predict risk• Cross-domain reporting

Oracle GenAI

Oracle Database Vision

With Generative AI (LLM)

Make it easy to **generate** and run modern apps and analytics for all use cases at any scale



Oracle Database 23c

Sample of Marquee Features for DBAs, Analysts, & Developers

Now **GA** on OCI Base Database Service

NEW IN
23c

DBA



Real-time SQL Plan Management
SQL Firewall
True Cache

ANALYST



Automatic Materialized Views
Improved ML Algorithms

DEVELOPER



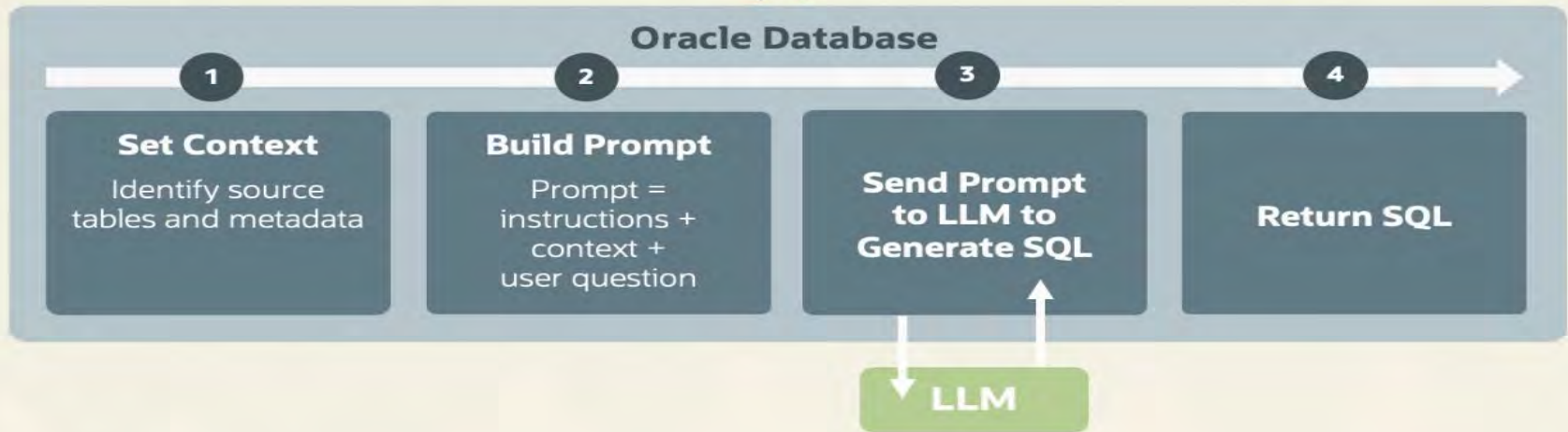
JSON Duality Views
Property Graph Views
JavaScript Stored Procedures
SQL Domains
AI Vector Search* (coming soon)

* Sign up to preview AI Vector Search



SQL Generation from Natural Language *using LLM*

→ “Give me the average salary of employees in each department”



Think of it as an Assistant (60-70%)



What's next from
Oracle Data Management

Example of a *Prompt*

Instructions: You are an Oracle SQL expert. Given an input question, first create a syntactically correct Oracle SQL query to run. You must query only the columns that are needed to answer the question.

Pay attention to use only the column names you can see in the tables below. Be careful to not query for columns that do not exist. Also, pay attention to which column is in which table.

Use the following format:

Question: Question here

SQL: Generated SQL query

Context: Only use the following tables and columns

Table: HR.DEPARTMENTS, **Columns:** DEPARTMENT_ID, DEPARTMENT_NAME, MANAGER_ID, LOCATION_ID

Table: HR.EMPLOYEES, **Columns:** EMPLOYEE_ID, FIRST_NAME, LAST_NAME, SALARY, MANAGER_ID, DEPARTMENT_ID

Primary keys: HR.DEPARTMENTS.DEPARTMENT_ID, HR.EMPLOYEES.EMPLOYEE_ID

Foreign keys: HR.EMPLOYEES.DEPARTMENT_ID -> HR.DEPARTMENTS.DEPARTMENT_ID, HR.DEPARTMENTS.MANAGER_ID -> HR.EMPLOYEES.EMPLOYEE_ID

Question: Give me the average salary of employees in each department ←

SQL:

Think of it as an Assistant (60-70%)

The screenshot displays the Oracle APEX SQL Workshop interface. At the top, the navigation bar includes 'APEX', 'App Builder', 'SQL Workshop', 'Team Development', and 'Gallery'. A search bar and user profile icons are on the right. Below the navigation bar, a green banner indicates 'Settings saved'. The main workspace is divided into two panes: 'Quick SQL' on the left and 'SQL' on the right. The 'Quick SQL' pane shows a table definition for 'departments' and 'employees'. The 'SQL' pane shows the corresponding SQL script for creating these tables. A 'Settings' menu is highlighted in the top toolbar. A promotional banner for 'What's next from Oracle Data Management' is visible on the right side of the workspace.

Quick SQL

```
1 departments
2 name
3 function /check IT, Sales, Accounting, Finance
4 employees
5 name /nn
6 email /lower
7 location
8 cost center num
9 date hired
10
```

SQL

```
1 -- create tables
2
3 create table hr_departments (
4     id                number generated by default on null as identity
5     |                | constraint hr_departments_id_pk primary key,
6     name              varchar2(255 char),
7     function          varchar2(4000 char) constraint hr_departments_function_ck
8     |                | check (function in ('IT','Sales','Accounting','Finance')),
9     row_version       integer not null,
10    created            date not null,
11    created_by         varchar2(255 char) not null,
12    updated            date not null,
13    updated_by        varchar2(255 char) not null
14 );
15
16
17 create table hr_employees (
18     id                number generated by default on null as identity
19     |                | constraint hr_employees_id_pk primary key,
20     department_id     number
21     |                | constraint hr_employees_department_id_fk
22     |                | references hr_departments,
23     name              varchar2(255 char) not null,
24     email             varchar2(255 char),
25     location          varchar2(4000 char),
26     cost_center       number,
27     date_hired        date,
28     row_version       integer not null,
```

What's next from Oracle Data Management

Review and Run

In whole example: He wrote 47 lines & it built 1200

APEX App Builder SQL Workshop Team Development Gallery

Utilities Quick SQL

Quick SQL Help Settings SQL Diagram

What's next from Oracle Data Management Review and Run

```
1 jobs
2   job_id          vc10 /pk
3   job_title       vc35 /nn
4   min_salary      num
5   max_salary      num
6
7 regions
8   region_id       num /nn /pk
9   region_name     vc25
10
11 countries
12  country_id      /nn /pk
13  country_name    vc40
14  region_id       num /fk regions
15
16 locations
17  location_id     num /nn /pk
18  street_address  vc40
19  postal_code     vc12
20  city            vc30 /nn
21  state_province  vc25
22  country_id      /fk countries
23
24 departments
25  department_id   num /nn /pk
26  department_name vc30 /nn
27  manager_id      num /fk employees
28  location_id     num /fk locations
29
30 employees
```

HR_JOB_HISTORY

ID	NUMBER
EMPLOYEE_ID	NUMBER
START_DATE	DATE
END_DATE	DATE
JOB_ID	INTEGER
DEPARTMENT_ID	NUMBER
ROW_VERSION	INTEGER
CREATED	DATE
CREATED_BY	VARCHAR2 (255 CHAR)
UPDATED	DATE
UPDATED_BY	VARCHAR2 (255 CHAR)

HR_JOBS

JOB_ID	INTEGER
JOB_TITLE	VARCHAR2 (35 CHAR)
MIN_SALARY	NUMBER
MAX_SALARY	NUMBER
ROW_VERSION	INTEGER
CREATED	DATE
CREATED_BY	VARCHAR2 (255 CHAR)
UPDATED	DATE
UPDATED_BY	VARCHAR2 (255 CHAR)

HR_EMPLOYEES

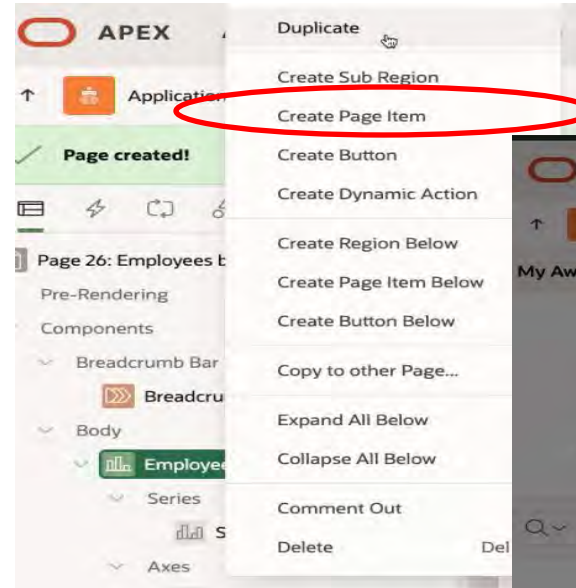
EMPLOYEE_ID	NUMBER
FIRST_NAME	VARCHAR2 (20 CHAR)
LAST_NAME	VARCHAR2 (25 CHAR)
EMAIL	VARCHAR2 (25 CHAR)
PHONE_NUMBER	NUMBER
HIRE_DATE	DATE
JOB_ID	INTEGER
SALARY	NUMBER
COMMISSION_PCT	NUMBER
DEPARTMENT_ID	NUMBER
ROW_VERSION	INTEGER
CREATED	DATE
CREATED_BY	VARCHAR2 (255 CHAR)
UPDATED	DATE
UPDATED_BY	VARCHAR2 (255 CHAR)

HR_DEPARTMENTS

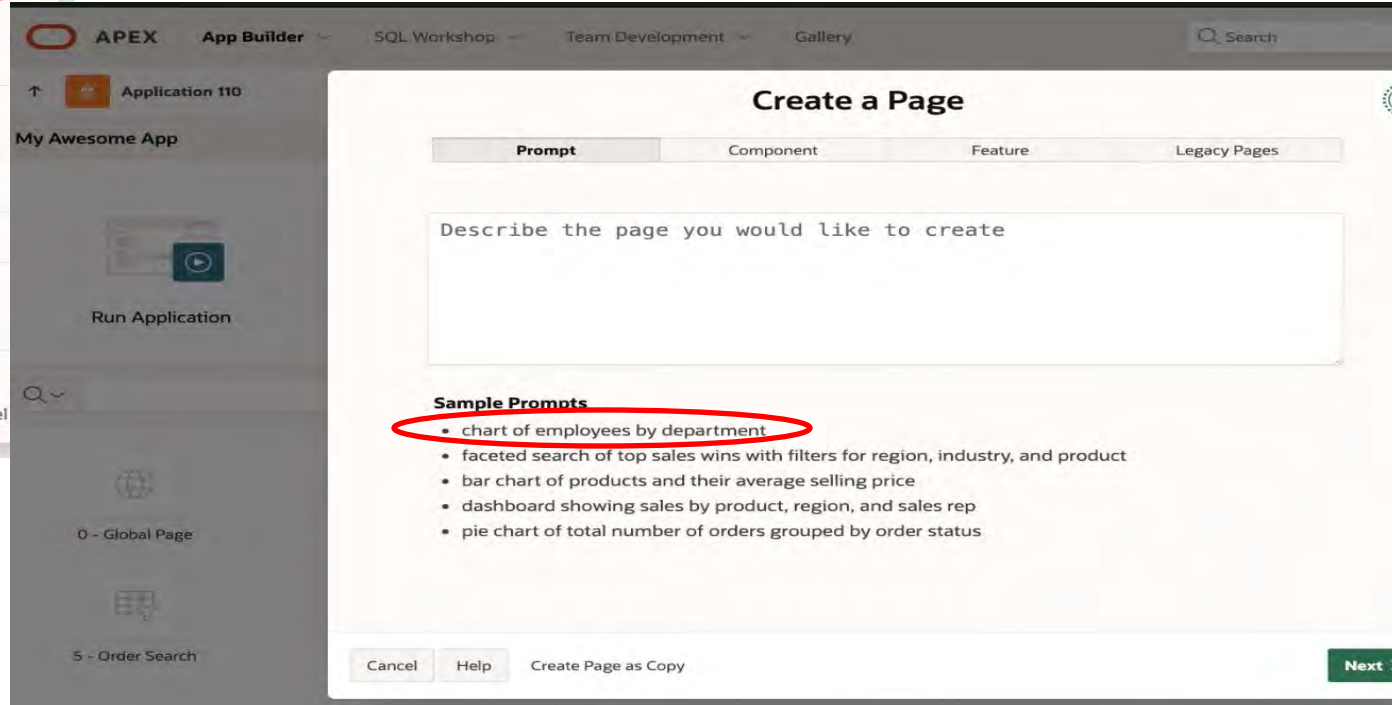
DEPARTMENT_ID	NUMBER
DEPARTMENT_NAME	VARCHAR2 (30 CHAR)
MANAGER_ID	NUMBER
LOCATION_ID	NUMBER
ROW_VERSION	INTEGER
CREATED	DATE
CREATED_BY	VARCHAR2 (255 CHAR)
UPDATED	DATE
UPDATED_BY	VARCHAR2 (255 CHAR)

HR_LOCATIONS

APEX Development Speed- GenAI



A screenshot of the APEX development environment. A context menu is open over an application item, listing various actions such as Duplicate, Create Sub Region, Create Page Item, Create Button, Create Dynamic Action, Create Region Below, Create Page Item Below, Create Button Below, Copy to other Page..., Expand All Below, Collapse All Below, Comment Out, and Delete. The 'Create Page Item' option is circled in red.



A screenshot of the 'Create a Page' dialog in the APEX development environment. The dialog has tabs for 'Prompt', 'Component', 'Feature', and 'Legacy Pages'. The 'Prompt' tab is selected, showing a text input field with the placeholder text 'Describe the page you would like to create'. Below the input field, there is a section titled 'Sample Prompts' with a list of suggestions. The first item, 'chart of employees by department', is circled in red.

Create a Page

Prompt Component Feature Legacy Pages

Describe the page you would like to create

Sample Prompts

- chart of employees by department
- faceted search of top sales wins with filters for region, industry, and product
- bar chart of products and their average selling price
- dashboard showing sales by product, region, and sales rep
- pie chart of total number of orders grouped by order status

Cancel Help Create Page as Copy Next

APEX Development Speed- GenAI



What's next from
Oracle Data Management

APEX App Builder SQL Workshop Team Development Gallery

Application 110

My Awesome App

Run Application

0 - Global Page

5 - Order Search

Create a Page

Prompt

Component

Feature

Legacy Pages

chart of employees by department

Sample Prompts

- chart of employees by department
- faceted search of top sales wins with filters for region, industry, and product
- bar chart of products and their average selling price
- dashboard showing sales by product, region, and sales rep
- pie chart of total number of orders grouped by order status

Cancel

Help

Create Page as Copy

Next >

APEX Development Speed- GenAI



What's next from
Oracle Data Management

Create Page Wizard

Title

Employees by Department

SQL

```
1 select department_name as chart_label,  
2     count(employee_id) as chart_value  
3 from oehr_employees  
4 join oehr_departments on oehr_employees.department_id = oehr_departments.department_id  
5 group by department_name  
6 order by count(employee_id) desc
```

Navigation

About

Use the App
to run, edit,
copy, or del
Select a pag
properties, c
to add a pag
Learn More

Tasks

Delete A
Copy Ap
Add Fea
Browse
Workflo

Working Co

Create Work

APEX Development Speed- GenAI

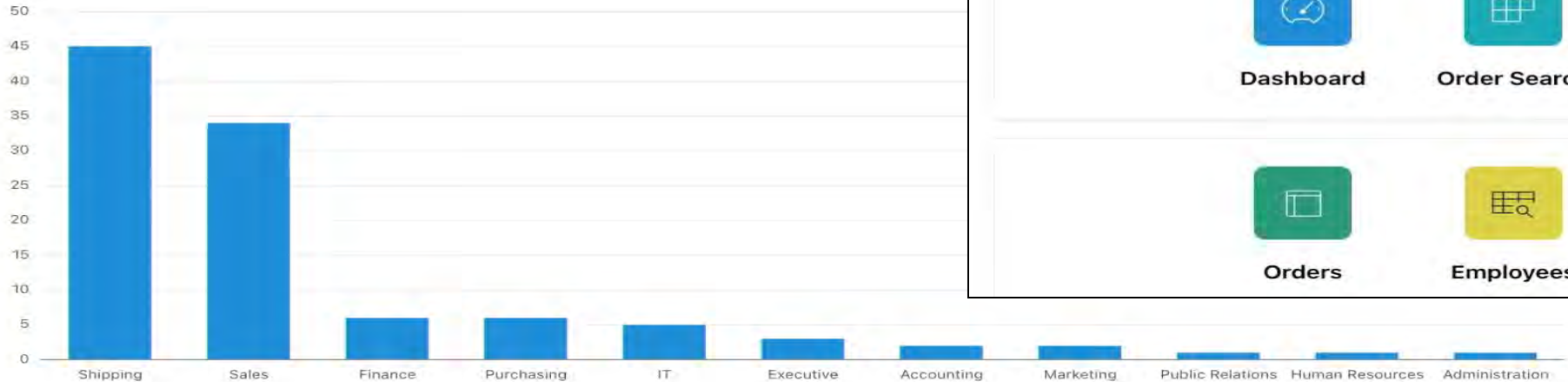


What's next from
Oracle Data Management

My Awesome App

Employees

Employees by Department



My Awesome App



My Awesom

Generated based on a S:



Dashboard



Order Search




Orders



Employees

APEX Development Speed- GenAI

 **What's next** from Oracle Data Management

APEX App Builder

Application 110 \ Page Designer

Page created!

Page 26: Employees by Department

- Pre-Rendering
- Components
 - Breadcrumb Bar
 - Breadcrumb
 - Body
 - Employees by Department
 - Series
 - Series 1
 - Axes
 - x
 - y
 - Average Salary by Job
 - Series
 - Series 1

Page Designer

Layout Page Search Help

Code Editor - SQL Query

APEx Assistant


```
1 show me average salaries by job
```

Save and Run Page

Cancel OK

Generative AI will also enable low-code developers to declare intent using natural language

Data Professional Developer **Low-Code Developer** End-User



APEX Development Speed- GenAI

Code Editor - SQL Query

```
1 select job_title as chart_label,  
2     avg(salary) as chart_value  
3 from oehr_employees,  
4     oehr_jobs  
5 where oehr_employees.job_id = oehr_jobs.job_id  
6 group by job_title  
7 order by avg(salary) desc
```



What's next from
Oracle Data Management

APEX App Builder - SQL Workshop - Team Development - Galle

Application 110 Page Designer

Page created!

Page 26: Employees by Department

- Pre-Rendering
- Components
 - Breadcrumb Bar
 - Breadcrumb
 - Body
 - Employees by Department
 - Series
 - Series 1
 - Average Salary by Job
 - Series
 - Series 1

Execution

Sequence 10

Source

Location Local Database

Type SQL Query

SQL Query

```
select job_title as chart_label,  
       avg(salary) as chart_value  
from oehr_employees,  
     oehr_jobs  
where oehr_employees.job_id =  
      oehr_jobs.job_id  
group by job_title  
order by avg(salary) desc
```

APEX Development Speed- GenAI



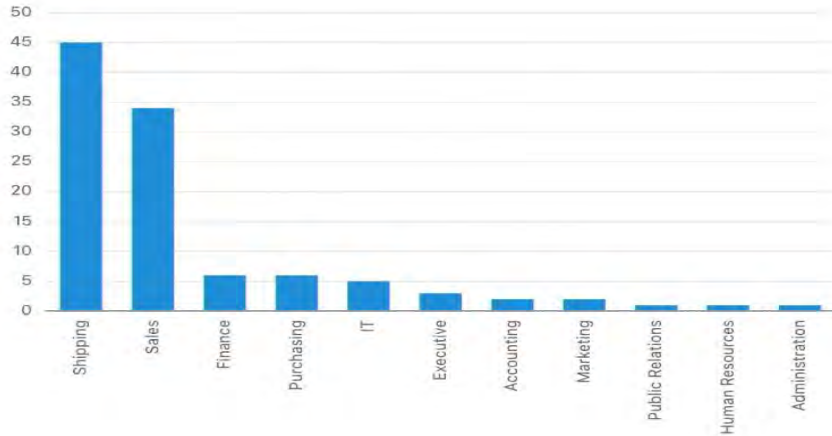
What's next from
Oracle Data Management

☰ My Awesome App

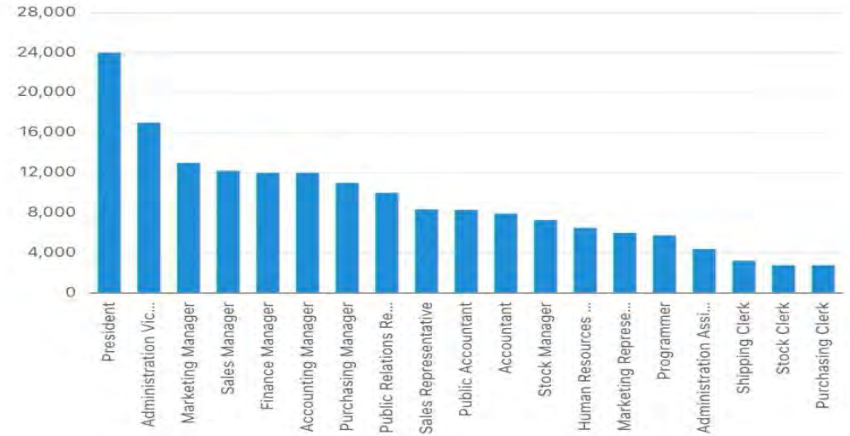
🔍 👤 demo

Employees

Employees by Department



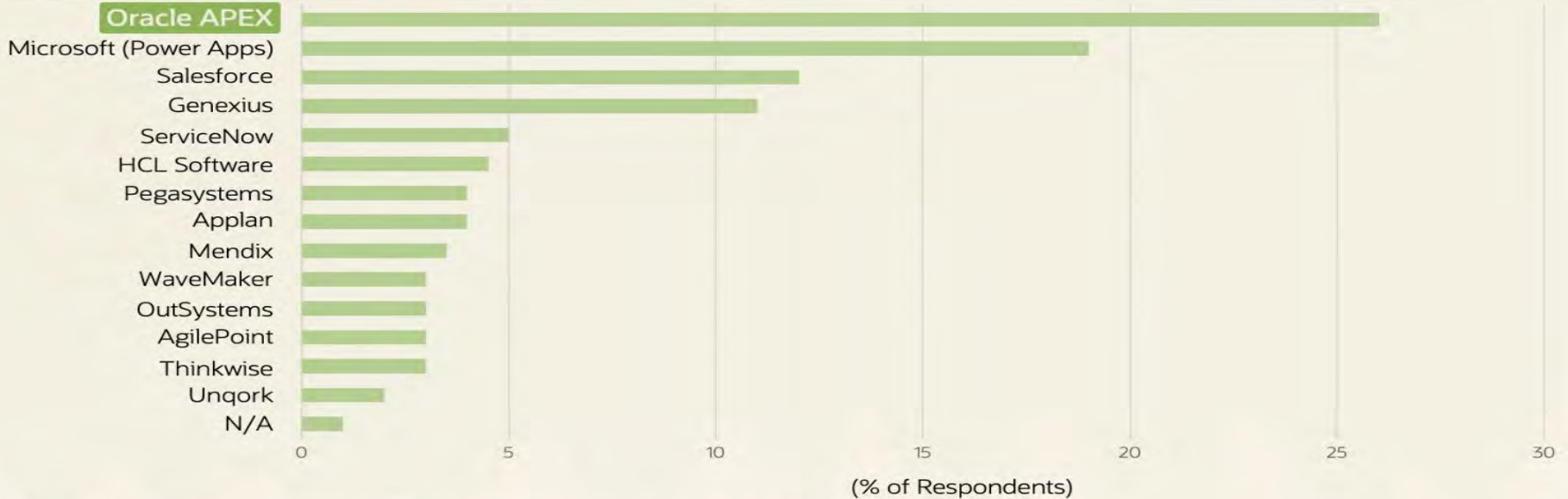
Average Salary by Job



APEX THE low code Dev Tool: 2M APEX Apps = 3K/day

IDC Survey Spotlight

Regardless of your organization's maturity with low code/no code app dev, which of the following platforms does your organization **predominantly utilize** today (or would likely use in the next 24 months) to build **low code/no code** applications?

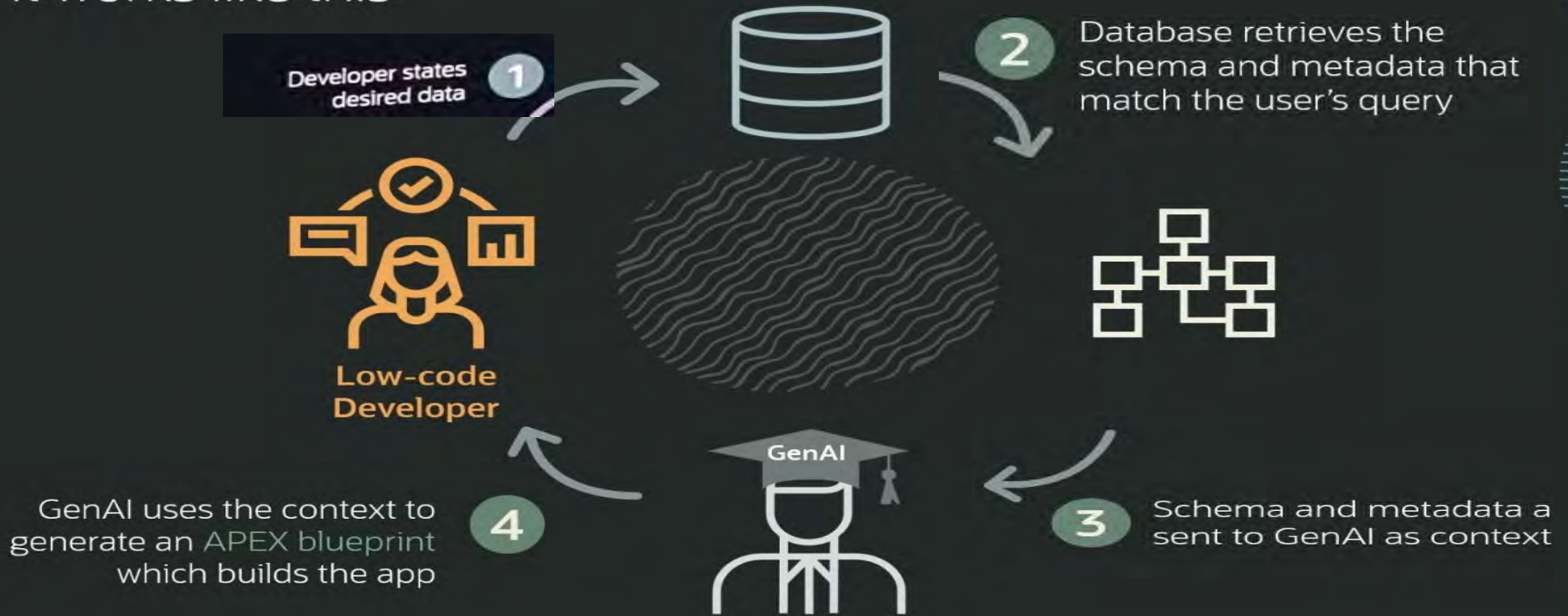


GenAI Writes the SQL

Generative AI can't generate a complete solution

It generates a first draft that experts validate and evolve

It works like this



GenAI will get to the End-User in ALL products

Data Professional



Developer



Low-Code Developer



End-User

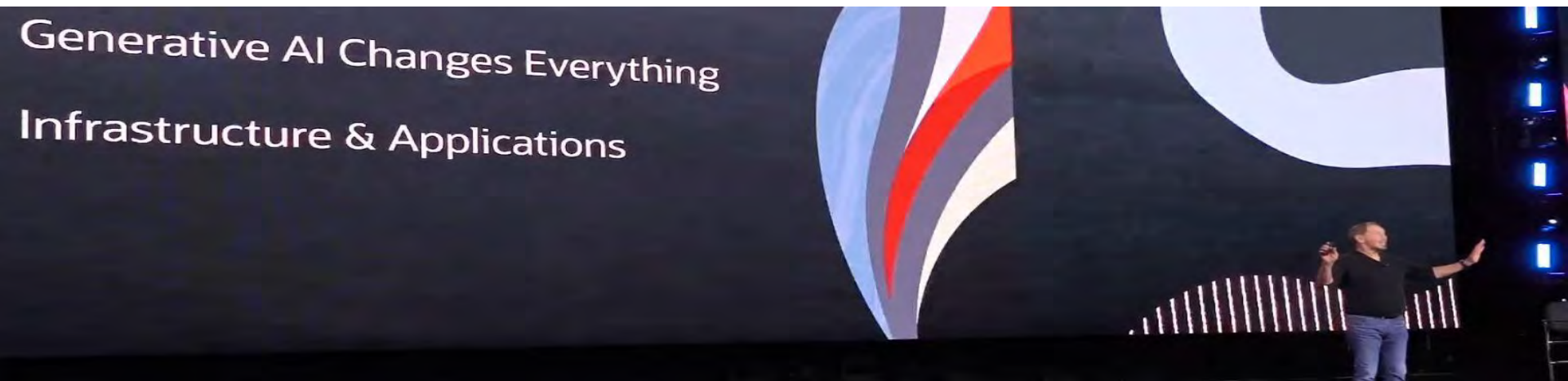


↓
Declare User Intent

Using AI, everyone can now search unstructured data, and query in natural language



What's next: A Worldwide race to build AI.



Is GenAI Most Important Ever? You're About to Find Out!

Generative Artificial Intelligence – Question?

Is Generative AI the Most Important New Computer Technology Ever? Probably!

- Countless Billions \$\$\$: Being Invested in Generative AI and Large Language Models
- Last Year: ChatGPT 4.0, Cohere, other LLMs Generated Images, Music, Computer Code ...
- Coming Soon: Self-Driving Cars, Antiviral Drug Design, Clinical Voice Digital Assistant ...

More Big Surprises !!



Healthcare Driving Oracle to Better AI Products

IoT Sensors: Collect Huge Amounts of AI Training Data

- Healthcare Training Data: Some Data Not Saved

- Patient Monitoring: Blood Pressure, Pulse Oximeter, Smart Watch

- Diagnostic Images: X-Rays, MRI Scans, Biopsy Slides ...

- Genomics: Gene Sequences, Variations and Mutations

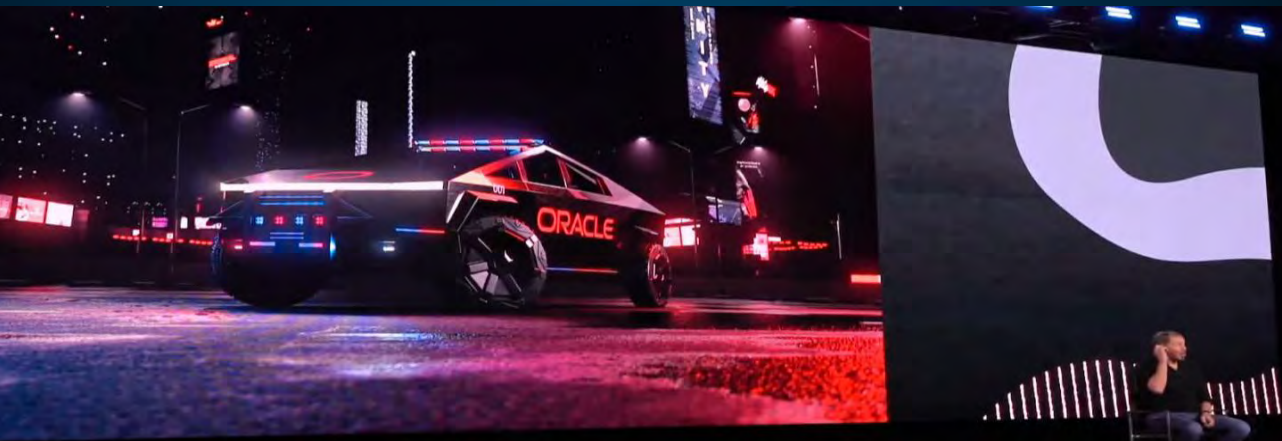
- Diagnostic Image Storage Cost Problem

- Diagnostic Images & Genomics: Not Saved Due to Cloud Storage Costs

- **Oracle Cloud: Provides an Economic Solution to this Problem**



Oracle Driving First Responders with Tesla

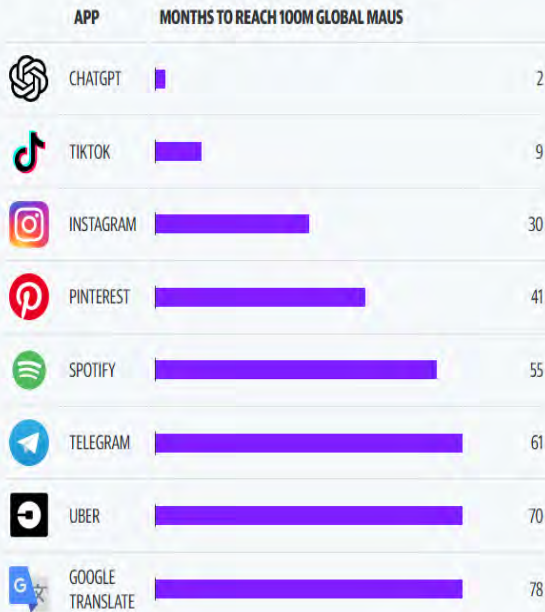


OpenAI - Ten Years Later... the Baby Talks!

HOW LONG IT TOOK TOP APPS TO HIT 100M MONTHLY USERS

ChatGPT is estimated to have hit 100M users in January, 2 months after its launch.

Here's how long it took other top apps to reach that:



About One Year Ago: ChatGPT 3.5 was Born ... the Baby Talks

- Developed by OpenAI: A Company Founded by Elon Musk 10 Years Ago
- Huge Artificial Neural Network: Multi-Billion Parameter Large Language Model (LLM)
- Trained on Language: Trillions of Data Elements ... Wikipedia + Public Internet
- ChatGPT 3.5 Talks: Generates Answers to Questions ... Engages in Conversations

A Big Surprise
ChatGPT's Scale Enabled it to Understand and Generate Language

Main Difference Between GPT-4 and ChatGPT

Unlike ChatGPT, GPT-4 can see and understand images, whereas the former is limited to text.

While ChatGPT had several issues upon its arrival, Microsoft claims that GPT-4 has been trained to avoid several malicious prompts.

Generative AI - Things to Know...

- **ChatGPT predicts the next word ... looks at less stuff.**
- **Oracle announces Vector DB to store words, images+ & use LLM to search**
- LLM's - Every Text / Email has auto-complete; translate languages, using a NN
- GenAI, LLM, ChatGpt (Generates News, articles, etc.)
- Will get us to AGI - Artificial Generative Intelligence. Language gets us to AGI.
- **“Time of Exponential Development”** - Similar to how Computers started to Exponentially get bigger & faster.
- ***We use language - even when thinking about something***
- LLM's are predictive machines, training data of Wikipedia & the Internet
- 100 layers of neurons for ChatGPT4 - **But, still can “Hallucinate”**
- Trains the NN by reading the internet ... etc.
- GenAI, Gpt4 (LLM), DeepMind WaveNet (text to audio), OpenAI DALL·E 2 - (text to images), Google DeepMind Google AlphaFold (phara - protein folding - same timing as ChatGPT 3.5) discoveries, ChatPDF, SlideGPT, MusicLM

ChatGPT

GPT-4 has arrived. It will blow ChatGPT out of the water *The Washington Post* 3/14/2023

ChatGPT was launched on November 30, 2022, by San Francisco-based OpenAI (the creator of the GPT series of large language models; DALL·E 2, a diffusion model used to generate images; and Whisper, a speech transcription model).



OpenAI Research Product Developers Safety Company Search Log in Sign up

Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

Try ChatGPT Read about ChatGPT Plus

OpenAI

ChatGPT → DALL·E → API →

Interact with our flagship language models in a conversational interface

Create realistic images and art from a description in natural language

Integrate OpenAI models into your application or business

Main Difference Between GPT-4 and ChatGPT

Unlike ChatGPT, GPT-4 can see and understand images, whereas the former is limited to text.

While ChatGPT had several issues upon its arrival, Microsoft claims that GPT-4 has been trained to avoid several malicious prompts.

RI what are the top 10 databases

As of my knowledge cutoff in September 2021, the top 10 databases were:

1. Oracle Database: A relational database management system (RDBMS) known for its scalability, security, and comprehensive features.
2. MySQL: An open-source RDBMS widely used for web applications due to its simplicity, reliability, and compatibility.
3. Microsoft SQL Server: A popular RDBMS developed by Microsoft, known for its strong integration with Windows Server and seamless integration with other Microsoft products.
4. PostgreSQL: An open-source object-relational database management system (ORDBMS) known for its extensibility, robustness, and compliance with SQL standards.
5. MongoDB: A NoSQL database that uses a flexible, document-oriented model and is known for its scalability, performance, and ease of use.
6. IBM Db2: A family of database management systems developed by IBM, known for its reliability, scalability, and support for a wide range of applications.
7. SQLite: A lightweight, file-based, embedded database engine that requires no separate server process and is commonly used in mobile and embedded systems.
8. Redis: An in-memory data structure store that can be used as a database, cache, or message broker. It is known for its high performance and flexibility.
9. Cassandra: A distributed and highly scalable NoSQL database designed for handling large amounts of data across commodity servers and providing high availability and fault tolerance.
10. Amazon DynamoDB: A fully managed NoSQL database service provided by Amazon Web Services (AWS).

Regenerate response

Google's Bard



☰ Bard Experiment

↻ Reset chat

🕒 Bard Activity

❓ FAQ

📄 Updates

❗ Help

R What are the top 10 databases

Here is a brief overview of each of the top 10 databases:

- **Oracle** is a commercial database that is known for its scalability and performance. It is often used by large enterprises that need to store and manage large amounts of data.
- **MySQL** is an open-source database that is known for its ease of use and flexibility. It is often used for web applications and other applications that need to be scalable and easy to maintain.
- **Microsoft SQL Server** is a commercial database that is known for its integration with other Microsoft products. It is often used by businesses that use Microsoft's Windows operating system and other Microsoft products.
- **PostgreSQL** is an open-source database that is known for its power and flexibility. It is often used for applications that need to store complex data or that need to be highly scalable.
- **MongoDB** is a NoSQL database that is known for its flexibility and scalability. It is often used for applications that need to store large amounts of unstructured data.
- **Redis** is a NoSQL database that is known for its speed and performance. It is often used for applications that need to store and manage large amounts of data in real time.
- **IBM DB2** is a commercial database that is known for its scalability and performance. It is often used by businesses that need to store and manage large amounts of data.
- **Elasticsearch** is a search engine database that is known for its speed and scalability. It is often used for applications that need to search and index large amounts of data.
- **Cassandra** is a NoSQL database that is known for its scalability and fault tolerance. It is often used for applications that need to store and manage large amounts of data in a distributed environment.
- **MariaDB** is a fork of MySQL that is known for its stability and performance. It is often used as a replacement for MySQL when MySQL is not available or when a more stable database is needed.

Enter a prompt here



● Illinois, USA

From your IP address • [Update location](#)

Bard may display inaccurate or offensive information that doesn't represent Google's views.

Transformers - Google, 2017

FYI

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jacob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukaszkaizer@google.com

Illia Polosukhin* ‡
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

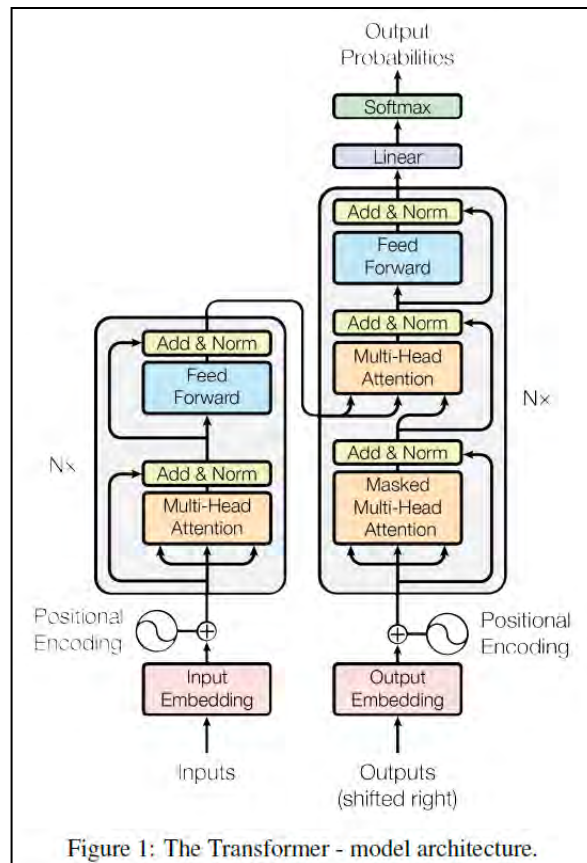


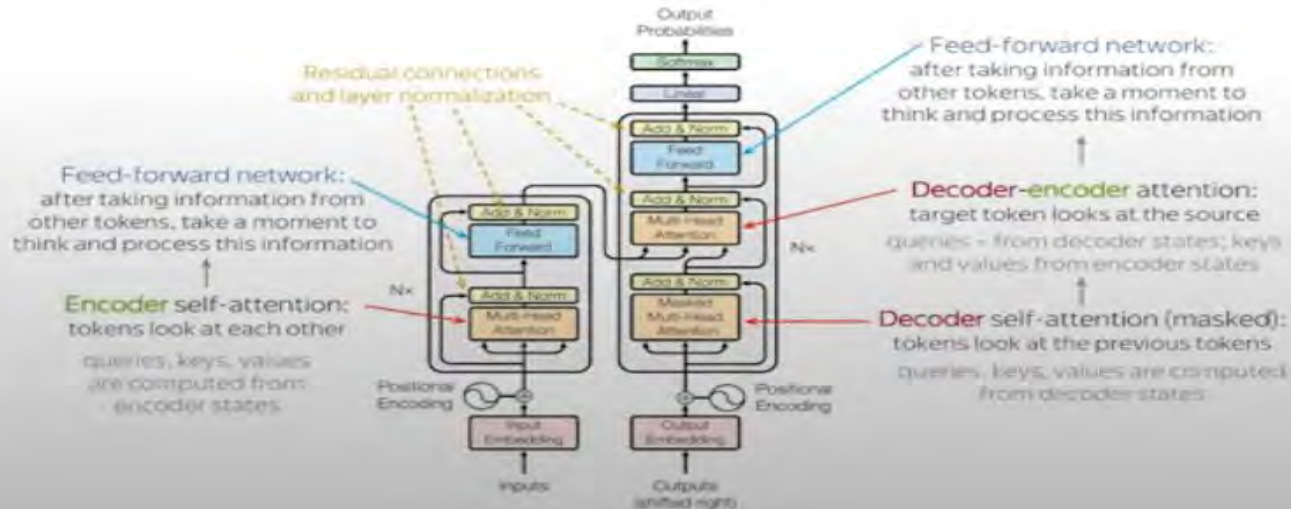
Figure 1: The Transformer - model architecture.

Leverages Parallelism to train model

Transformer Model led to Dominant Design

- All main models today are transformer models

The Transformer Architecture: Gene's dominant design?



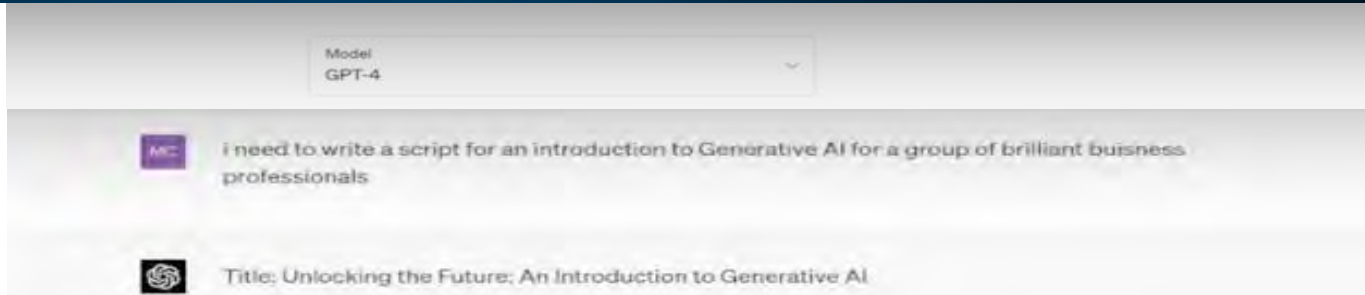
Generative AI: GPT & ChatGPT

FYI



- **GPT = Generative Pre-trained Transformer**
- **ChatGPT is a Foundation Model** (Stanford paper 2021/2022) **for LLM (Large Language Model) Chatbots.** At end of 2018, NLP changed with start of foundation models begins.
- **Transformer model architecture** [Vaswani et al. 2017] that **leverages the parallelism of the hardware to train** much more expressive models than before.
- Foundation models have also led to surprising emergence which results from **scale**. For example, **GPT-3 [Brown et al. 2020], with 175 billion parameters** compared to **GPT-2's 1.5 billion**, permits in-context learning. (DALL-E for images)
- **GPT-4 is 1 trillion parameters** & you refine using your data.
- **GenAI Generates new content/output (GANs & Deep Fake)**

ChatGPT-4



GPT-4 Technical Report



This report focuses on the capabilities, limitations, and safety properties of GPT-4. GPT-4 is a Transformer-style model [39] pre-trained to predict the next token in a document, using both publicly available data (such as internet data) and data licensed from third-party providers. The model was then fine-tuned using Reinforcement Learning from Human Feedback (RLHF) [40]. **Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar.**




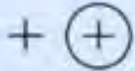














ChatGPT 4 is Bi-Modal (Images/Words)

Predictions: Potential Capabilities of

GPT-4



Latent Space

 <p>What's in this picture?</p> <p>Looks like a duck.</p> <p>That's not a duck. Then what's it?</p> <p>Looks more like a bunny.</p> <p>Why?</p> <p>It has bunny ears.</p>	<p>Here are eight images:</p>    <p>The following image is:</p> <p><input checked="" type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C</p>    <p>D <input type="checkbox"/> E <input type="checkbox"/> F</p>   	<p>Input Prompt</p>  <p>Question: Explain why this photo is funny? Answer:</p> <p>Completion</p> <p>The cat is wearing a mask that gives the cat a smile.</p> <p>(1)</p>	<p>Input Prompt</p>  <p>Question: Why did the little boy cry? Answer:</p> <p>Completion</p> <p>Because his scooter broke.</p> <p>(2)</p>	<p>Input Prompt</p>  <p>Question: What is the hairstyle of the blond called? Answer:</p> <p>Completion</p> <p>pony tail</p> <p>(3)</p>	<p>Input Prompt</p>  <p>Question: When will the movie be released? Answer:</p> <p>Completion</p> <p>On June 27</p> <p>(4)</p>
		<p>Input Prompt</p>  <p>What is TorchScale?</p> <p>Completion</p> <p>A library that allows transformers to work efficiently and effectively at any scale.</p> <p>(5)</p>	<p>Input Prompt</p>  <p>Question: The result is? Answer:</p> <p>Completion</p> <p>5 + 4 = 9</p> <p>(6)</p>	<p>Input Prompt</p>  <p>Question: What is the heart rate in this picture? Answer:</p> <p>Completion</p> <p>57 bpm</p> <p>(7)</p>	<p>Input Prompt</p>  <p>The time now is</p> <p>Completion</p> <p>10:10 on a large clock</p> <p>(8)</p>

Oracle Vector Database Introduced!

Oracle Vector Database: Easily Specialize AI Models & Build AI Apps

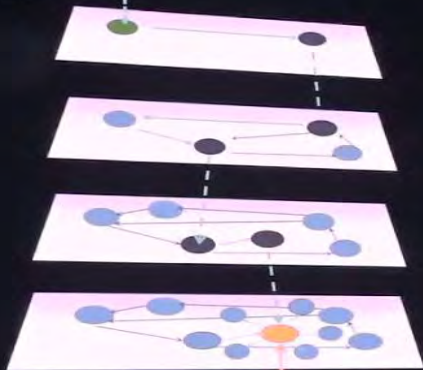
- Designing & Training Foundational AI Models: Complex and Expensive
 - Design Multi-Billion Parameter Artificial Neural Network: OpenAI, Cohere, xAI ...
 - Ingest Trillions of Training Data Elements: Wikipedia + Everything on the Internet
 - Oracle Cloud: 16,000 Node H100 NVIDIA Supercluster with RDMA Interconnect
- Specializing AI Models: Pretrained Foundational Model + Supplementary Training Data
 - Oracle Vector Database: Stores Vectorized Supplementary Training Data
 - Specialized AI Model with EHR Data: Generate Doctors' Orders – Cerner New Millennium
 - Specialized AI Model with Diagnostic Image Data: Rapid Cancer Detection – Imagen

Most Customers will Train Specialized AI Models for Specific Applications



Vector Indexes Leverage Partitions / RAC / Exadata

Entry Point -----



Query Vector

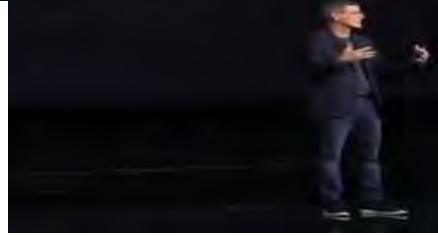


Oracle database accelerates
AI vector search using
sophisticated vector indexes



From Juan Loaiza Interview

AI vector search using
sophisticated vector indexes



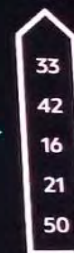
- Time of Big Transitions
 - **Generating Apps with GenAI**
 - **AI & AI Search**
- **Dramatic! Orders of benefit - at least 10x**
- GenAI - build directly into DB - No need to be a data scientists
- **AI Vector search - take a document/image with the content (not pixels) - build a vector for this & compare things better.** Very new tech!!!
- Used to need to be a data scientist... now just use SQL with vector searches.
- **Talk to the system like ChatGPT to pose questions to the database & get answers back using vector search (in 23c only)**
- AI benefit - does things like **Generate Apps** - describe it & they will build the first version of it. You get a DRAFT of it. You just need to check everything (like chatGPT).
- **It will keep improving!** Hard to predict where it is going - not as esoteric - used to be for corner cases - easier now & better.
- **Won't replace experts - gives a first draft** - verify draft / update it / own it
- AI won't get it 100% right
- Traditionally via Relational Data - App users want to use **JSON & Graph: ALL ok now.**

Search on Data with AI Vectors & Vector Indexes

Vectors are used to represent the semantic content of images, documents, videos, etc.



Vector



A vector is a sequence of numbers, called dimensions, used to capture the important "features" of the data



Oracle Vector Database

Example: the features for a house image could be

Vector

Features

House



Type of Roof

Decorations

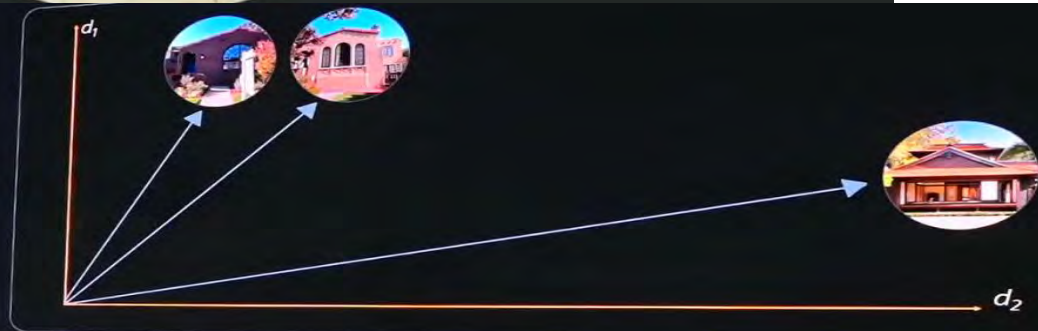
Number of Stories

Building Materials



Each dimension represents a different feature of the house

House vectors when collapsed into 2 dimensions instead of hundreds could look like this

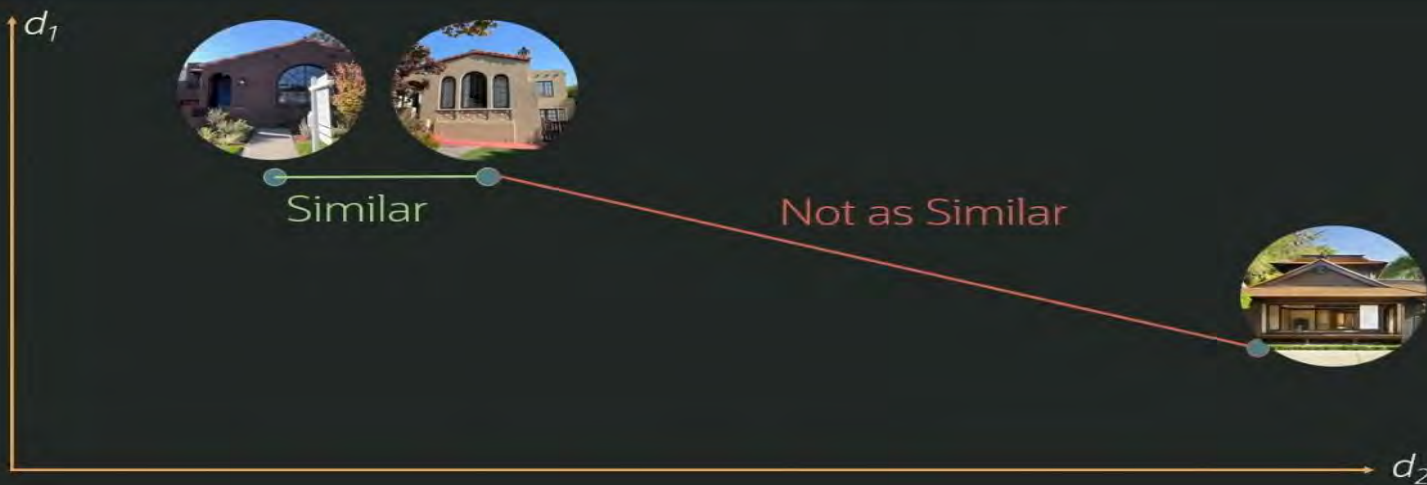


Vector Search

Imagine a house-hunting app that helps customers find houses for sale that are similar to a picture the customer uploads



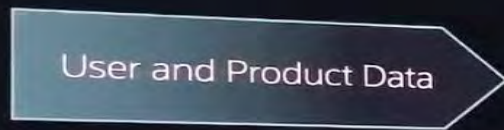
The distance between the vectors is proportional to their semantic similarity



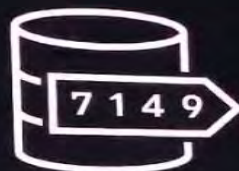
Search on vector (GenAI), add prices/locations



Business Database



Continuously send your business data to a vector database



Vector Database

Searches on a combination of business and semantic data are more effective if both types of data are stored together



Create table with Vector Data Type & Blob

```
CREATE TABLE house_for_sale (house_id      number,  
                             price         number,  
                             city          varchar2(400),  
                             house_photo  blob,  
                             house_vector vector  
);
```

Find houses that are similar to this picture and match the customer's preferred city and budget



```
SELECT ...  
FROM   house_for_sale  
WHERE  price <= (SELECT budget          FROM customer ...)  
AND    city  in (SELECT search_city    FROM customer ...)  
ORDER BY vector_distance(house_vector, :input_vector);
```

AI Vector search Preview!



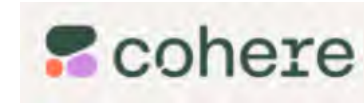
Announcing:
AI Vector Search in
Oracle Database 23c

*Sign up
for Preview Now*



Cohere & Oracle GenAI Partnership (Signed June 2023)

- **Aiden Gomez, Cohere CEO** worked on “Attention is all you need” paper on Transformers Aidan while at Google Brain
- **Cohere turns words to numbers with semantic knowledge**
- Knowledge Augmentation - Don't read 100 docs, **feed into Cohere LLM models** (Working with ALL Oracle Teams).
- Vector DB uses your information & knowledge that you have
- **Challenges - Hallucinations (made up “facts”)**
- **RAG** - most promising **solution to hallucinations**
- Models trained on the web can now use your own data with RAG (Retrieval Augmented Generation)
- Training vs. RAG - Training good to **get model to speak in your brand voice** - Adds new knowledge coming in every msec when you update the DB & RAG keeps it up to date.
- **Future - Embedding Models - performs twice as well in speed & accuracy. Also compress 32x with 96% accuracy.**
- **Noisy data** (pdf scrapes / errors); bad emails; No matter how noisy, you still get accurate results with what we're building.



Officially launched partnership at Cloud World (Sept 2023)

ORACLE®



Retrieval Augmented Generation

Vector Databases augment Generative AI by retrieving detailed, often private content needed to answer questions

Called: Retrieval Augmented Generation (RAG)

Retrieval Augmented Generation How it works

The user's question is encoded as a vector and sent to a Vector DB

1



2

Vector DB finds private content that closely match the user's question

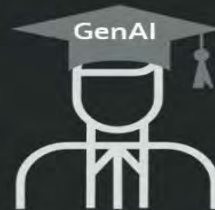


User



3

The content is sent to the GenAI to help answer the user's question



4

GenAI uses the content plus general knowledge to provide an informed answer

Coming soon: OCI Generative AI

State-of-the-art foundational models

We have partnered with Cohere to bring SOTA foundational models to both OCI customers and Oracle SaaS customers.

Fine-tuning support

Tune custom models with your own data to solve specific problems or target specific domains.

Fully hosted on OCI

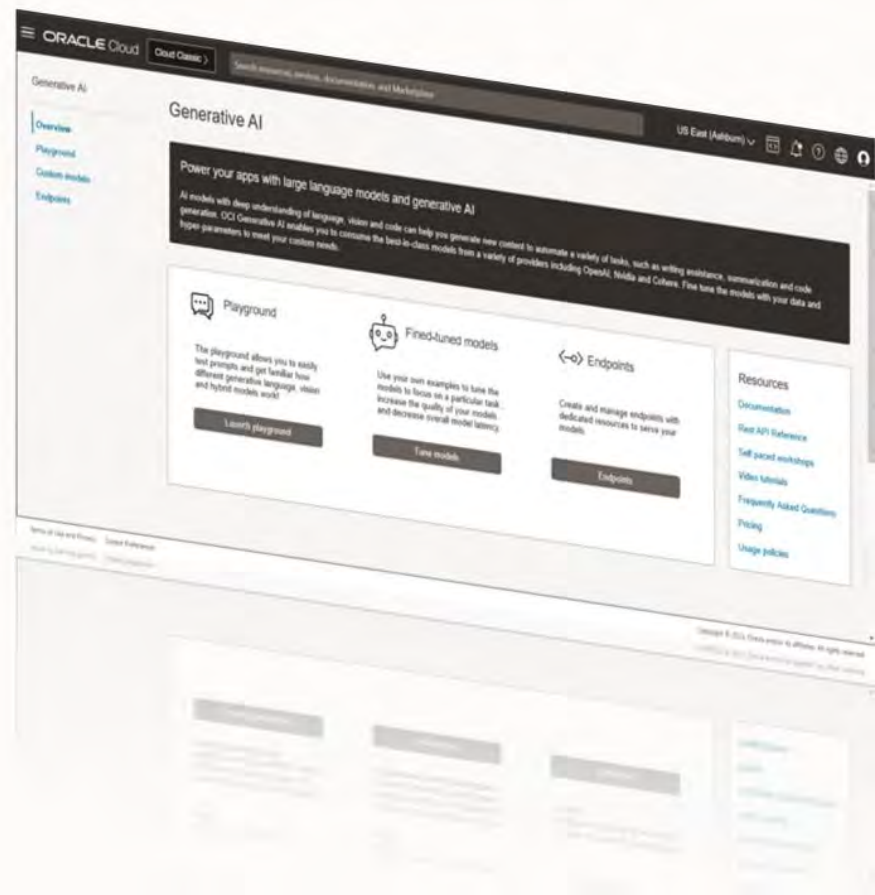
All processing and data storage happens inside OCI, no cross-region or cross-cloud communication.

Respect for customer privacy

Customer training and inference data is owned by the customer and only used to serve the customer.

PAYG and single tenant, dedicated hardware

Pay per token or host the models on dedicated hardware for stable performance.



The Oracle AI Stack



Business applications, Oracle SaaS portfolio

ORACLE
Cerner

ORACLE
NETSUITE

ORACLE
Aconex

...

AI services



Digital Assistant



Speech



Language



Vision



Document
Understanding



Anomaly Detection



Forecasting



+
cohere

OCI Generative AI

Machine learning services



OCI Data Science



ML in Oracle Database



MySQL Heatwave



OCI Data Labeling

Data Platform

AI infrastructure



Compute bare metal instances and VMs



Cluster networking



Block, object, and file storage; HPC filesystems



GPUs, OCI Supercluster



NVIDIA
A100

Answer Detailed Questions / Supply Manuals

Kate's List

Your AI powered source for home repairs

Natalie Beckham

Welcome to Kate's List. How can I help you?

Hello, my power generator does not seem to be starting up on saturday mornings anymore. The model is X2341.



I am sorry to hear that. According to our database there are several potential reasons that may cause a generator not to start on it's scheduled time.

- Check if combustion can take place. According to 2, a generator won't start when it's not getting the fuel it needs, air or a spark.
- If your generator is programmable, check that it is programmed correctly. 3
- Checked that the generator is actually on. 4

Please refer to the X2341 generator operator's manual 1 for safety precautions.

Hmm, it's definitely programmed correctly. I think it's the spark plug. Which one do you recommend?







For an X2341 sparkplug, I recommend an NGK Spark Plug, NGK BCPR5ES. 1

ok, where can I fi



Home Repair Resources

Use our automated chat to the left to get assistance now!

- >  1. User's Guide X2341 USER MANUAL
- >  2. X Series GeneApp Mobile Application USER MANUAL
- >  3. 12 Reasons why your generator may not be ready for Winter. ARTICLE
- >  4. User's Guide X2341 USER MANUAL

Answer Location Questions - Integrate Map

Kate's List

Your AI powered source for home repairs

Natalie Beckham

I am sorry to hear that. According to our database there are several potential reasons that may cause a generator not to start on it's scheduled time.

- Check if combustion can take place. According to 2, a generator won't start when it's not getting the fuel it needs, air or a spark.
- If your generator is programmable, check that it is programmed correctly. 3
- Checked that the generator is actually on. 4

Please refer to the X2341 generator operator's manual 1 for safety precautions.

Hmm, it's definitely programmed correctly. I think it's the spark plug. Which one do you recommend?



For an X2341 sparkplug, I recommend an NGK Spark Plug, NGK BCPR5ES. 1

ok, where can I find in my area?



The spark plug can be found at the following stores in the Huntsville area (See map to the right). I can also order one for you online with free shipping. It's only \$4.99.

OH - online, even better.



> 3. 12 Reasons why your generator may not be ready for Winter.pdf

ARTICLE

> 4. User's Guide X2341

USER MANUAL

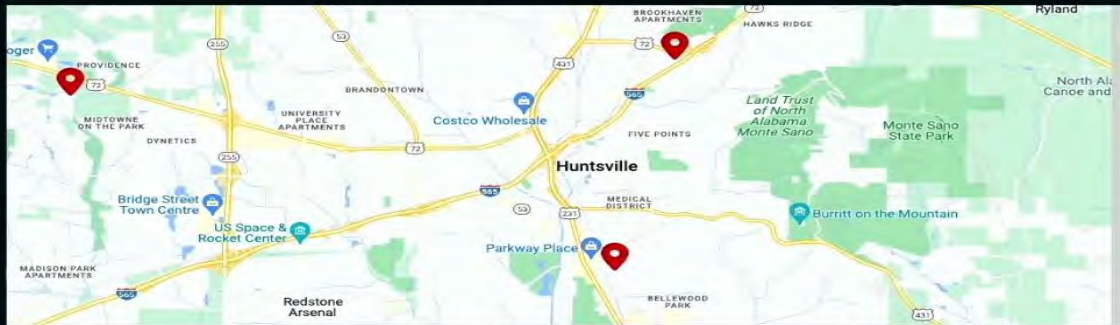
▼ Huntsville, AL 3 Locations:

MAP & LOCATION


Acme Parts
234 Hughes Rd
Huntsville, AL 55543

Foobar Hardware
456 2nd Street
Huntsville, AL 55543

The Third Wheel
987 Long Parkway
Huntsville, AL 55543



Ask Questions & Reference Document AI Searches

↓ Ask a Question 

Page 04 / 21



Chat with Nicky



Welcome to Acme Insurance. How can I help you today?

Are hotel stays covered in my policy?



Boarding is not covered as part of your current policy.
(See [Policy Exclusions](#), Page 13)

← Now displaying Page 13.

↓ FOUND ON THIS PAGE BELOW

WHAT IS NOT COVERED

Exclusions

We will not pay for costs associated with or resulting from the following:

- a. Aesthetic, cosmetic, endodontic, or orthodontic dental services such as caps, crowns or crown amputation, fillings, implants and root canals or planing.
- b. Organ or heart valve transplants.
- c. Boarding.
- d. **Conditions that occur** during a waiting period.
- e. Cosmetic and elective prostheses or procedures (including but not limited to claw removal, ear cropping and tail docking).
- f. Dental cleanings unless used to treat a covered **illness** or covered by an applicable endorsement.
- g. **Experimental** or investigational **treatment** or medication (including clinical trials) that is not generally accepted in the **veterinary** medical community as effective or proven.
- h. Funeral services, memorial items, urns, caskets, or burial plots/fees.
- i. Grooming or grooming supplies (including but not limited to non-prescription baths, ear cleanings, non-prescription shampoos and nail trims).
- j. Food products.
- k. House call fees, time and travel expenses to and from the **veterinarian's** premises or hospital.
- l. **Illness** or **injury** that results from intentional, malicious, or grossly negligent activities or from failure to perform actions commonly accepted as responsible **pet** care by **you**, a member of **your** household or a caregiver for **your pet**.
- m. **Ligament and knee conditions**, if any **ligament and knee condition occurred** prior to the first **effective date** of the applicable coverage or during a waiting period.
- n. Non-medical supplies such as but not limited to toys, leashes, ramps, bedding or other devices intended to prevent **injury** or **illness**, but that do not treat a **condition**.
- o. Non-**veterinary** services (including but not limited to administrative fees, medical records expenses, medical waste, discount package or membership fees, postage and tax).
- p. **Pre-existing conditions that occurred** on or before the first **effective date** of the applicable coverage or during a waiting period.

Stanford (113 authors / 215 pgs.) on Foundation Models 2021/2022

On the Opportunities and Risks of Foundation Models

Rishi Bommasani* Drew A. Hudson Ehsan Adeli Russ Altman Simran Arora
 Sydney von Arx Michael S. Bernstein Jeannette Bohg Antoine Bosselut Emma Brunskill
 Erik Brynjolfsson Shyamal Buch Dallas Card Rodrigo Castellon Niladri Chatterji
 Annie Chen Kathleen Creel Jared Quincy Davis Dorottya Demszky Chris Donahue
 Moussa Doumbouya Esin Durmus Stefano Ermon John Etchemendy Kawin Ethayarajh
 Li Fei-Fei Chelsea Finn Trevor Gale Lauren Gillespie Karan Goel Noah Goodman
 Shelby Grossman Neel Guha Tatsunori Hashimoto Peter Henderson John Hewitt
 Daniel E. Ho Jenny Hong Kyle Hsu Jing Huang Thomas Icard Saahil Jain
 Dan Jurafsky Pratyusha Kalluri Siddharth Karamcheti Geoff Keeling Fereshte Khani
 Omar Khattab Pang Wei Koh Mark Krass Ranjay Krishna Rohith Kuditipudi
 Ananya Kumar Faisal Ladhak Mina Lee Tony Lee Jure Leskovec Isabelle Levent
 Xiang Lisa Li Xuechen Li Tengyu Ma Ali Malik Christopher D. Manning
 Suvir Mirchandani Eric Mitchell Zanele Munyikwa Suraj Nair Avanika Narayan
 Deepak Narayanan Ben Newman Allen Nie Juan Carlos Niebles Hamed Nilforoshan
 Julian Nyarko Giray Ogut Laurel Orr Isabel Papadimitriou Joon Sung Park Chris Piech
 Eva Portelance Christopher Potts Aditi Raghunathan Rob Reich Hongyu Ren
 Frieda Rong Yusuf Roohani Camilo Ruiz Jack Ryan Christopher Ré Dorsa Sadigh
 Shiori Sagawa Keshav Santhanam Andy Shih Krishnan Srinivasan Alex Tamkin
 Rohan Taori Armin W. Thomas Florian Tramèr Rose E. Wang William Wang Bohan Wu
 Jiajun Wu Yuhuai Wu Sang Michael Xie Michihiro Yasunaga Jiaxuan You Matei Zaharia
 Michael Zhang Tianyi Zhang Xikun Zhang Yuhui Zhang Lucia Zheng Kaitlyn Zhou
 Percy Liang*¹

Center for Research on Foundation Models (CRFM)
 Stanford Institute for Human-Centered Artificial Intelligence (HAI)
 Stanford University

AI is undergoing a paradigm shift with the rise of models (e.g., BERT, DALL-E, GPT-3) trained on broad data (generally using self-supervision at scale) that can be adapted to a wide range of downstream tasks. We call these models foundation models to underscore their critically central yet incomplete character. This report provides a thorough account of the opportunities and risks of foundation models, ranging

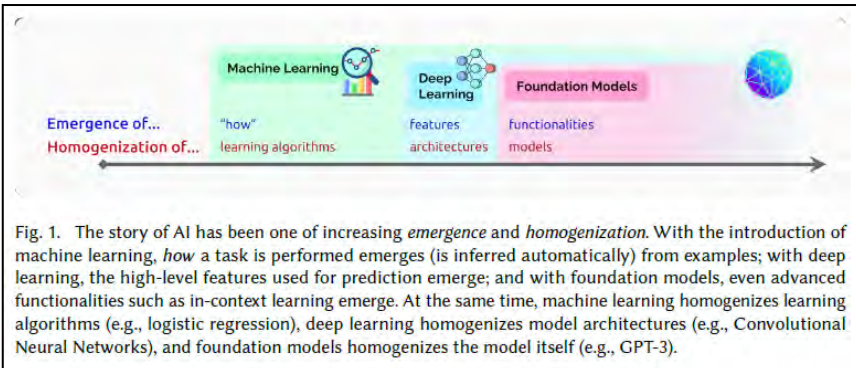
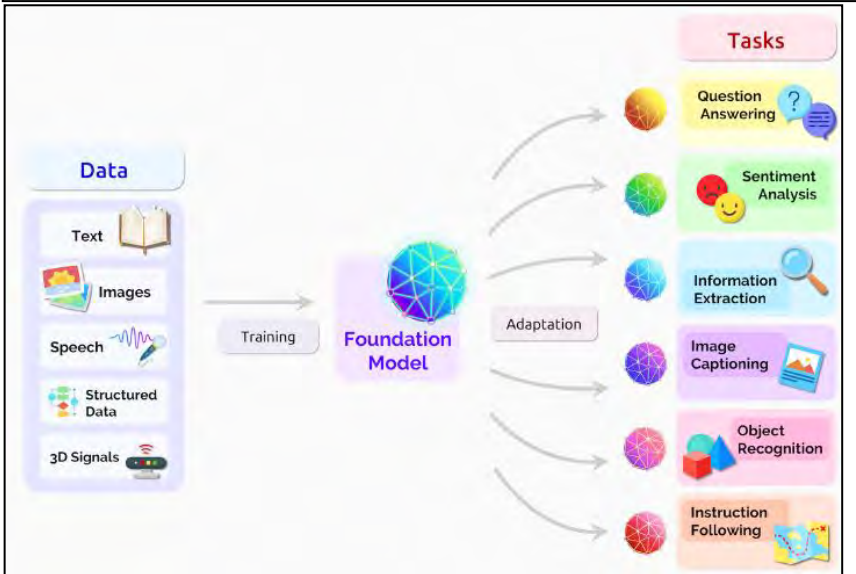


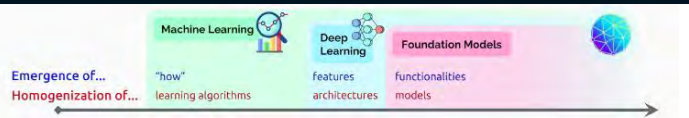
Fig. 1. The story of AI has been one of increasing *emergence* and *homogenization*. With the introduction of machine learning, *how* a task is performed emerges (is inferred automatically) from examples; with deep learning, the high-level features used for prediction emerge; and with foundation models, even advanced functionalities such as in-context learning emerge. At the same time, machine learning homogenizes learning algorithms (e.g., logistic regression), deep learning homogenizes model architectures (e.g., Convolutional Neural Networks), and foundation models homogenizes the model itself (e.g., GPT-3).



arXiv:2108.07258v3 [cs.LG] 12 Jul 2022

From Stanford Paper

FYI



- **Machine learning:** Most AI systems today are powered by machine learning, where predictive models are trained on historical data and used to make future predictions.
- **Deep neural networks** would be trained on the raw inputs (e.g., pixels), and higher-level features would emerge through training (a process dubbed “representation learning”).
- **Foundation models** are enabled by transfer learning [Thrun 1998] and scale. The idea of transfer learning is to take the “knowledge” learned from one task (e.g., object recognition in images) and apply it to another task (e.g., activity recognition in videos).

Fig. 1. The story of AI has been one of increasing *emergence* and *homogenization*. With the introduction of machine learning, how a task is performed emerges (is inferred automatically) from examples; with deep learning, the high-level features used for prediction emerge; and with foundation models, even advanced functionalities such as in-context learning emerge. At the same time, machine learning homogenizes learning algorithms (e.g., logistic regression), deep learning homogenizes model architectures (e.g., Convolutional Neural Networks), and foundation models homogenizes the model itself (e.g., GPT-3).

From Stanford Paper

FYI

124

Center for Research on Foundation Models (CRFM)

One Model

A finite number of **generalizable model mechanisms are combined** to produce behaviors across tasks.



Many Models

For each task, distinct model mechanisms are used to produce behaviors; akin to a **large collection of individual expert models**.



On the Opportunities and Risks of Foundation Models

4.7 Security and privacy

Authors: Florian Tramèr*, Rohith Kudithipudi*, Xuechen Li*



Fig. 20. Risks and opportunities raised by foundation models for security and privacy of ML systems.

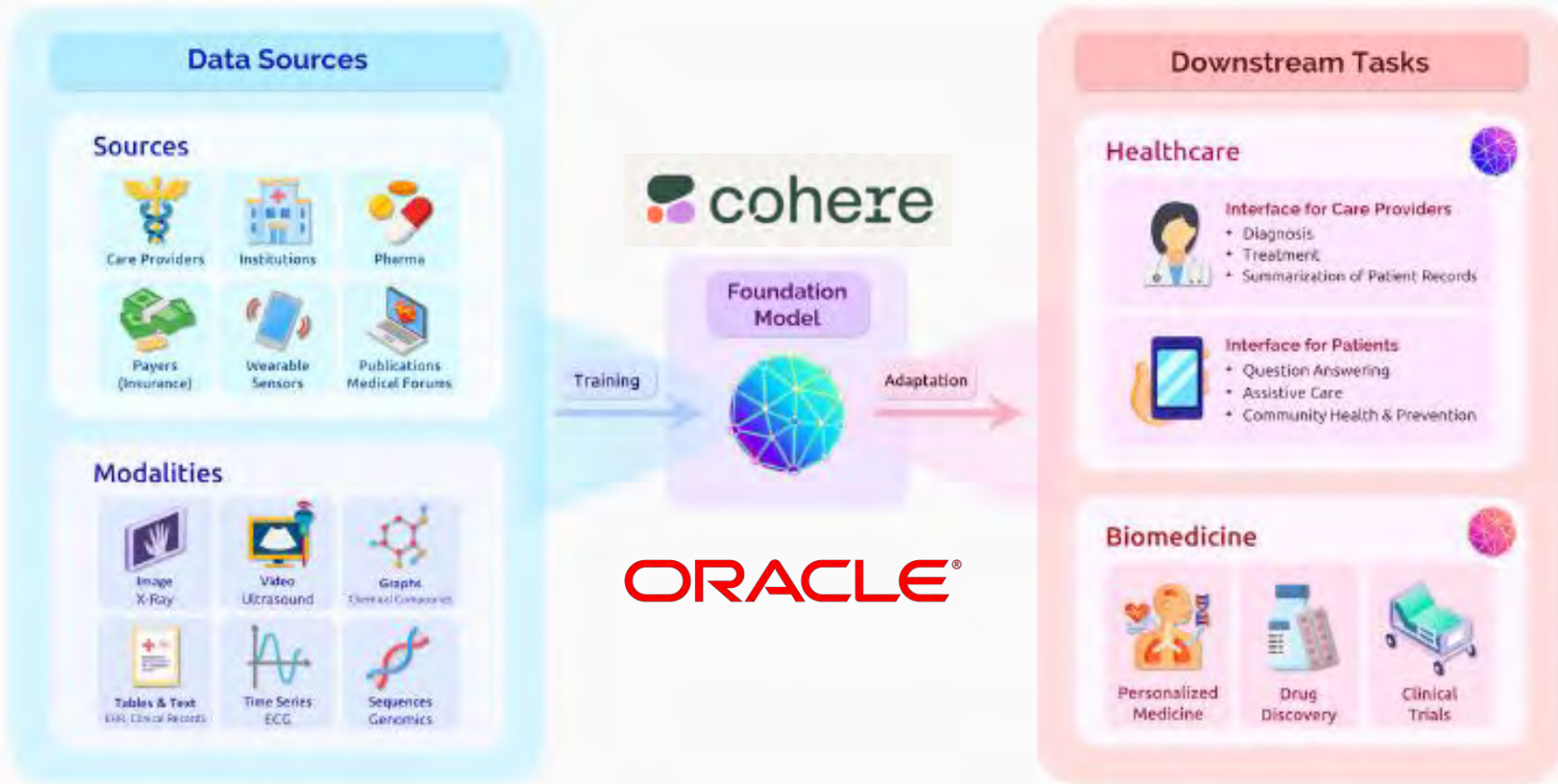
Will Businesses of the Future just be a Single Foundation Model to do all Tasks? (my own question)

Fig. 23. The one model–many model nature of foundation models: A central interpretability question is to understand where a foundation model lies on the spectrum between *one model* and *many models*. As one model, behavior can be made interpretable by identifying and characterising the finite number of generalizable model mechanisms used to produce behaviors across tasks (e.g., mechanisms that assign meaning to words, compare quantities, and perform arithmetic). As many models, explanations of model behavior in one task are not necessarily informative about behavior in other tasks, thus requiring the independent study of behavior in each task.

From Stanford Paper (Example Application)

FYI

Section on Healthcare & Biomedicine



Better World – Leverage Tech!

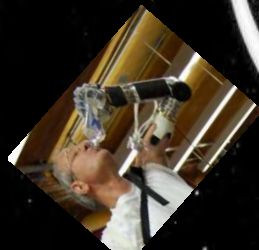
Steve Wozniak on AI: Will we be pets or mere ants to be squashed our robot overlords?

Apple co-founder Steve Wozniak predicts a 'scary' future when artificial intelligence takes

COMPUTERWORLD

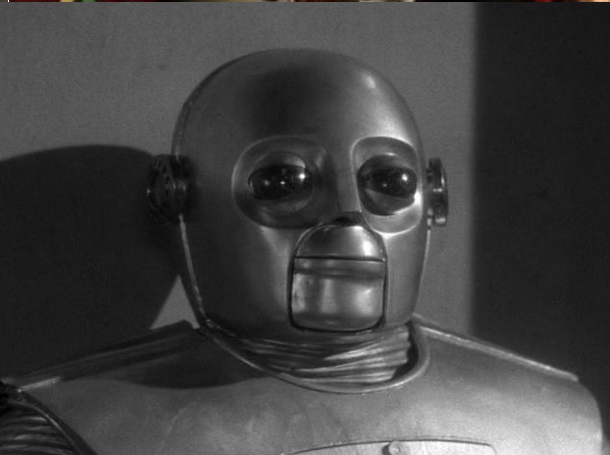
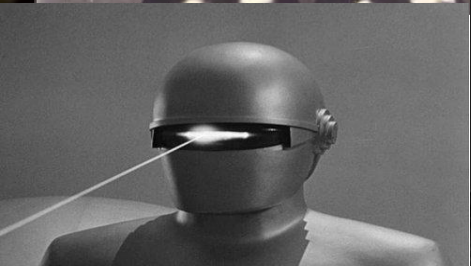
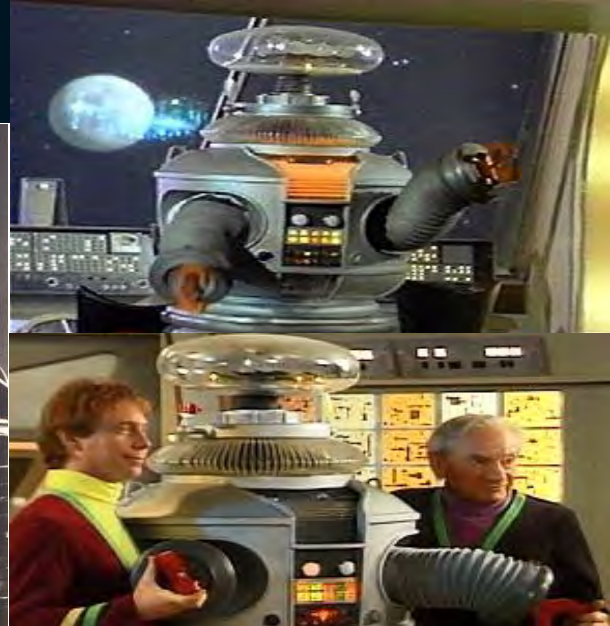
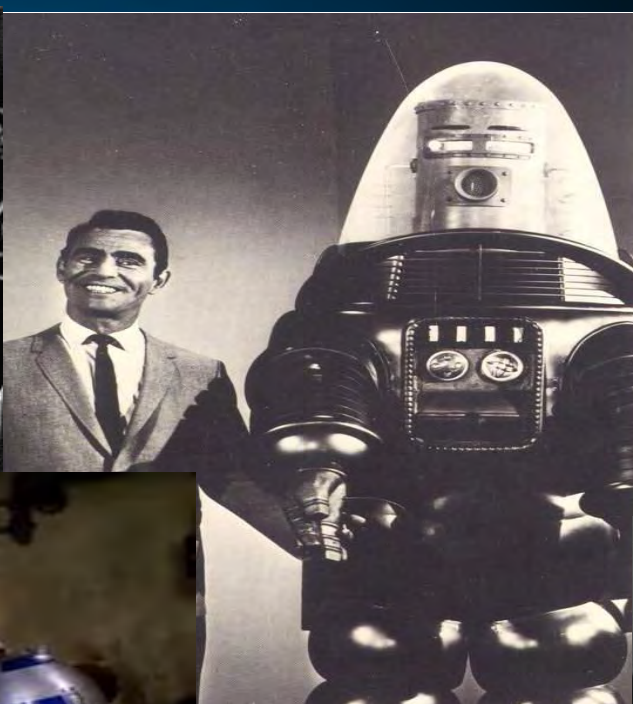
THE TWILIGHT ZONE

"You unlock this door with the key of imagination. Beyond it is another dimension - a dimension of sound, a dimension of sight, and substance. You're moving into a land of both shadow and ideas. You've just crossed over into the Twilight Zone."
-Rod Serling



E=MC²

Robots We Grew up With...



Movie Robots... Closer to the Future



MORAVEC PARADOX: WHAT'S EASY FOR MACHINES IS DIFFICULT FOR HUMANS AND WHAT'S EASY FOR HUMANS IS DIFFICULT FOR MACHINES: E.G. IT'S HARD TO BUILD A ROBOT WITH THE SENSORIMOTOR SKILLS OF A 1 YEAR OLD HUMAN



Moravec's Paradox



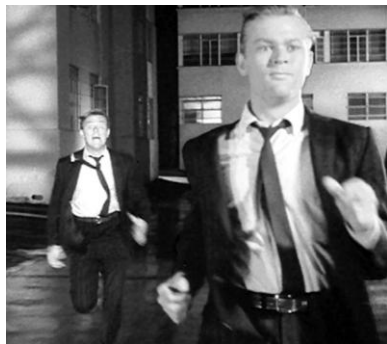
Logical thought is hard for humans and easy for computers, but picking a screw from a box of screws is an unsolved problem.



Actual Robots Now... Beyond Science Fiction



Mirror Image



Mirror Image



Prof. Hiroshi Ishiguro, right, inspects his robotic twin, Geminoid HI-1, also known as 'Gemmy,' who handles his teaching duties at Osaka University in Japan. Photo courtesy of ATR Intelligent Robotics and Communication Laboratories.



GenAI inside APEX, Sensors, Robots...etc.

Transformative Technologies: Applied to other Industries

- **Generative Artificial Intelligence**
 - APEX No Code Application Generation System
 - Autonomous Database Stores Vast Amounts AI Training Data
- **Internet of Things: Sensors & Robots**
 - Sensors Collect Vast Amounts of Training Data
 - Robots Lower Costs and Improve Product Quality & Job Quality



Use Oracle Virtual Assistant with Robots


ORACLE

SER-10
SERVICE

VIRTUAL ASSISTANT

Offer Natural, Conversational Interactions at Scale with a Smart Automated Solution

Oracle Virtual Assistant the Chatbot for Service



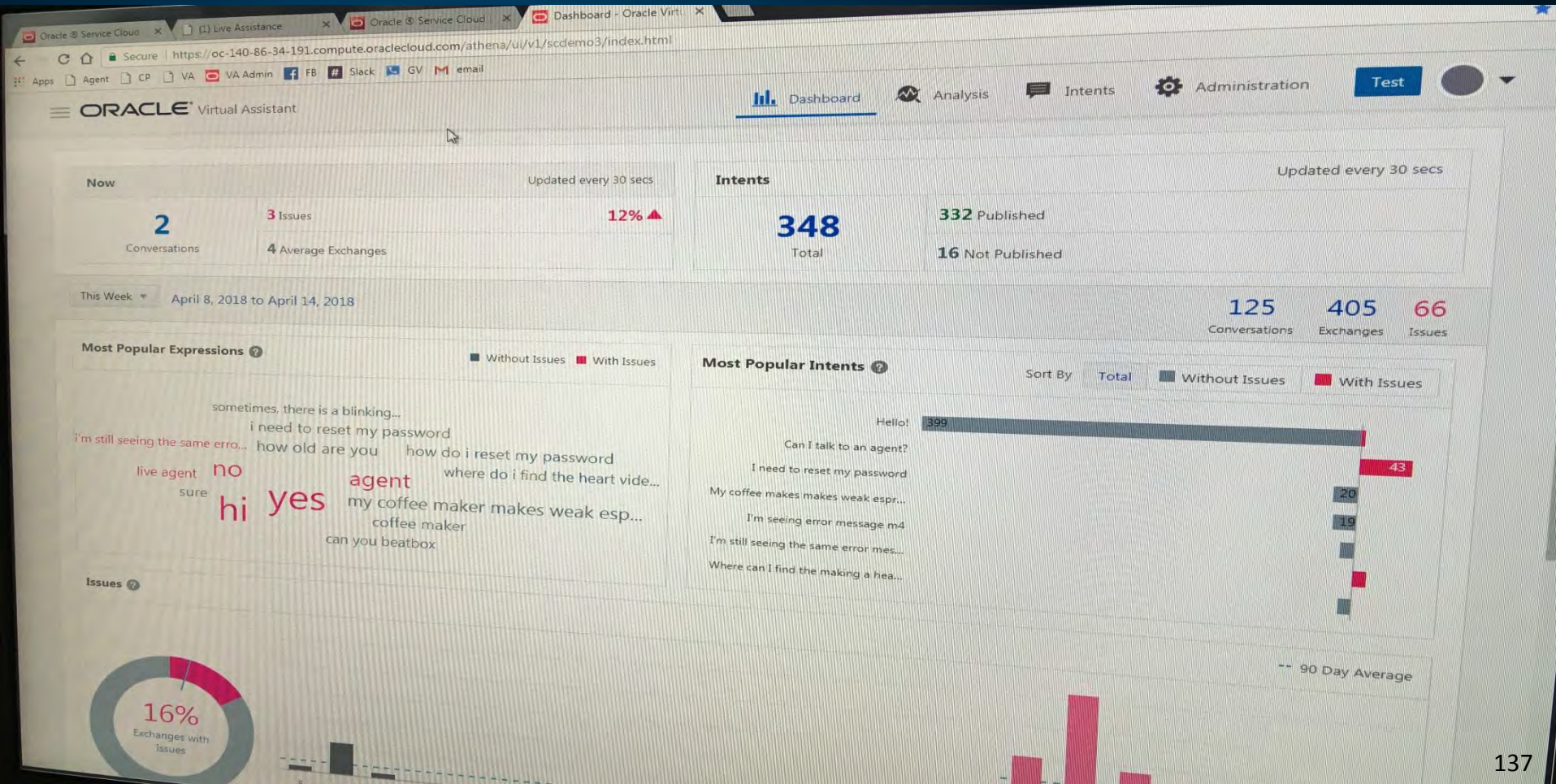
Oracle Virtual Assistant the Chatbot for Service



- SMS
- Facebook Messenger
- WeChat
- Kik
- WhatsApp
- Phone
- Alexa
- Slack
- HipChat
- Siri



Oracle Virtual Assistant Interface



Leveraging – DB, AI & Virtual Reality!



Virtual Reality



- > Immersion in virtual worlds
- > Total interaction with virtual
- > E.g. Oculus Rift

Mixed Reality



- > Virtual World integrated to reality
- > Interaction between reality and virtual
- > E.g. Microsoft HoloLens

Augmented Reality



- > Virtual on top of reality
- > Limited interaction with the virtual
- > E.g. Smartphones & Tablets

What Is XR (Extended Reality)?

XR describes virtual, augmented, and mixed-reality technologies

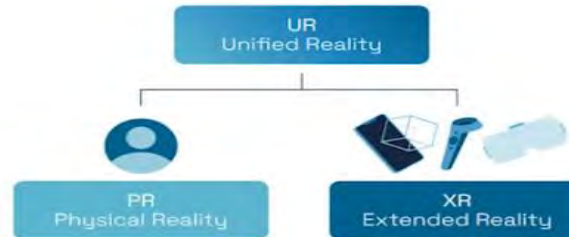
{ XR & Immersive Curiosity_

< Extended reality (or XR) is a suitcase term for the growing family of immersive technologies* that connect the bits & bytes of digital reality with the atoms and molecules of human reality to create opportunities for engagement and value creation />

*Including AR, VR, MR, METAVERSE | MULTIVERSE, 360 VIDEO, FILTERS, DIGITAL TWINS & OTHERS



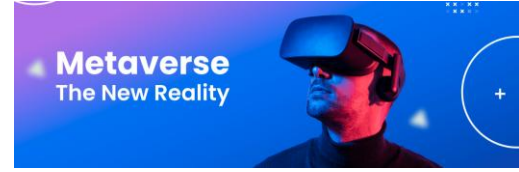
- Electricity crackles with this UR in the world it serves & was built.



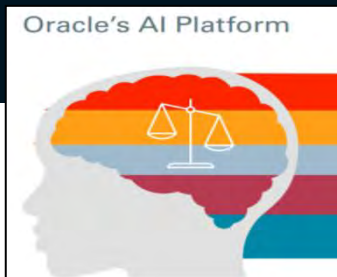
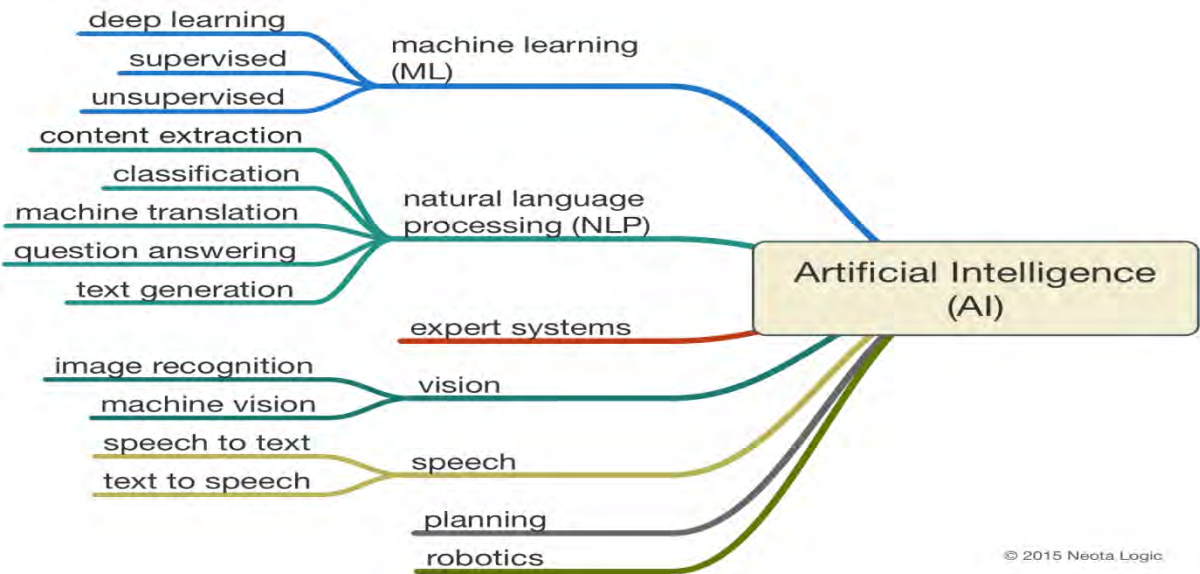
{ Unified Reality (UR)_



A World of Difference (Getting Closer)



Future Tech is in AI: You can Leverage this NOW!



© 2015 Neota Logic

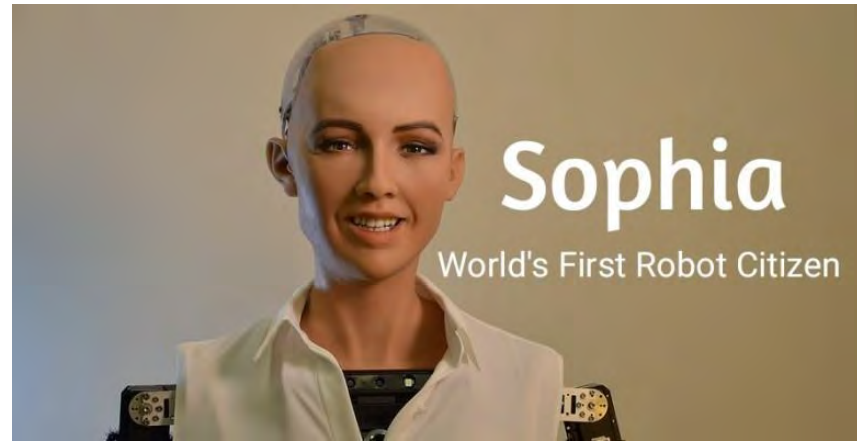


100%
of cloud applications will feature AI

>50%
of enterprise data managed autonomously

85%
of all interactions will be automated

The After Hours (*Future Sentience Issues Ahead*)



- Is now a Citizen
- Wants to get degree
- Wants to have kids

- What is Alive?
- Sentience Issues?
- Robot's Feelings?
- Ethical Issues +++



The Digital Transformation Ahead

Gerd

Digital Transformation 2000 to 2050

A historically significant
change in humanity...



Disconnected
The Dark Ages

Using
Digital

Wearing
Digital

Implanting
Digital

The Hive
Mind

Connection to the Brain – Currently Working

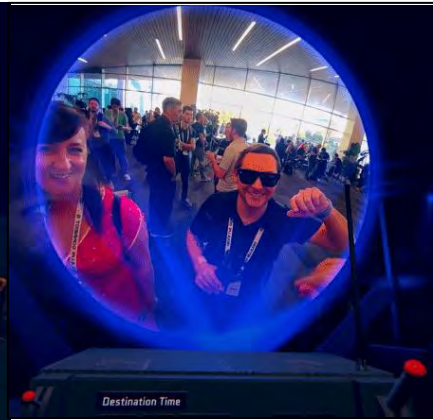


Walking Distance *(Create your past to view)*



Developer Creates VR Time Machine To Relive His Past Year

July 13, 2020 • by Bobby Carlton



Digital – How did we go from Magical to Toxic?



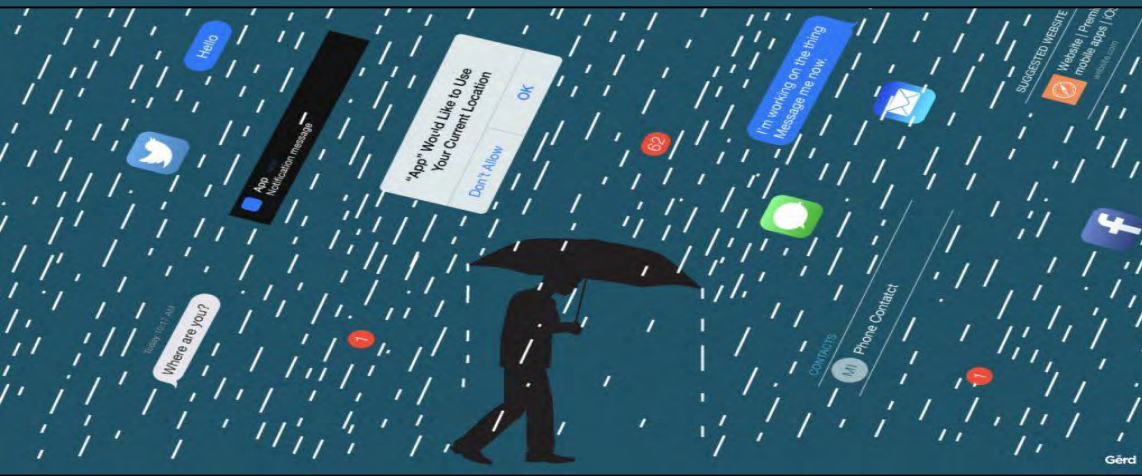
MAGIC



MANIC



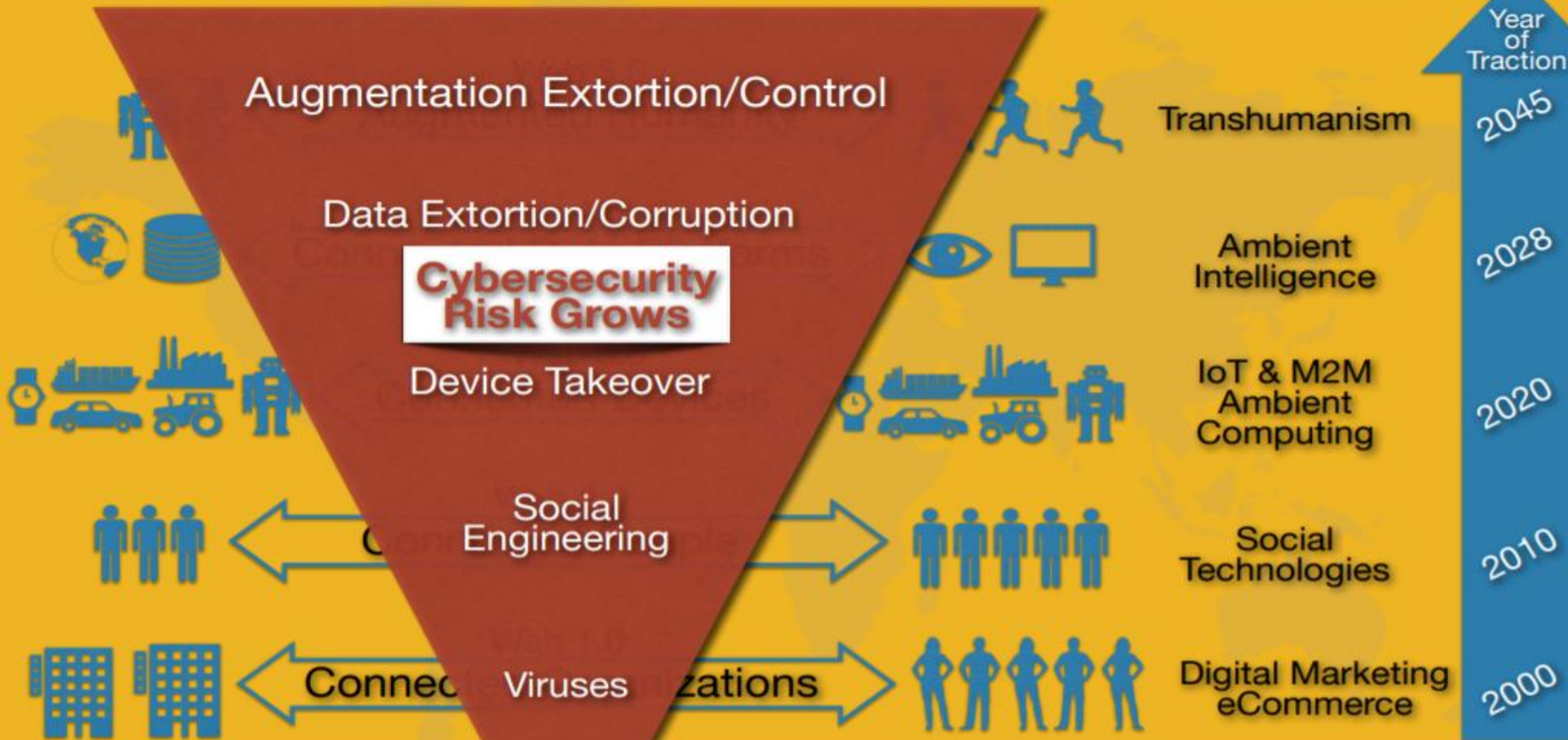
TOXIC



The Issues to Overcome moving Forward...

The Transformation of Connection

FPOV



Oracle Database Security

Built over MANY years...

1977

Database Auditing
Government customer

Native Network Encryption (Oracle7)

Strong authentication (PKI, Kerberos, RADIUS)

Database Encryption API

Virtual Private Database (8i)

Global roles

Enterprise User Security

Proxy authentication

Oracle Label Security (2000)

Client Identifier / Identity propagation

Secure application roles

Fine Grained Auditing (9i)

EM Configuration Scanning

Transparent Data Encryption

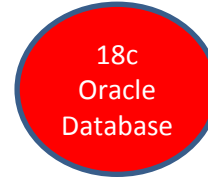
DB Security Evaluation #19

Oracle Database Vault

Oracle Audit Vault

Oracle Multi-Tenant Security (PDBs)

Encryption in the Silicon (M7)



18c: User-defined Master Encryption Key (bring your own key - software keystores)



19c: Oracle Database supplied schema-only accounts have passwords removed (not sample accounts)



21c: Global Fault-Tolerant Key Vault & Oracle Data Safe

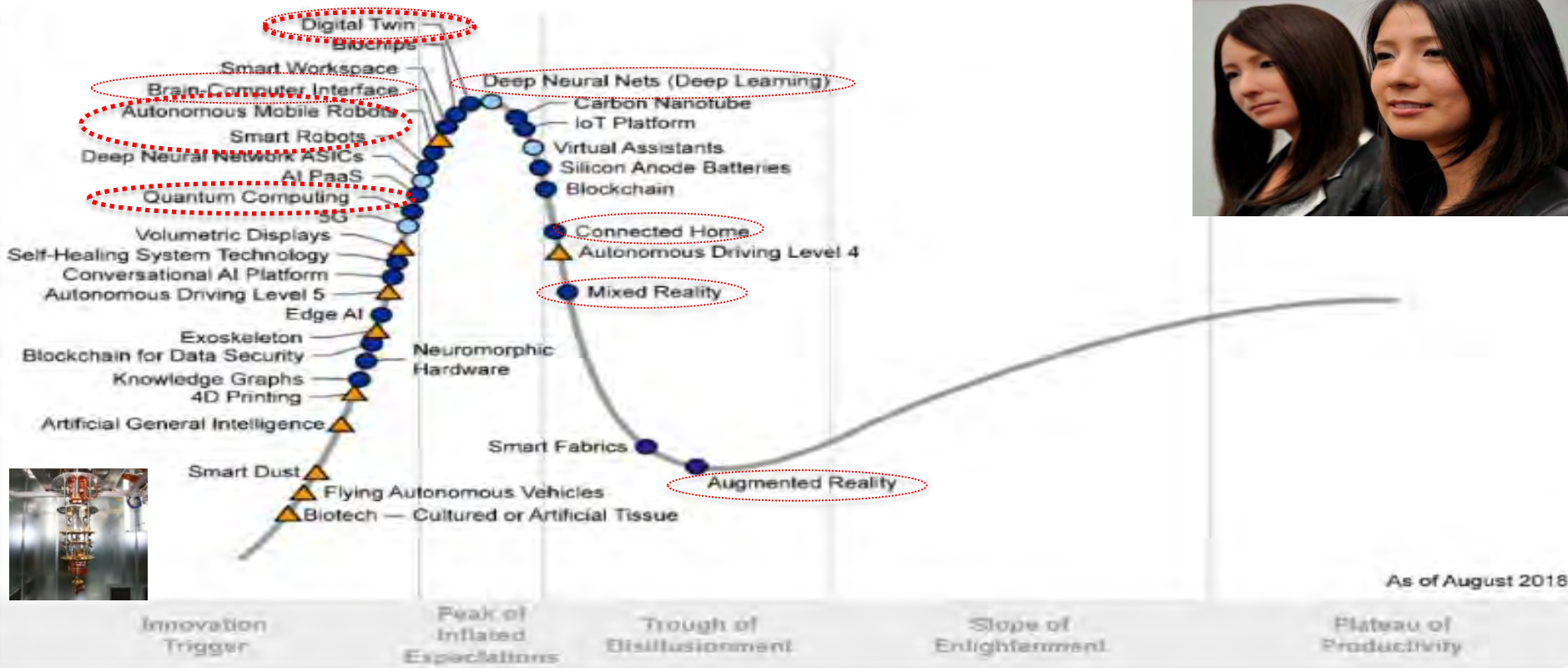


23c: In-Database SQL Firewall blocks unauthorized SQL & SQL Injection Attacks

2023 +

Gartner Hype Cycle August 2018

All about Tech Creating a New Reality



As of August 2018

Plateau will be reached:

- less than 2 years
- ◐ 2 to 5 years
- 5 to 10 years
- ▲ more than 10 years
- ⊗ obsolete before plateau

Gartner 2020 hype cycle

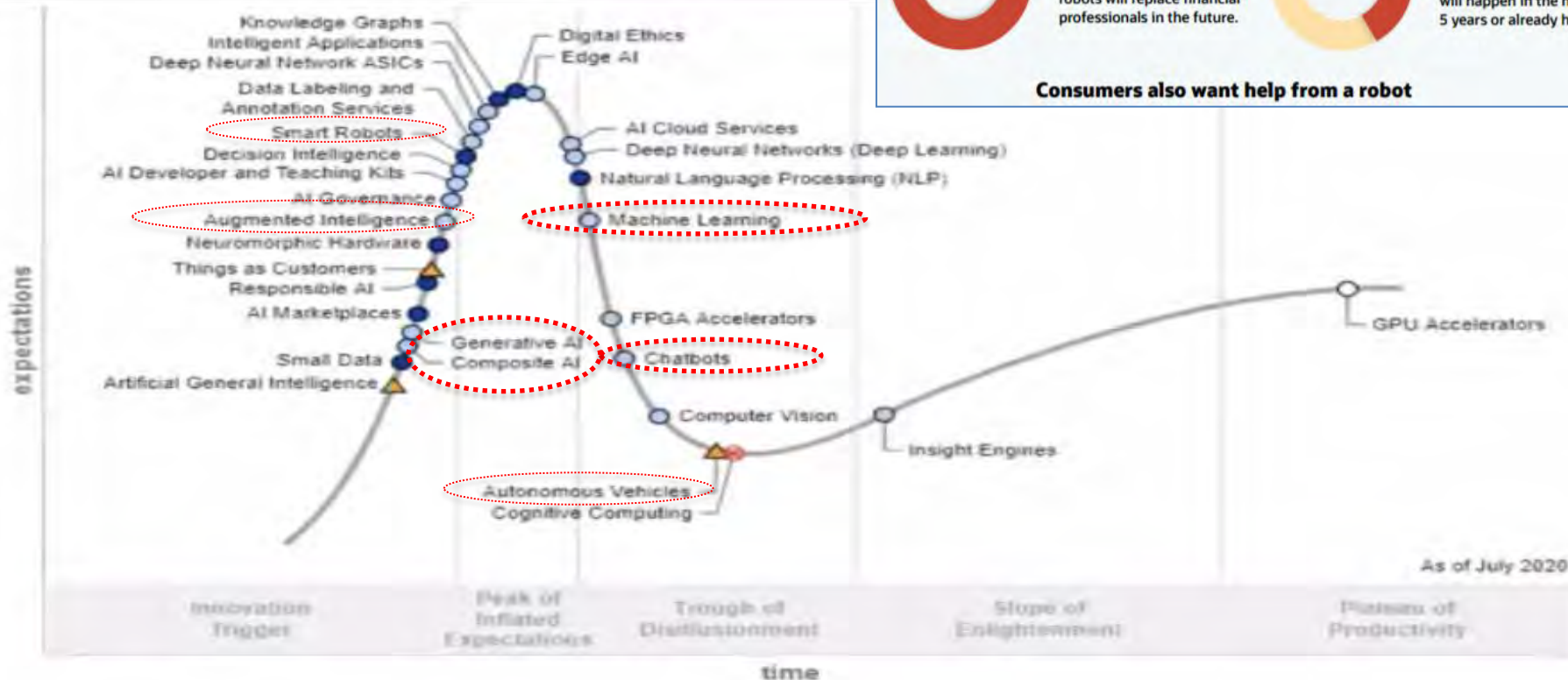


82%
of consumers believe that robots will replace financial professionals in the future.



42%
of consumers believe it will happen in the next 5 years or already has.

Consumers also want help from a robot



As of July 2020

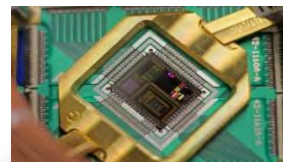
Plateau will be reached:

○ less than 2 years ● 2 to 5 years ● 5 to 10 years △ more than 10 years ⊗ obsolete before plateau

What's comes *after* the Exadata Zone?

YOU will soon be in for more...

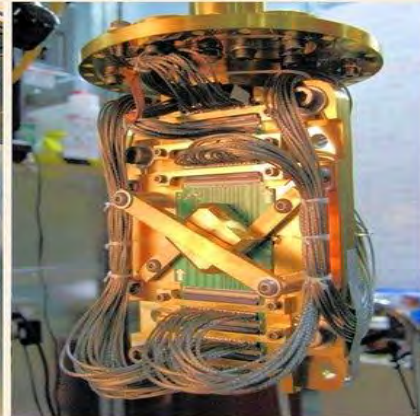
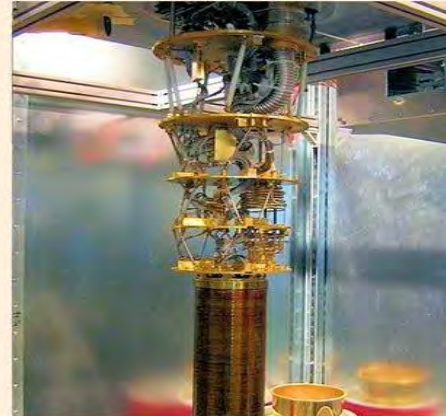
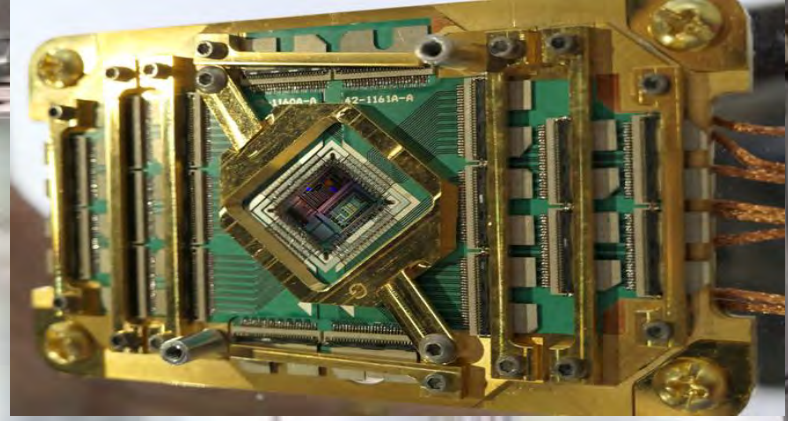
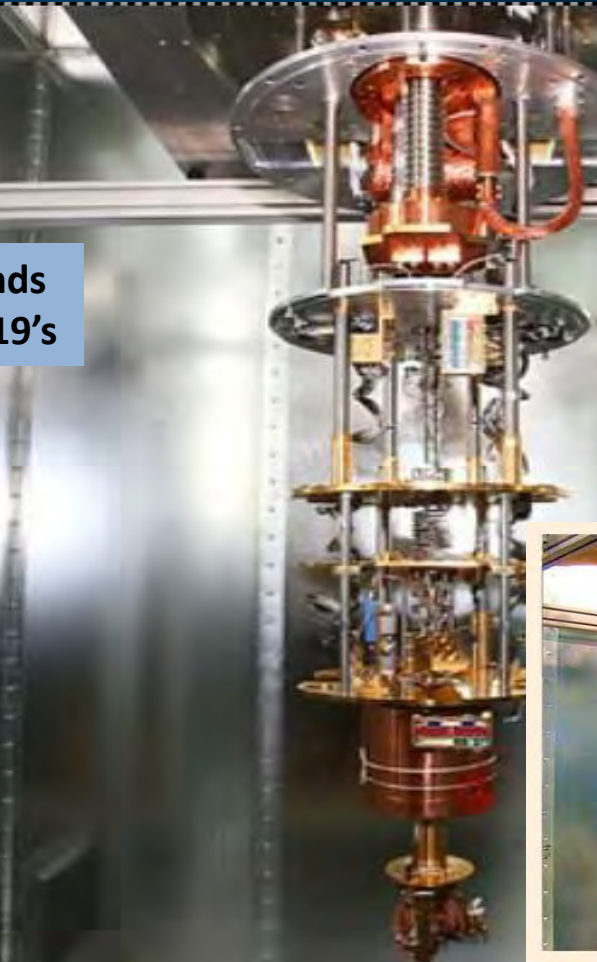
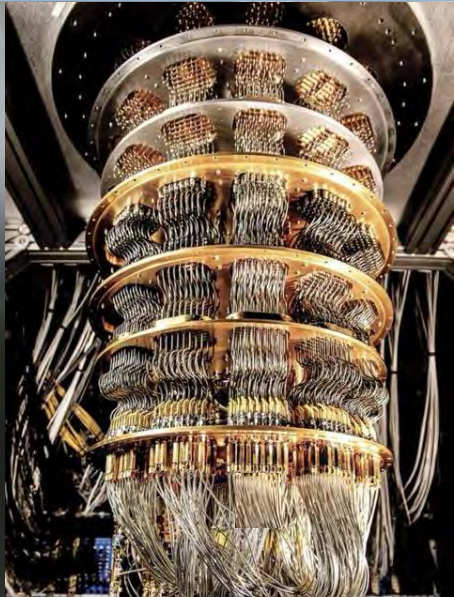
	<u>Directly Addressable</u>	<u>Indirect/Extended</u>
<u>4 Bit:</u>	16	(640)
<u>8 Bit:</u>	256	(65,536)
<u>16 Bit:</u>	65,536	(1,048,576)
<u>32 Bit:</u>	4,294,967,296	
<u>64 Bit:</u>	18,446,744,073,709,551,616	
<u>128 Bit:</u>	3.4 x 10e+38... quantum leap!	



- Qubits allow multiple states so that you can look at all of the possibilities/probabilities at one time (IBM Q is 20 qubits).
- The “**Quantum Zone**” next (Quantum Physics is incomplete – Einstein)
 - Just 512 qubits would store 512-bits of addressable memory or 2^{512} (which is well over a googol or 1 with 100 zero's after it – **a googol is about 2^{332}**).
 - Brush up on your Eigenvectors, Eigenvalues, Pauli Matrices & Grover's Algorithm
 - Create Singularity ... all atoms of a person by 2045 (I think earlier); *12-Monkeys*
 - Private universes – Is there one for each person? (Schroeder's cat – I think not)
 - Rearranging atoms to create new objects; Nanotech + Quantum Physics coming!

Quantum Computing makes ML Fast Enough!

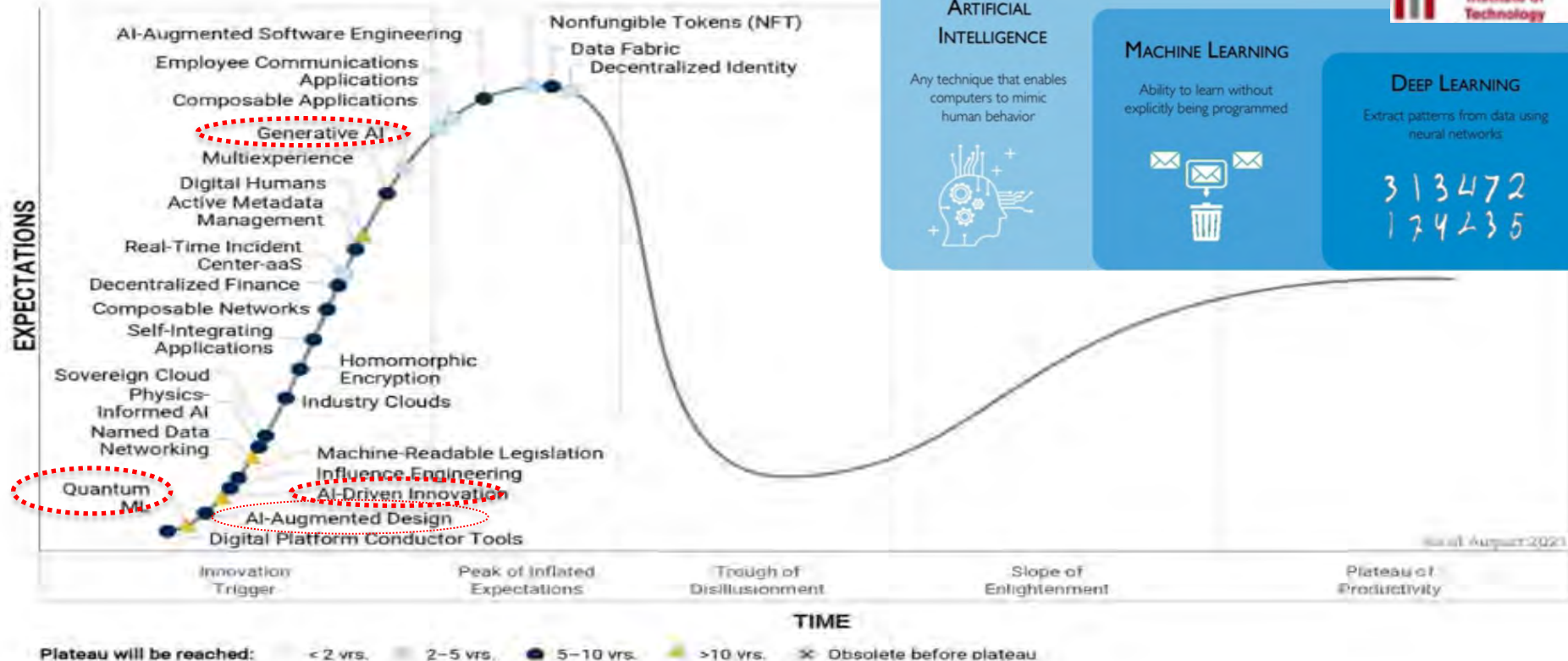
Google: 47 years to 6 seconds
241M times faster than 2019's



Gartner 2021 hype cycle for Emerging Tech



Hype Cycle for Emerging Technologies, 2021



3 Types of Artificial Intelligence

Artificial Narrow Intelligence (ANI)



Stage-1

Machine Learning

- ▶ Specialises in one area and solves one problem



Siri



Alexa



Cortana

Artificial General Intelligence (AGI)



Stage-2

Machine Intelligence

- ▶ Refers to a computer that is as smart as a human across the board

2023:
You're Here

Artificial Super Intelligence (ASI)



Stage-3

Machine Consciousness

- ▶ An intellect that is much smarter than the best human brains in practically every field

Final Thoughts... world changing fast!

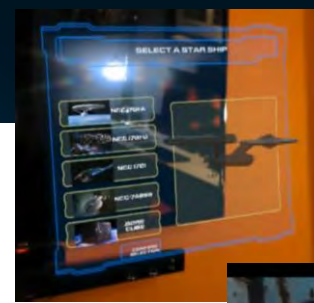
“Those who use things of the world should not become attached to them. For the world in its present form is passing away.”

1 Corinthians 7:31



Star Trek

- Communicator - Motorola Flip Phone
- Phaser - EPM or Stun Gun / Taser
- Tablets (Medical) - Tablet Computers
- Tricorders - Many Medical Devices (below)
- Translators - Google Translate (others)
- Tractor Beam - MIT has it in concept
- Telepresence - Zoom Virtual Conferences
- Geordi's Visor - Robotic Eyes / Implants
- Communicator Badges - Many Security Badges
- Food Replicator - 3D Printer
- Holodeck - VR
- Teleportation - Quantum Entanglement
- Big Screen TV - Everyone has it
- Hands Free Phone - Bluetooth headset / Airpods



Hololens VR Apps Above



Microsoft working on Holodeck



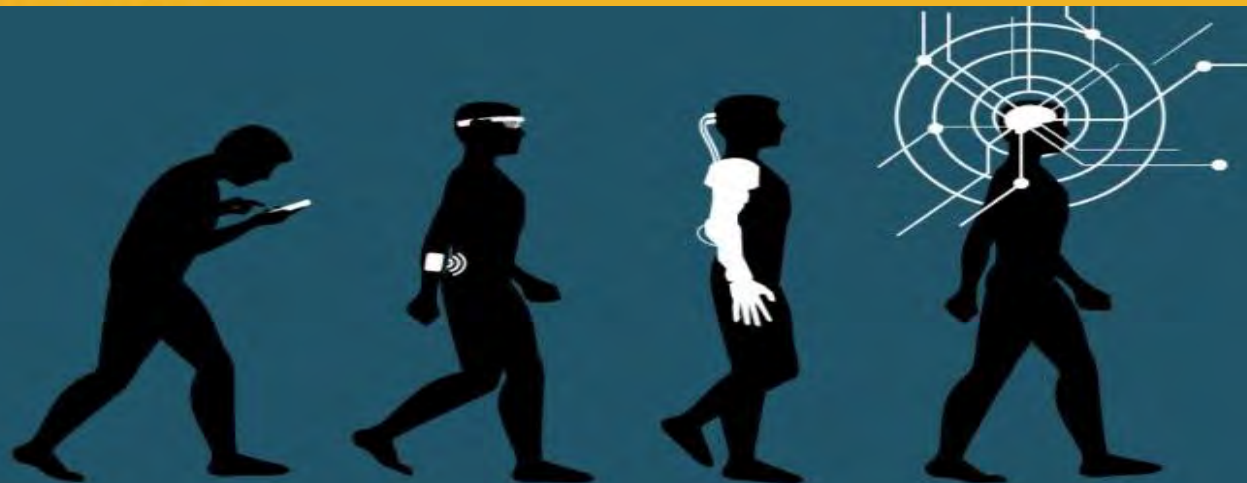
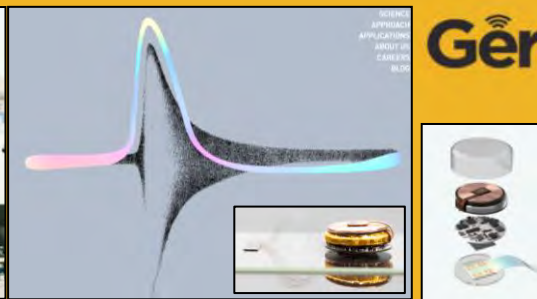
MouthLab



The Digital Transformation Ahead

Digital Transformation 2000 to 2050

A historically significant
change in humanity...



Disconnected
The Dark Ages

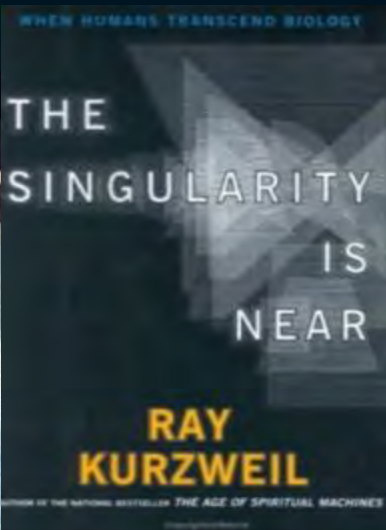
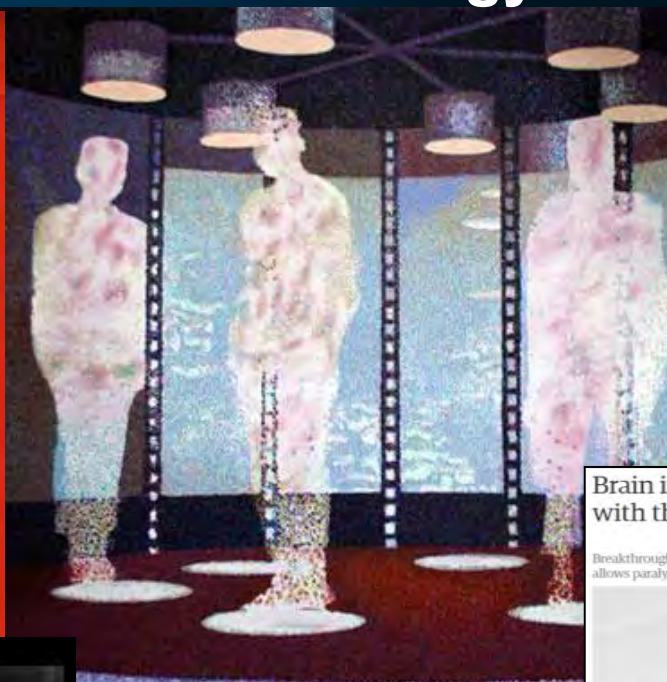
Using
Digital

Wearing
Digital

Implanting
Digital

The Hive
Mind

World now changing Faster! The Singularity: Humans Transcend Biology



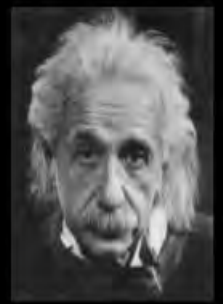
Brain implant controls robotic arm - with the power of thought

Breakthrough as neuro-prosthetic device implanted in "higher" brain region, allows paralysed man to control robotic arm just by thinking about it



Quantum Entanglement

"Spooky action-at-a-distance"
(A. Einstein)



A neuro-prosthetic device implanted in Eric Soren's brain allowed him to drink alcohol for the first time in his life. Photographs: Science Photo and Christine E. Hoff

Summary – Apply Technology & W.I.N.

SUBMITTED
FOR YOUR
APPROVAL

- ❑ Economic Potential of GenAI
- ❑ Economic Impact of Robots
- ❑ ML & Oracle - Overview & ADB
- ❑ Oracle Generative AI Coming
- ❑ Transformers, ChatGPT & Foundation Models
- ❑ Oracle Vector DB, GenAI, & Cohere
- ❑ Machine Learning Future, Robots & VR



Quantum Entanglement

“Spooky action-at-a-distance”
(A. Einstein)



FYI

Black Mirror out Dec. 2011- Dystopian



- Nosedive: (Social Meeting Addiction & Ratings)



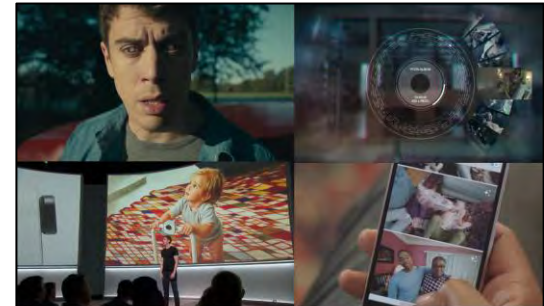
- Hated in the Nation (Cancel Culture)



- Be Right Back (ETER9)



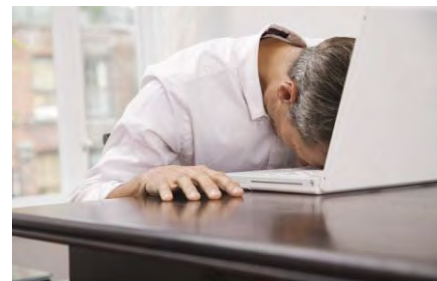
- The Entire History of You (Kapture / Google Glass / Google Clips)



- Metalhead (Robot Dog)



A Thing About Machines - People Frustrated



HELPDESK

HAVE U TRIED TURNING IT OFF
AND ON AGAIN?

Final Thoughts... Catch your Ride!



“Things may come to those who wait, but only the things left by those who hustle.”

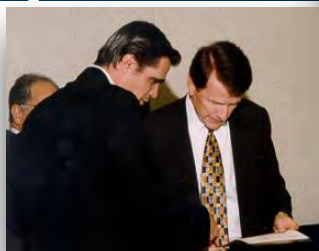
— Abraham Lincoln



Oracle never caught from behind

Oracle's 46th Anniversary in 2023

- Great Sales/Marketing
- Great Database
- Applications Leader
- BI Leader
- In the lead except Cloud
- **GAME OVER**
- Hardware/Software Engineering!
- Have Everything to Win in Cloud + AI!



 **Richard Niemiec**
Chief Innovation Officer, Viscosity North America
2d • 🌐

Dr. DBA, Ken Jacobs, He will be missed!! Oracle's best advocate of User Groups; Kind, Giving, Positive, Honest, Humble & Smartest Guy in the Room! Sadly, Ken suffered a pulmonary embolism and passed away last Friday at the age c ...see more

 Who is the Dr. pictured below?



Dr. DBA – Ken Jacobs (in the middle)

🔗 🗨️ 📍 104 • 43 comments

In Memory of Ken Jacobs, Dr. DBA & Joel Kallman

Join Thousands Of Happy Customers And End The Frustration In Tuning Oracle Databases

Our services give you the training you need so your Oracle database runs faster and more efficiently. And, you get the credit!



MEMBERSHIPS



TRAINING



FREE TOOLS

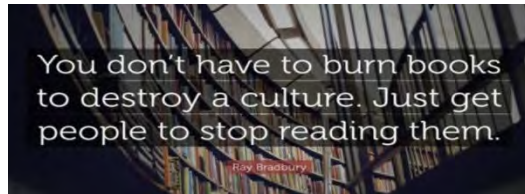


BOOKS

References

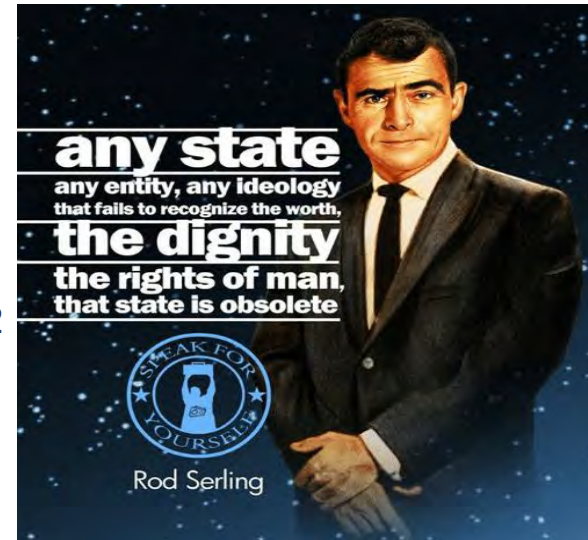


- The Emerging Technology Roadmap, Scott Klososky
- Futurist Gerd Leonhard The Futures Agency, Technology vs. Humanity, Gerd Leonhard, oracle.com & Juan Loiza / Doug Hood presentations, amazon.com, smartcitiescouncil.com, youtube.com, business coach, libelium.com, monetate, en.wikipedia.org, Netflix, Black Mirror, hometoys.com, FPOV, huffingtonpost.com, theguardian.com, nationalgeographic.com, newscientist.com, enswmu.blogspot.com, dailymail.co.uk, FutureRobot, theguardian.com, thinkhealthwireless.blogspot.com, ge.com, cmswire.com, runningsupplement.co.uk, quickmeme.com, nike.com, thisiswhyimbroke.com, businessinsider.com, slideshare.com, forrester.com, spiceworks.com, mwaintel.com, humancapitalist.com, wired.com and Wired Magazine, shodanhq.com, developer.nokia.com, extremetech.com, Getty Images, Dr. Quantum, Leo Kouwenhoven, Telecom Tech News, BBC Radio, Actimage, MIT & any other company products are the property of their respective companies.



Copyright Information

- Neither the Viscosity nor the author guarantee this document to be error-free. Please provide comments/questions to rich.niemiec@viscosityna.com & richniemiec@gmail.com. I'm always looking to improve!
- Rich Niemiec ©2023. This document cannot be reproduced without expressed written consent from Rich Niemiec, but may be reproduced or copied for presentation and conference use.
- References include Rich Niemiec's Exadata Presentation & Oracle 12cR2 Database Performance Tuning Tips & Techniques book, www.oracle.com, en.wikipedia.org, slashgear.com, gifsoup.com, www.amazon.com, Tech Crunch, www.rolta.com, The Twilight Zone, Information Week, Gartner, Computerworld, Quest, Data and Technology Today, Forbes, Quest, Orbit, Computer Weekly, Redmond, Database Trends & Applications, dsp, Dataversity, zdnet, DBVisit, Steve Jones, Kerry Osborne, Julian Dontcheff, Accenture, Quora, Brent Ozar, & Oracle OpenWorld



**THERE IS A FIFTH DIMENSION
BEYOND THAT WHICH IS KNOWN TO
MAN. IT IS A DIMENSION AS VAST
AS SPACE AND AS TIMELESS AS INFINITY.**

Getting Started—Oracle ML/AI Resources

ORACLE

Oracle Advanced Analytics Overview Information

- [Oracle Machine Learning Newest Features and Road Map.pptx](#) presentation
- Blog post: [Simple Guide to Oracle's Machine Learning and Advanced Analytics](#)
- Oracle Advanced Analytics Public [Customer References](#)
- [Oracle's Machine Learning and Advanced Analytics Data Management Platforms](#) white paper on OTN



YouTube recorded Oracle Advanced Analytics Presentations and Demos, White Papers

- [Oracle's Machine Learning & Advanced Analytics 12.2 & Oracle Data Miner 4.2 New Features YouTube video](#)
- [Library of YouTube Movies on Oracle Advanced Analytics, Data Mining, Machine Learning \(7+ "live" Demos e.g. Oracle Data Miner 4.0 New Features, Retail, Fraud, Loyalty, Overview, etc.\) Overview YouTube video of Oracle's Advanced Analytics and Machine Learning](#)

ORACLE
UNIVERSITY

Getting Started/Training/Tutorials

- [OAA/Oracle Data Miner Workflow GUI Online \(free\) Tutorial Series](#) on OTN
- [OAA/Oracle R Enterprise \(free\) Tutorial Series](#) on OTN
- [Try the Oracle Cloud Now!](#)
- [Getting Started w/ ODM blog entry](#)
- [New OAA/Oracle Data Mining 2-Day Instructor Led Oracle University course](#)
- [Oracle Data Mining Sample Code Examples](#)

ORACLE Help Center

Additional Resources, Documentation & OTN Discussion Forum

- [Oracle Advanced Analytics Option on OTN page](#)
- [OAA/Oracle Data Mining on OTN page, ODM Documentation](#)
- [OAA/Oracle R Enterprise page on OTN page, ORE Documentation](#)
- [Oracle SQL based Basic Statistical functions on OTN](#)
- [Oracle R Advanced Analytics for Hadoop \(ORAHA\) on OTN](#)



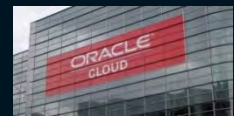
[Analytics and Data Summit](#) , All Analytics, All Data, No Nonsense.

March 12-14, 2019, Redwood Shores, CA

Rich's Overview @richniemiec



ORACLE
ACE Director



- Chief Innovation Officer, Viscosity North America: rich.niemiec@viscosityna.com
- Former CEO of TUSC
 - Inc. 500 Company (Fastest Growing 500 Private Companies)
 - 10 Offices in the United States (U.S.); Based in Chicago
 - Oracle Advantage Partner in Tech & Applications
- Former President Rolta TUSC & President Rolta EICT International
- Author (5 Oracle Best Sellers – #1 Oracle Tuning Book for over a Decade):
 - Oracle Performing Tips & Techniques (Covers Oracle7 & 8i)
 - Oracle9i Performance Tips & Techniques
 - Oracle Database 10g Performance Tips & Techniques
 - Oracle Database 11g Performance Tips & Techniques
 - Quick Start Guide to Oracle Query Tuning (2015)
 - Oracle Database 12cR2 Performance Tips & Techniques
- Former President of the International Oracle Users Group
- IOUG Top Speaker in 1991, 1994, 1997, 2001, 2006, 2007
- MOUG Current President & Top Speaker Twelve Times
- National Trio Achiever award - 2006
- Oracle Certified Master & Oracle Ace Director
- Chris Wooldridge Award – 1998, 2012
- Chicago Entrepreneur Hall of Fame - 1998
- MIT Machine Learning & AI
- Purdue Outstanding Electrical & Computer and Engineer Alumni – 2007
- Board Member – TEC, Entrigna, Ask DB Experts
- E&Y Entrepreneur of Year & National Hall of Fame - 2001



Quick FREE notes

Send email to (for slides):

hello@viscosityna.com

(richniemiec@gmail.com)

@richniemiec - twitter



CONF42



Rich Niemiec

Chief Innovation Officer
at Viscosity North America

Understand & Where to use
AI & Machine Learning 101

Conf42 DevSecOps 2023

Thursday • November 30th • 5PM GMT



CONF42