



Business Intelligence

Ezequiel Muñoz Krsulovic
Diciembre 2020

Ezequiel T. Muñoz Krsulovic

Ingeniero Civil Industrial, UChile
Magister en Ingeniería de Negocios con TI, UChile
Diplomado en Modelamiento Predictivo y Machine Learning, PUCV

Gerente Analytics, Forus
Director y Profesor Diploma en Gerencia TIC, UAI
Profesor Magister en Ingeniería Industrial, UAI
Profesor Magíster en Ingeniería de Negocios con TI, UChile
Director Analítica para la Minería, Alaya Digital Solutions
Consultor y Relator Internacional (AR, BO, BR, CL, CO, EC, PY, PE)

etmunoz@gmail.com

Skype: @ezequieltomas

ICQ: 104098411

<https://www.linkedin.com/in/ezequielmunoz/>



Tópicos

¿Qué es BI?

BI como fuente
de valor

Enfoques
metodológicos

Capacidades
requeridas

Los problemas
de los datos

Machine
Learning

Mercado:
herramientas /
arquitecturas

A person in a dark suit and tie is holding a tablet horizontally. Overlaid on the person's chest is a glowing blue graphic of a brain composed of circuit lines. The background is dark with out-of-focus blue circles.

¿Qué es BI?

#bienelmundo

Business intelligence (BI) is an umbrella term that includes the **applications, infrastructure** and **tools**, and **best practices** that enable access to and analysis of information to improve and optimize decisions and performance.

Gartner®

Hoy existen diferentes aproximaciones

Business

INTELLIGENCE

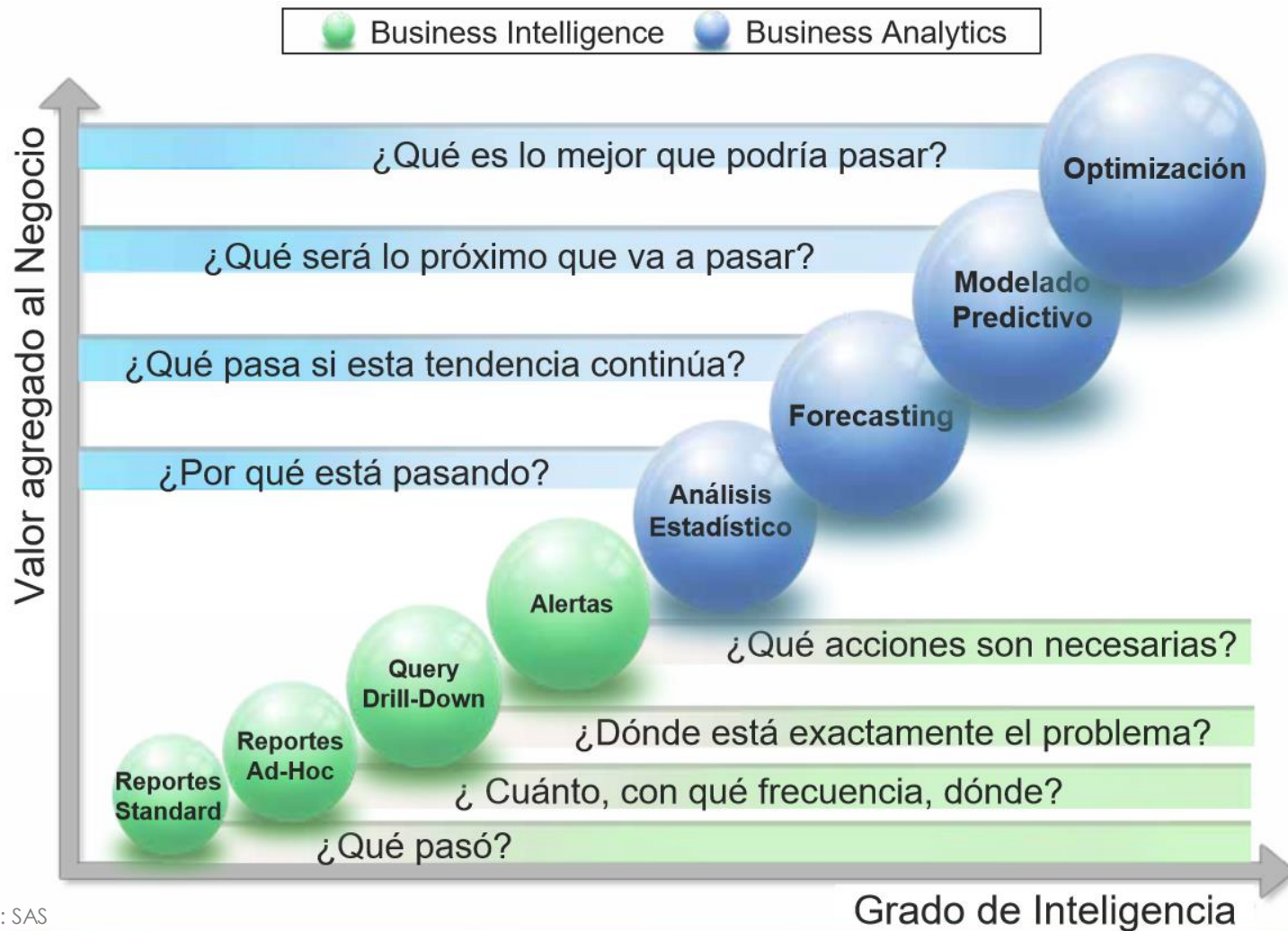
VS

ANALYTICS

Business analytics is comprised of **solutions** used to build **analysis models** and **simulations** to create scenarios, understand realities and predict future states.

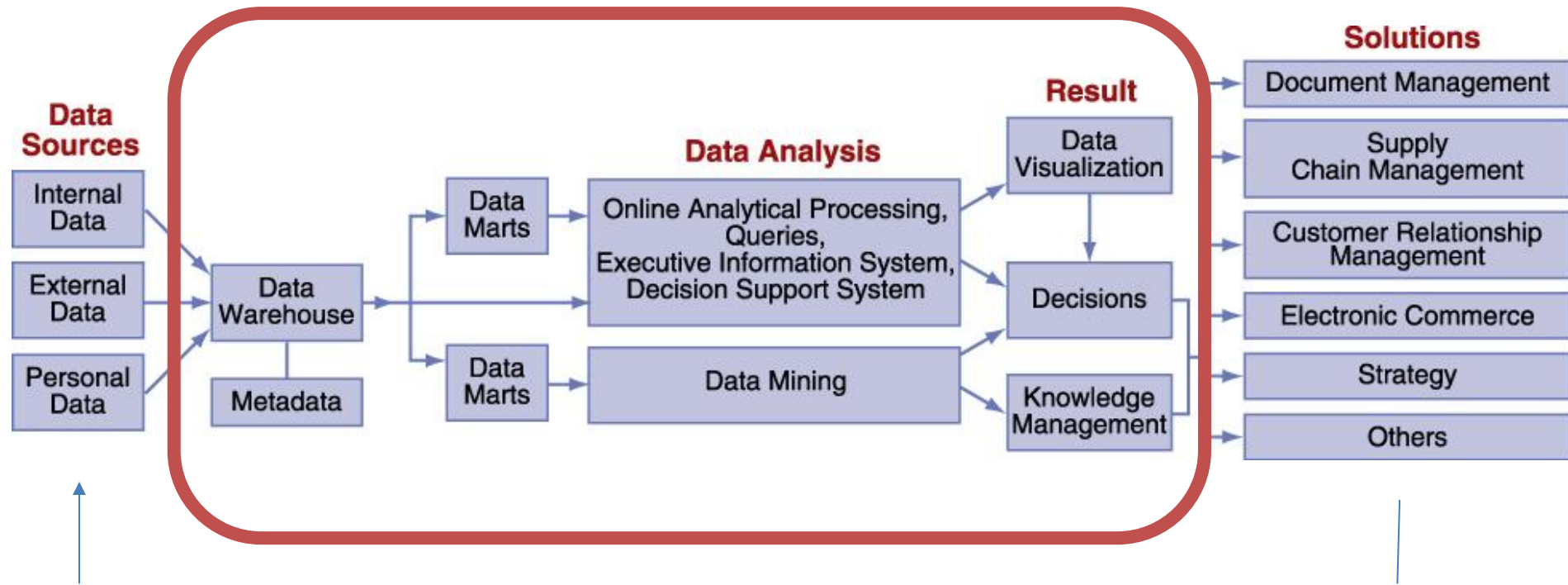
Gartner®

Lo podemos observar en forma integrada como...



Si hablamos sólo de Bl...

BI/BA se centra principalmente en:



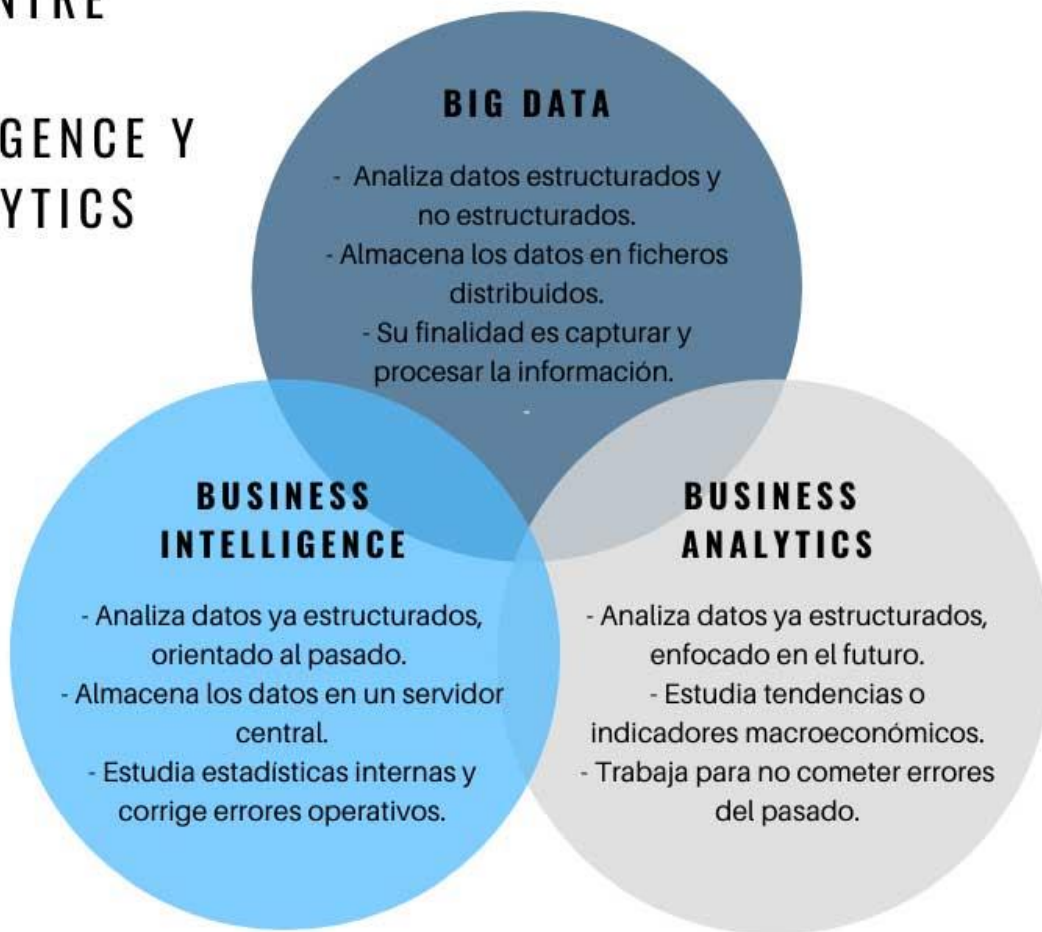
Pero hay un tema que tenemos que tocar...

Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.

Gartner®

Entonces... resumiendo...

DIFERENCIAS ENTRE BIG DATA BUSINESS INTELLIGENCE Y BUSINESS ANALYTICS





Características de BI/BA

¿Qué significa entonces
adentrarme a un BI/BA?



Desarrollo de
Capacidades
Diversas

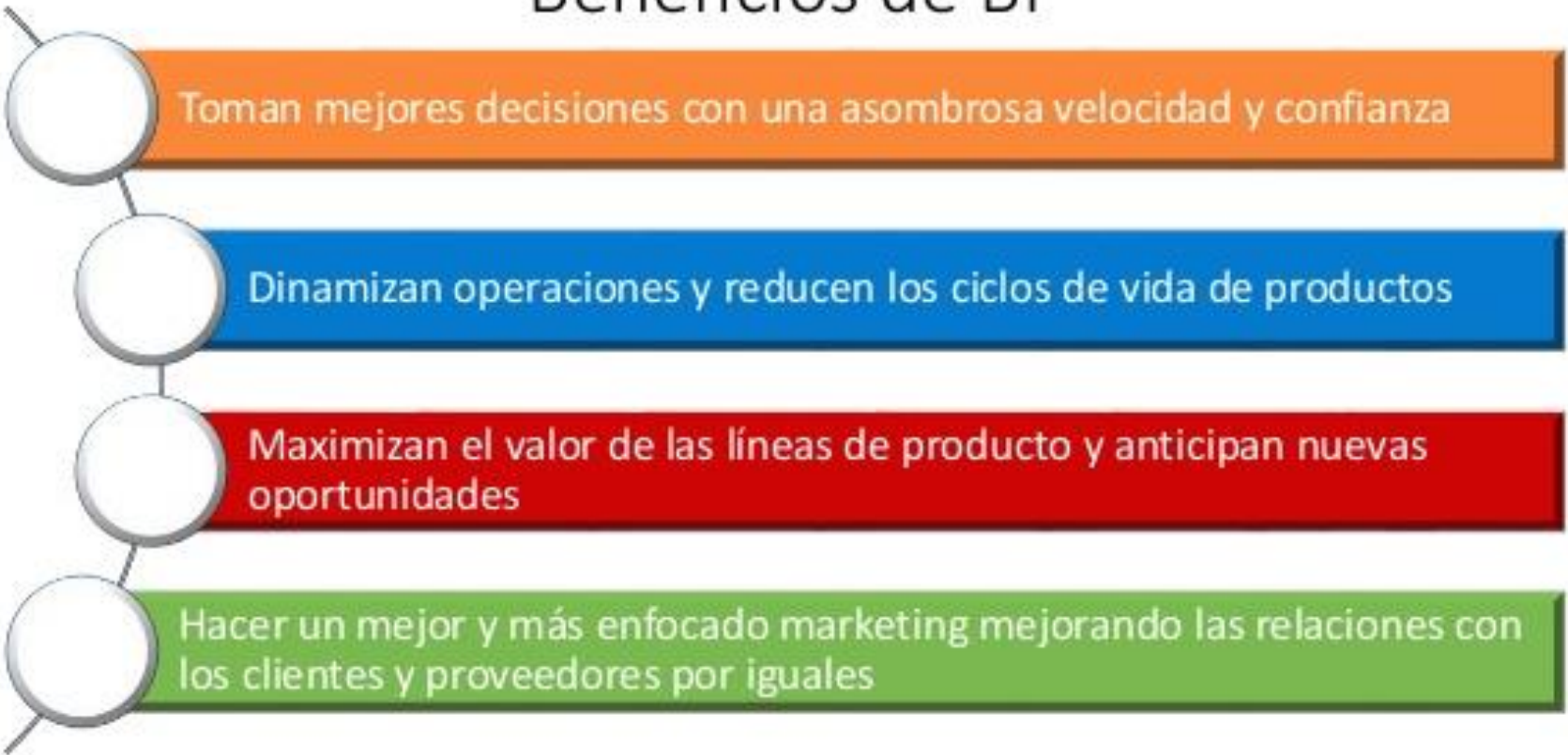


¿Por qué BI?

En el artículo “Business Intelligence Definition and Solutions” de Ryan Mulcahy (en E.E.U.U.):

- Las cadenas de restaurantes como Hardee's, Wendy's, Ruby Tuesday y T.G.I. Friday's usan BI para tomar decisiones estratégicas, como qué nuevos productos añadir al menú, qué platos quitar o qué locales deben cerrar.
- Harrah's utiliza el business intelligence tanto para construir mega-casinos hasta para analizar la fidelidad de los clientes.
- Amazon se basa en la inteligencia de negocios para llevar a cabo cambios comerciales.
- Capital One lleva a cabo más de 30.000 experimentos anuales para identificar a los clientes solventes y ofrecerles tarjetas de crédito.

Beneficios de BI



Toman mejores decisiones con una asombrosa velocidad y confianza

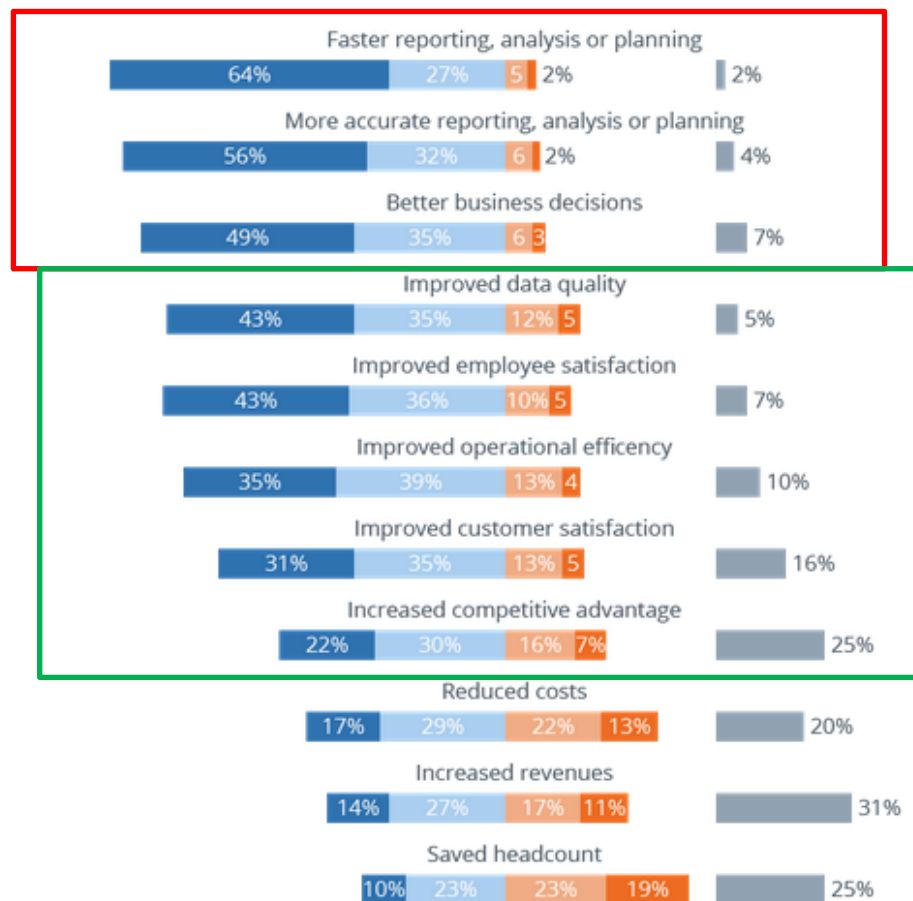
Dinamizan operaciones y reducen los ciclos de vida de productos

Maximizan el valor de las líneas de producto y anticipan nuevas oportunidades

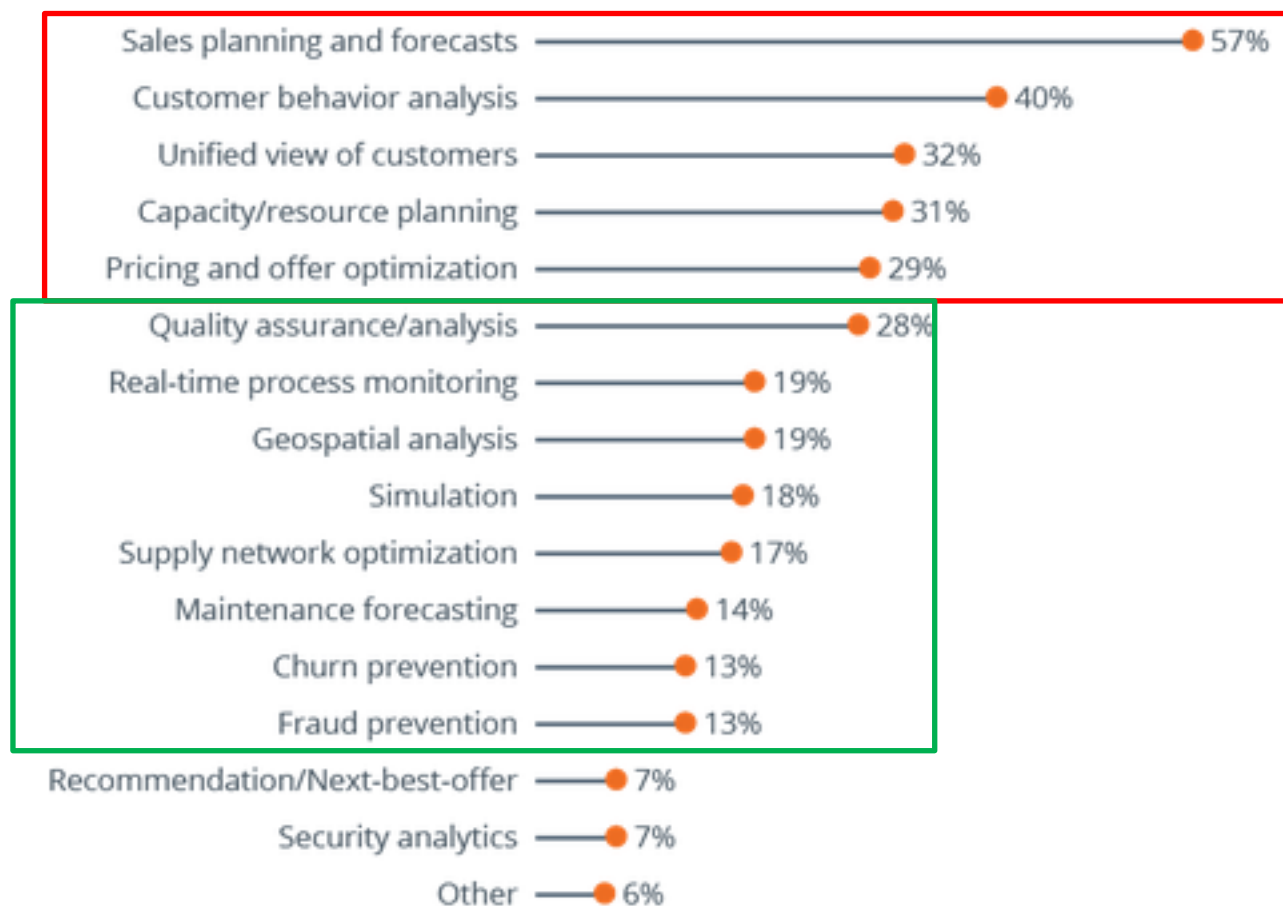
Hacer un mejor y más enfocado marketing mejorando las relaciones con los clientes y proveedores por iguales

Pero, qué ha motivado a las
empresas a abordar a un BI/BA

Good Moderate Low Not achieved Don't know

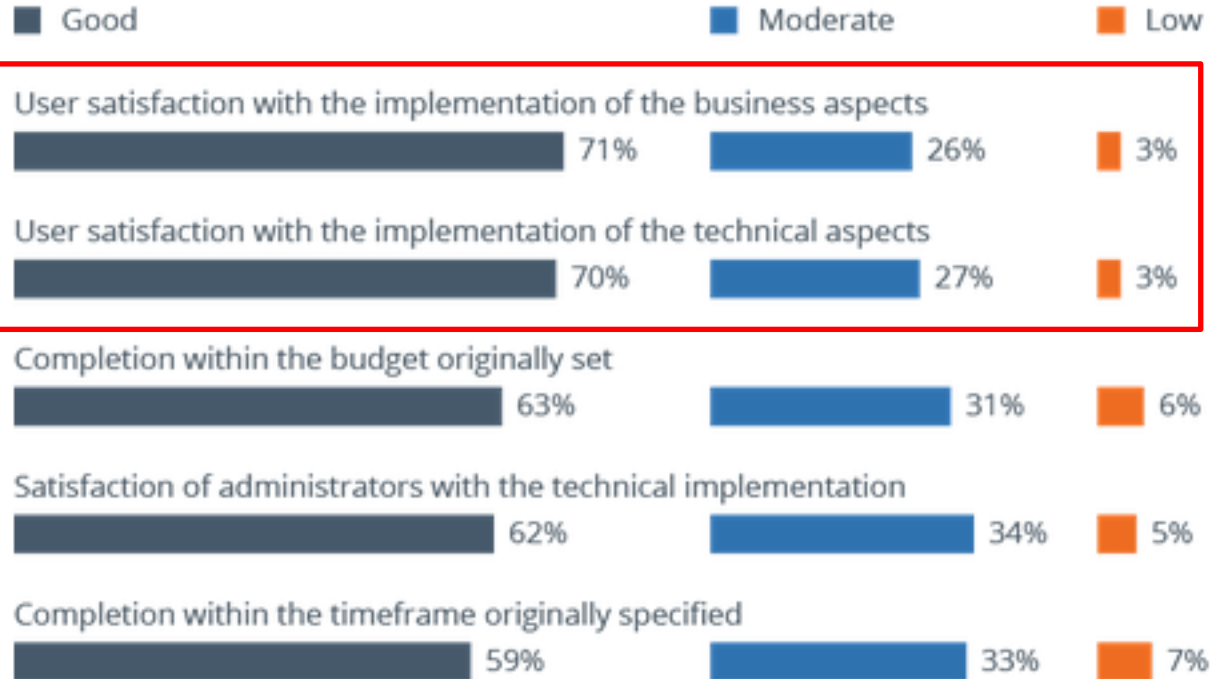


To what level have you achieved the following benefits with BI? (n=2,618)



In which BI use cases is your company investing in? (n=2,210)

Y, cómo ha sido el resultado de
estos proyectos de BI/BA



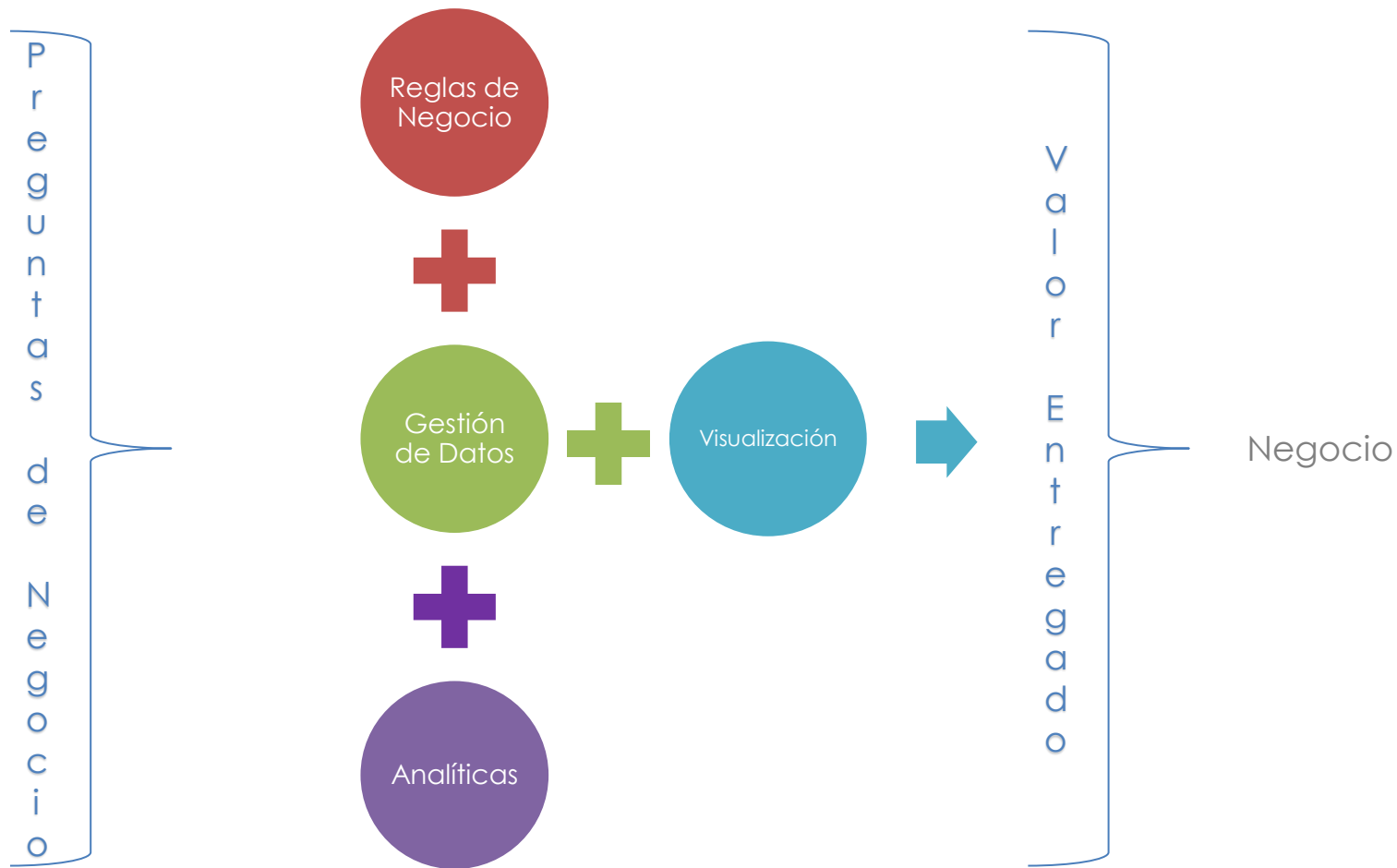
Satisfaction with aspects of BI projects (n=2,579)



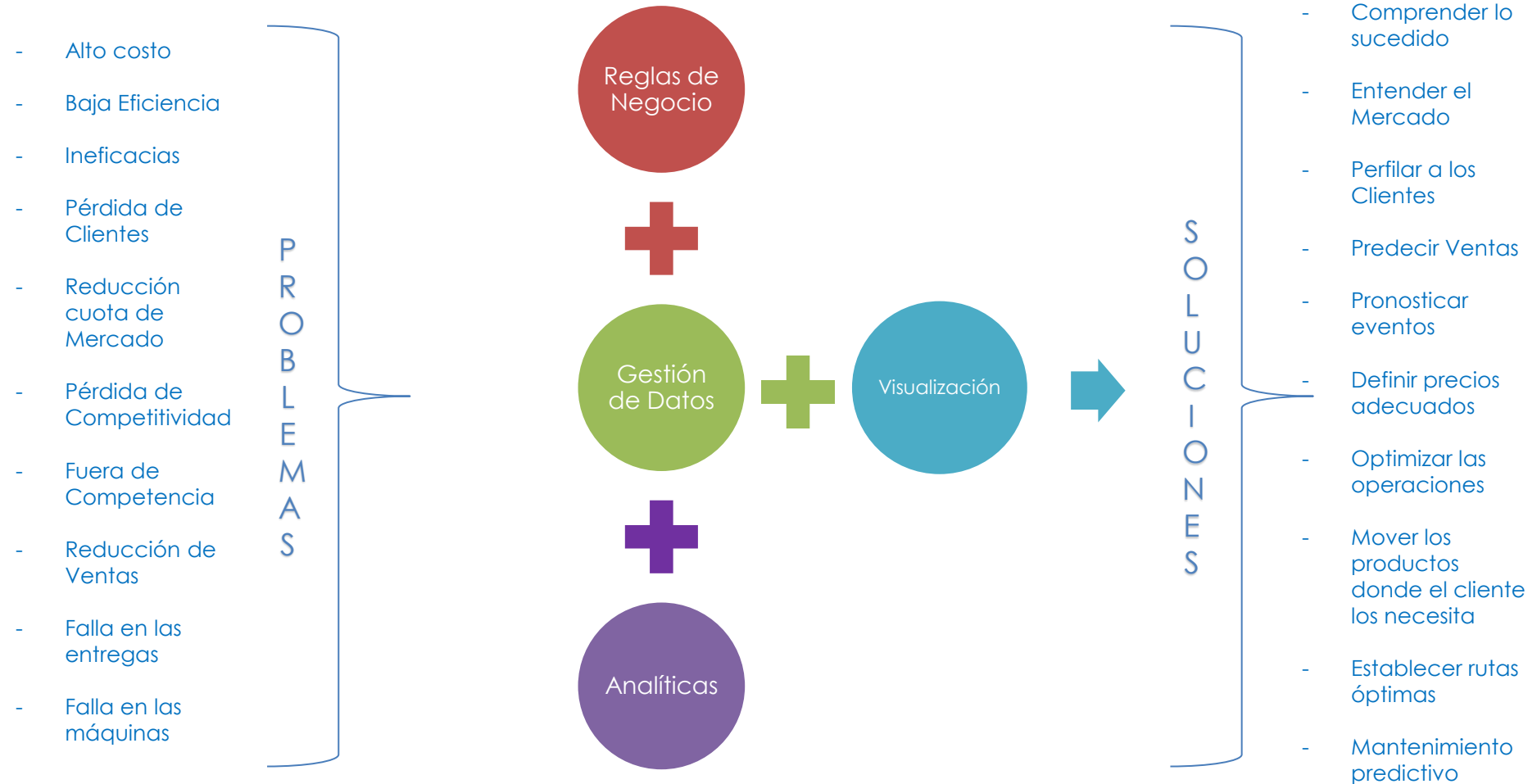
BI Value Stream

Entonces, en la práctica,
¿qué es lo que es necesario
considerar?

BI Value Stream



BI Value Stream

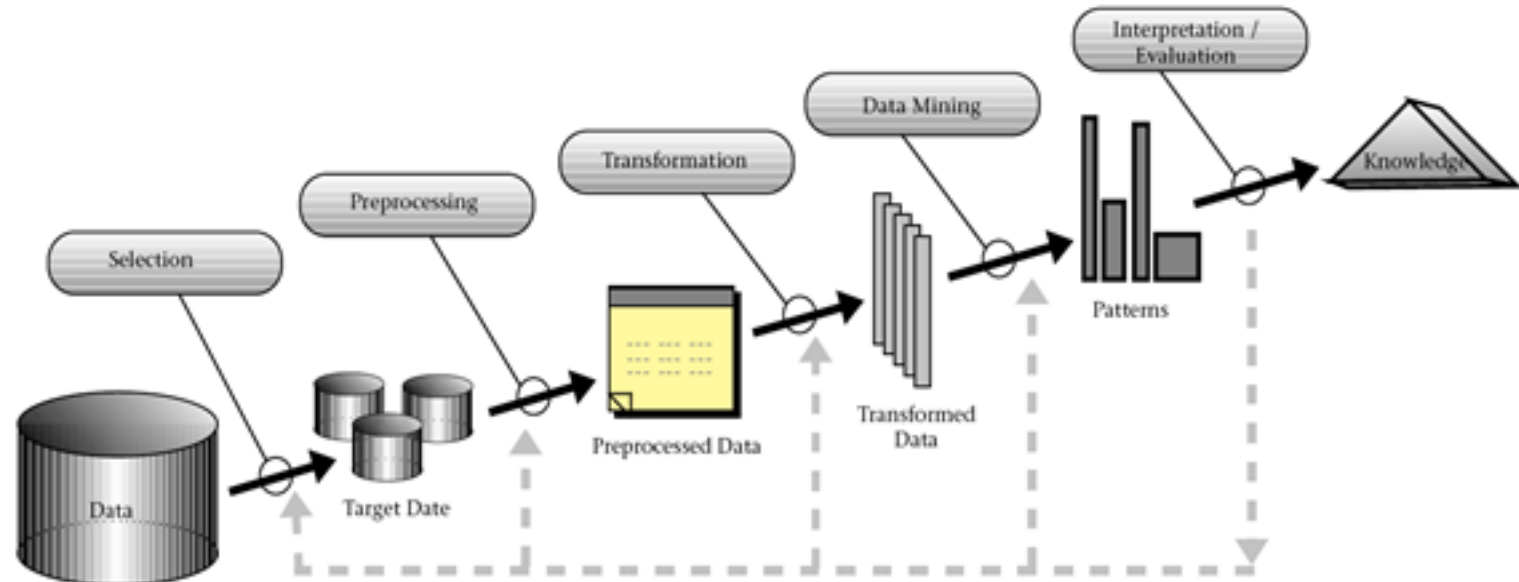


A person in a dark suit is holding a tablet. The tablet screen is blue and displays various business intelligence (BI) visualizations. At the top left, there is a magnifying glass icon with the word "SEARCH" inside. To its right, the words "BUSINESS INTELLIGENCE" are written in a bold, white, sans-serif font. Below this, the screen shows a complex dashboard with several elements: a line graph with data points, a bar chart, a pie chart, and a process flow diagram with numbered steps (02, 03, 04, 05, 06). There are also percentage values like "20%", "15%", and "5%" associated with different parts of the dashboard. The background of the image is a blurred office setting with windows and other people, giving it a professional, modern feel.

BI Como Processo

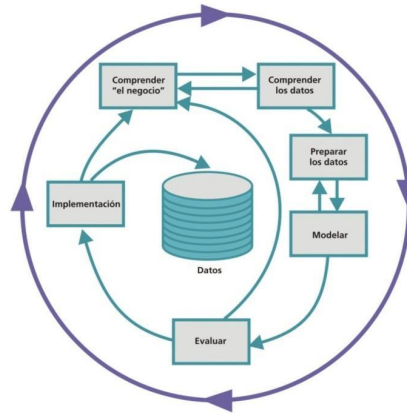
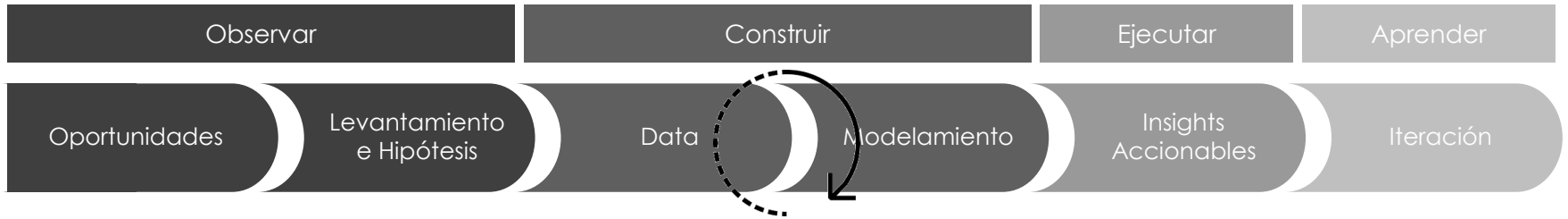
Inicialmente se consideró de
esta manera

Una mirada alternativa: KDD

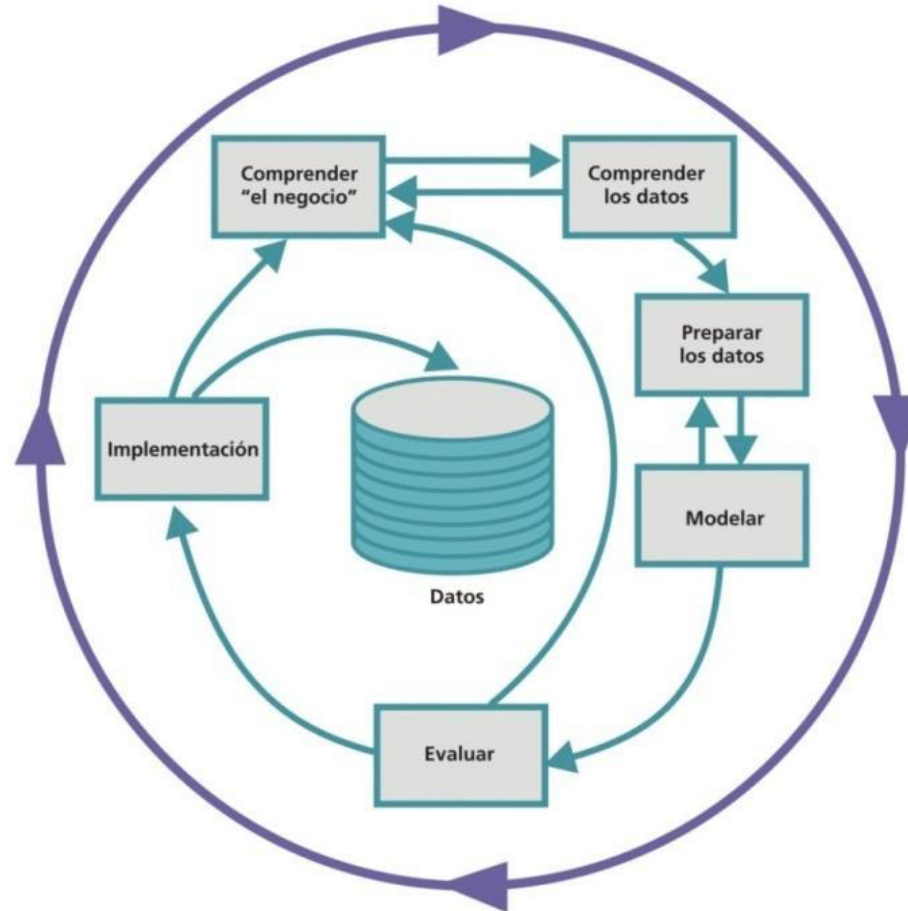


Pero con el tiempo, llega la
mirada de proceso para BI/BA

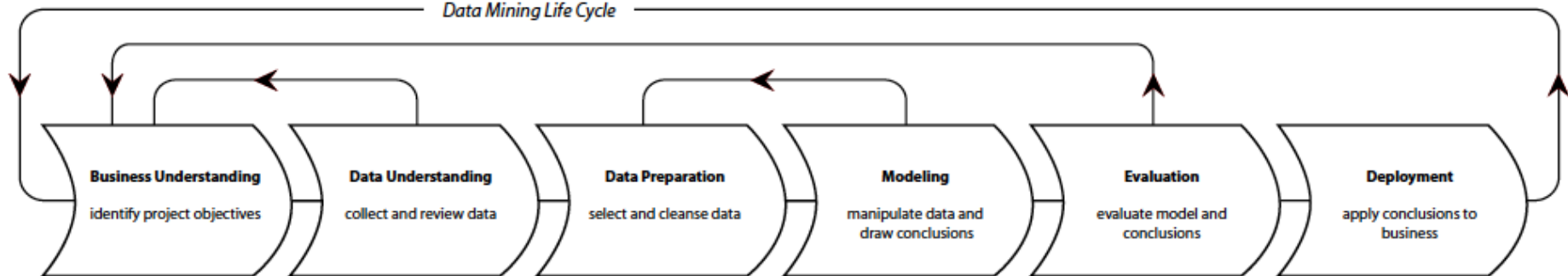
BI como Proceso



BI como Proceso



Hay enfoques
alternativos como
SEMMA de SAS

**Determine Business Objectives**

Background
Business Objectives
Business Success Criteria
(Log and Report Process)

Assess Situation

Inventory of Resources,
Requirements, Assumptions,
and Constraints
Risks and Contingencies
Terminology
Costs and Benefits
(Log and Report Process)

Determine Data Mining Goals

Data Mining Goals
Data Mining Success Criteria
(Log and Report Process)

Produce Project Plan

Project Plan
Initial Assessment of Tools and
Techniques
(Log and Report Process)

Collect Initial Data

Initial Data Collection Report
(Log and Report Process)

Describe Data

Data Description Report
(Log and Report Process)

Explore Data

Data Exploration Report
(Log and Report Process)

Verify Data Quality

Data Quality Report
(Log and Report Process)

Data Set

Data Set Description
(Log and Report Process)

Select Data

Rationale for Inclusion/
Exclusion
(Log and Report Process)

Clean Data

Data Cleaning Report
(Log and Report Process)

Construct Data

Derived Attributes
Generated Records
(Log and Report Process)

Integrate Data

Merged Data
(Log and Report Process)

Format Data

Reformatted Data
(Log and Report Process)

Select Modeling Technique

Modeling Technique
Modeling Assumptions
(Log and Report Process)

Generate Test Design

Test Design
(Log and Report Process)

Build Model Parameter Settings

Models
Model Description
(Log and Report Process)

Assess Model

Model Assessment
Revised Parameter
(Log and Report Process)

Evaluate Results

Align Assessment of Data
Mining Results with
Business Success Criteria
(Log and Report Process)

Approved Models

Review Process
Review of Process
(Log and Report Process)

Determine Next Steps

List of Possible Actions
Decision
(Log and Report Process)

Plan Deployment

Deployment Plan
(Log and Report Process)

Plan Monitoring and Maintenance

Monitoring and
Maintenance Plan
(Log and Report Process)

Produce Final Report

Final Report
Final Presentation
(Log and Report Process)

Review Project

Experience
Documentation
(Log and Report Process)

Generic Tasks

Specialized Tasks
(Process Instances)

a visual guide to CRISP-DM methodology

SOURCE CRISP-DM 1.0

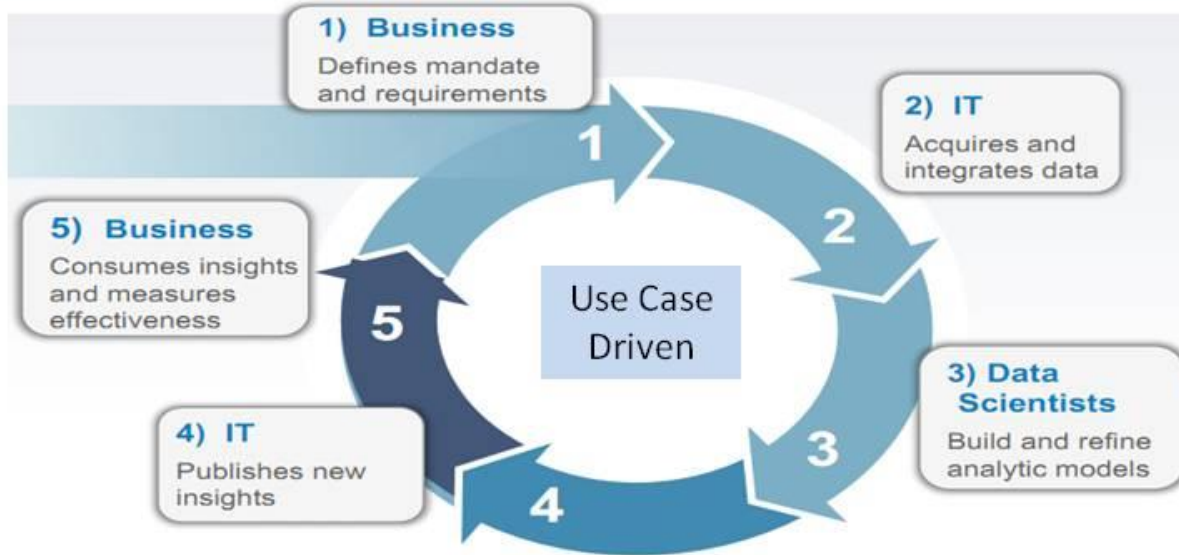
<http://www.crisp-dm.org/download.htm>

DESIGN Nicole Leaper

<http://www.nicoleleaper.com>



Use Case: focusing on a compelling business opportunity



Focus on an organization's **Strategic Business Initiatives** in order to be relevant to the business

- Public statement of business intent
- Delivers compelling business value
- Cross-functional
- Championed by a senior executive
- Has measurable goals
- Has well-defined delivery timeframe



Use "**visioning methodology**" to identify "where" and "how" Big Data Analytics can power an organization's Strategic Business Initiatives

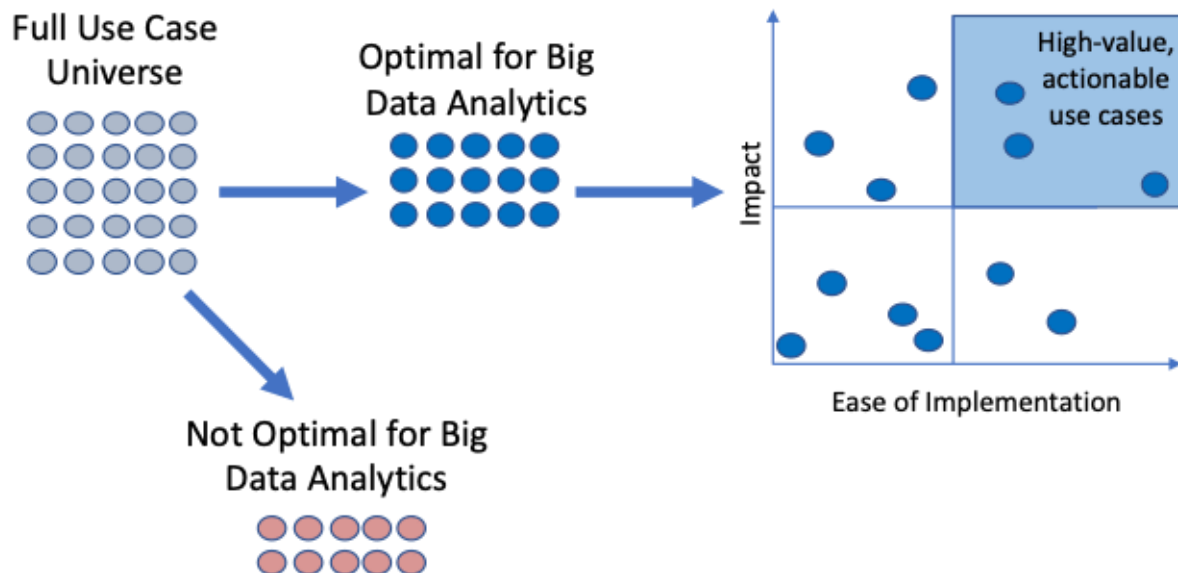
1. Identify the universe of use cases that would provide value to the organization
2. Estimate the value of each use case
3. Review the **impact** and **complexity** for each use case

- **Impact**

- Financial
- Management
- Sales
- Marketing
- Procurement
- IT

- **Complexity**

- Data source complexity
- Analytical complexity
- Business complexity
- Regulatory complexity
- Implementation complexity



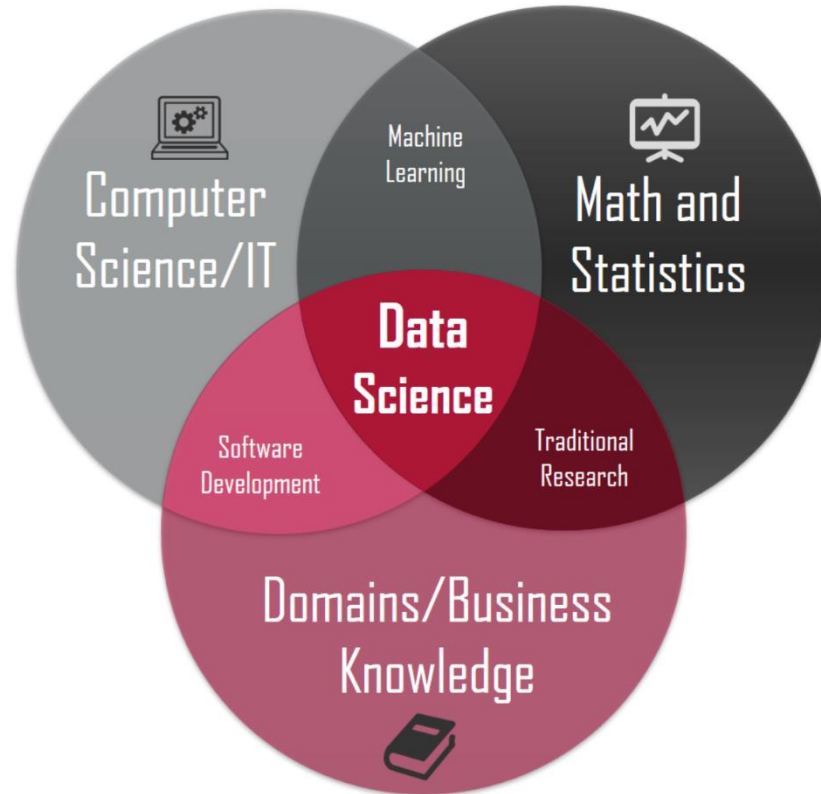


BUSINESS INTELLIGENCE

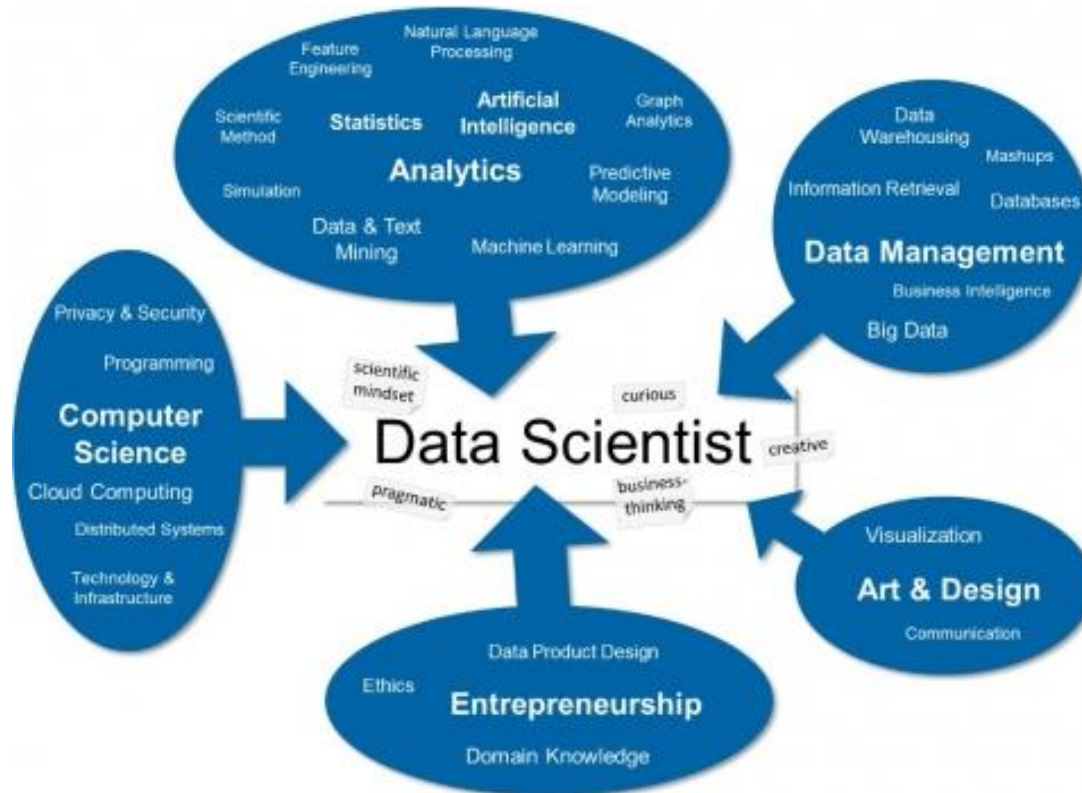
Capacidades BI



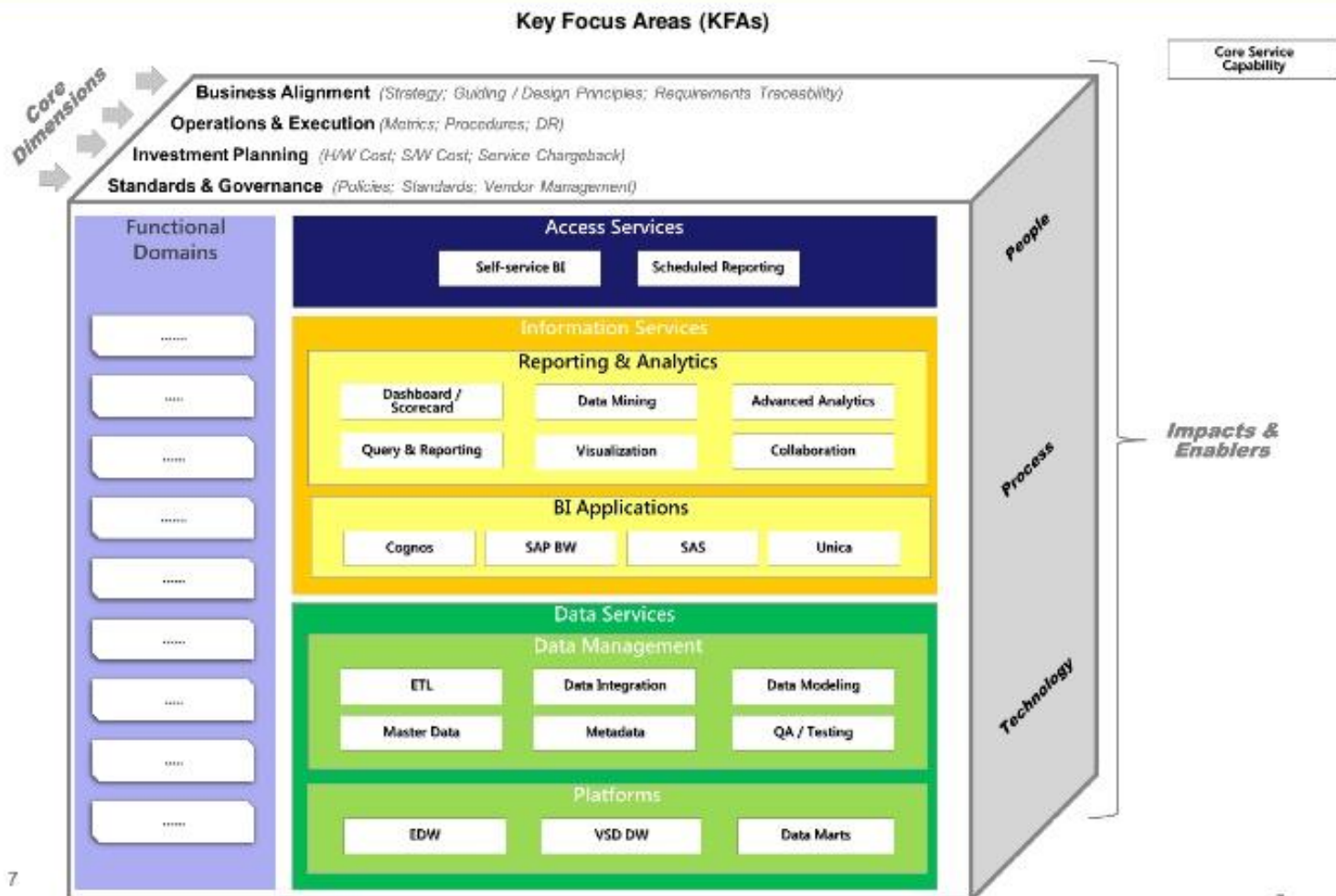
Capacidades BI

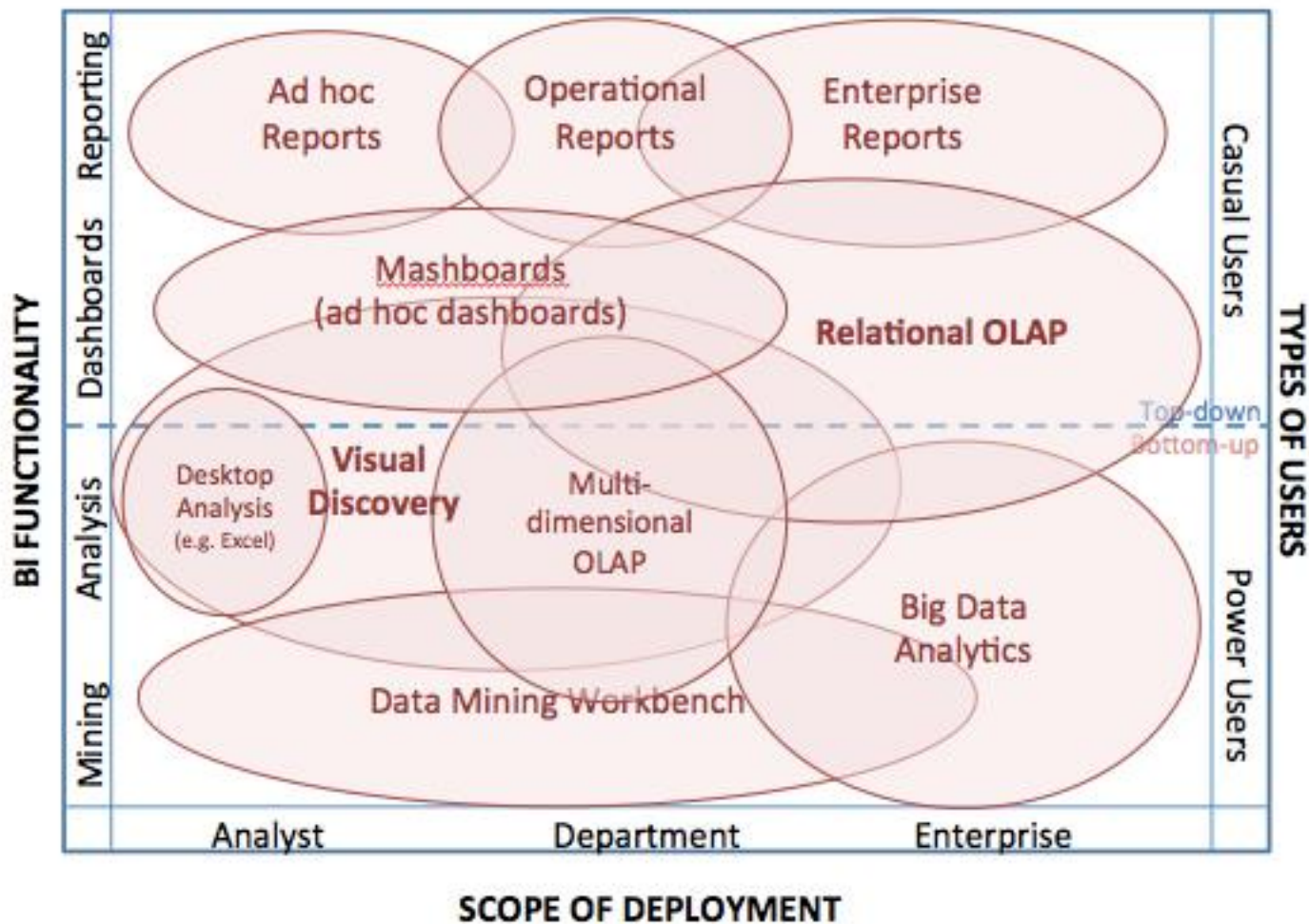


Capacidades BI



BI Service Capability Framework





Veamos esto en roles...

Fuente: <https://todobi.com/tipos-de-roles-en-analytics-business/>

Languages
SQL

Skills & Talents

- ✓ Basic tools (e.g. MS Office)
- ✓ Data visualization tools (e.g. Tableau)
- ✓ Conscious listening and storytelling
- ✓ Business Intelligence understanding
- ✓ Data modeling



HIRED BY
UBER  ORACLE

BUSINESS ANALYST

CHANGE AGENT

Role

Improves business processes as intermediary between business and IT

Mindset

Resilient project juggler

DATA ANALYST

DATA DETECTIVE

Role

Collects, processes and performs statistical data analyses

Mindset

Intuitive data junkie with high "figure-it-out" quotient



HIRED BY
  

Languages

R, Python, HTML, Javascript, C/C++, SQL

Skills & Talents

- ✓ Spreadsheet tools (e.g. Excel)
- ✓ Database systems (SQL and NO SQL based)
- ✓ Communication & visualization
- ✓ Math, Stats, Machine Learning

DATA ARCHITECT

THE CONTEMPORARY DATA MODELLER

Languages
SQL, XML, Hive, Pig, Spark

- Skills & Talents**
- ✓ Data warehousing solutions
 - ✓ In-depth knowledge of database architecture
 - ✓ Extraction Transformation and Load (ETL), spreadsheet and BI tools
 - ✓ Data modeling
 - ✓ Systems development



Role:

Creates blueprints for data management systems to integrate, centralize, protect and maintain data sources

Mindset:

Inquiring ninja with a love for data architecture design patterns

HIRED BY
VISA *Coca-Cola* logitech

DATA ENGINEER

SOFTWARE ENGINEERS BY TRADE

Role
Develops, constructs, tests and maintains architectures (such as databases and large-scale processing systems)

Mindset
All-purpose everyman



Languages

SQL, Hive, Pig, R, Matlab, SAS, SPSS, Python, Java, Ruby, C++, Perl

Skills & Talents

- ✓ Database systems (SQL & NO SQL based)
- ✓ Data modeling & ETL tools
- ✓ Data APIs
- ✓ Data warehousing solutions

HIRED BY
 Spotify  

DATA SCIENTIST

"AS RARE AS UNICORNS"

Languages
R, SAS, Python, Matlab, SQL,
Hive, Pig, Spark

- Skills & Talents**
- ✓ Distributed computing
 - ✓ Predictive modeling
 - ✓ Story-telling and visualizing
 - ✓ Math, Stats, Machine Learning



Role
*Cleans, massages and organizes
(big) data*

Mindset
Curious data wizard

HIRED BY



DATABASE ADMINISTRATOR

"DATABASE CARETAKER"

Role
*Ensures that the database is
available to all relevant users, is
performing properly and is being
kept safe*

Mindset
Master of Disaster Prevention



Languages
SQL, Java, Ruby on Rails, XML, C#,
Python

- Skills & Talents**
- ✓ Backup & recovery
 - ✓ Data modeling and design
 - ✓ Distributed Computing (Hadoop)
 - ✓ Database systems (SQL and NO SQL based)
 - ✓ Data security
 - ✓ ERP & business knowledge

HIRED BY



STATISTICIAN

"HISTORIC LEADERS OF DATA"

Languages

R, SAS, SPSS, Matlab, Stata, Python,
Perl, Hive, Pig, Spark, SQL

Skills & Talents

- ✓ Statistical theories & methodology
- ✓ Data mining & machine learning
- ✓ Distributed Computing (Hadoop)
- ✓ Database systems (SQL and NO SQL based)
- ✓ Cloud tools



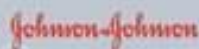
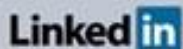
Role

*Collects, analyzes and interprets-
qualitative as well as quantitative
data with statistical theories and
methods*

Mindset

*Logical and enthusiastic stats
genius*

HIRED BY



PEPSICO

DATA AND ANALYTICS MANAGER

"DATA SCIENCE TEAM LEADER"

Role

*Manages a team of analysts and
data scientists*

Mindset

Data Wizards' Cheerleader



Languages

SQL, R, SAS, Python, Matlab,
Java

Skills & Talents

- ✓ Database systems (SQL and NO SQL based)
- ✓ Leadership & project management
- ✓ Interpersonal communication
- ✓ Data mining & predictive modeling

HIRED BY

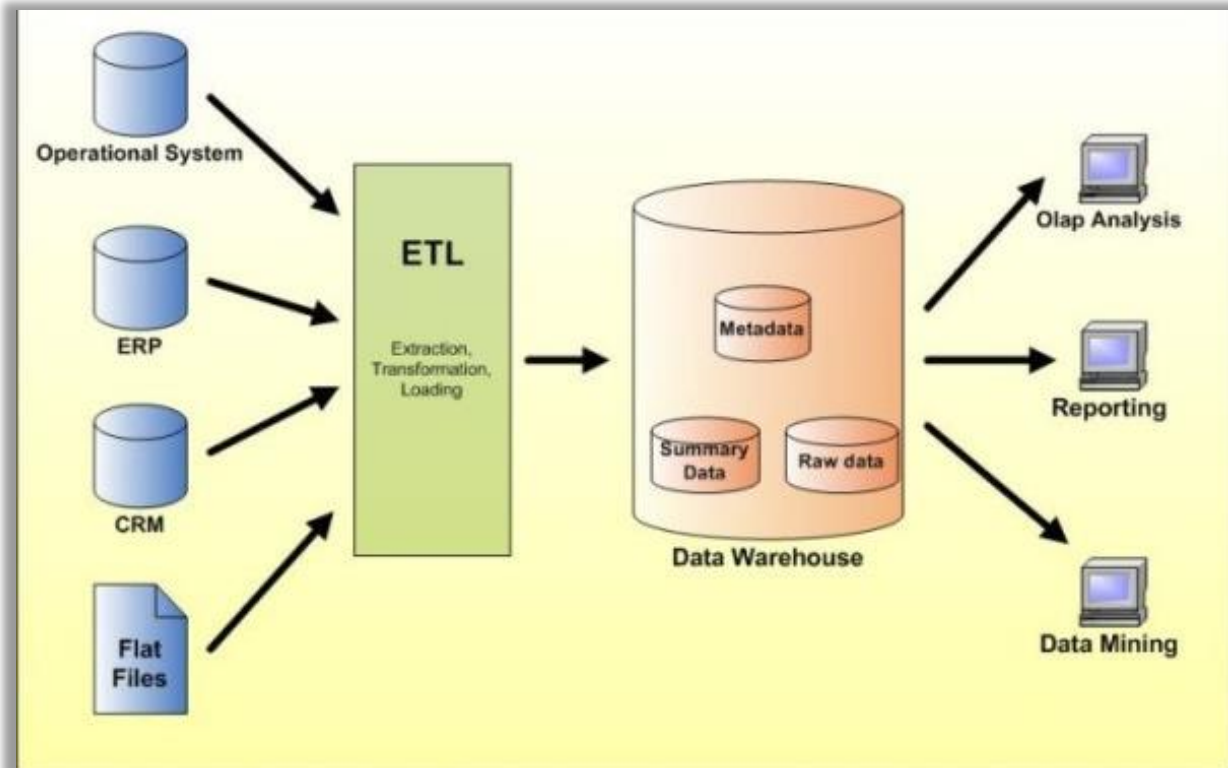


MOTOROLA SOLUTIONS

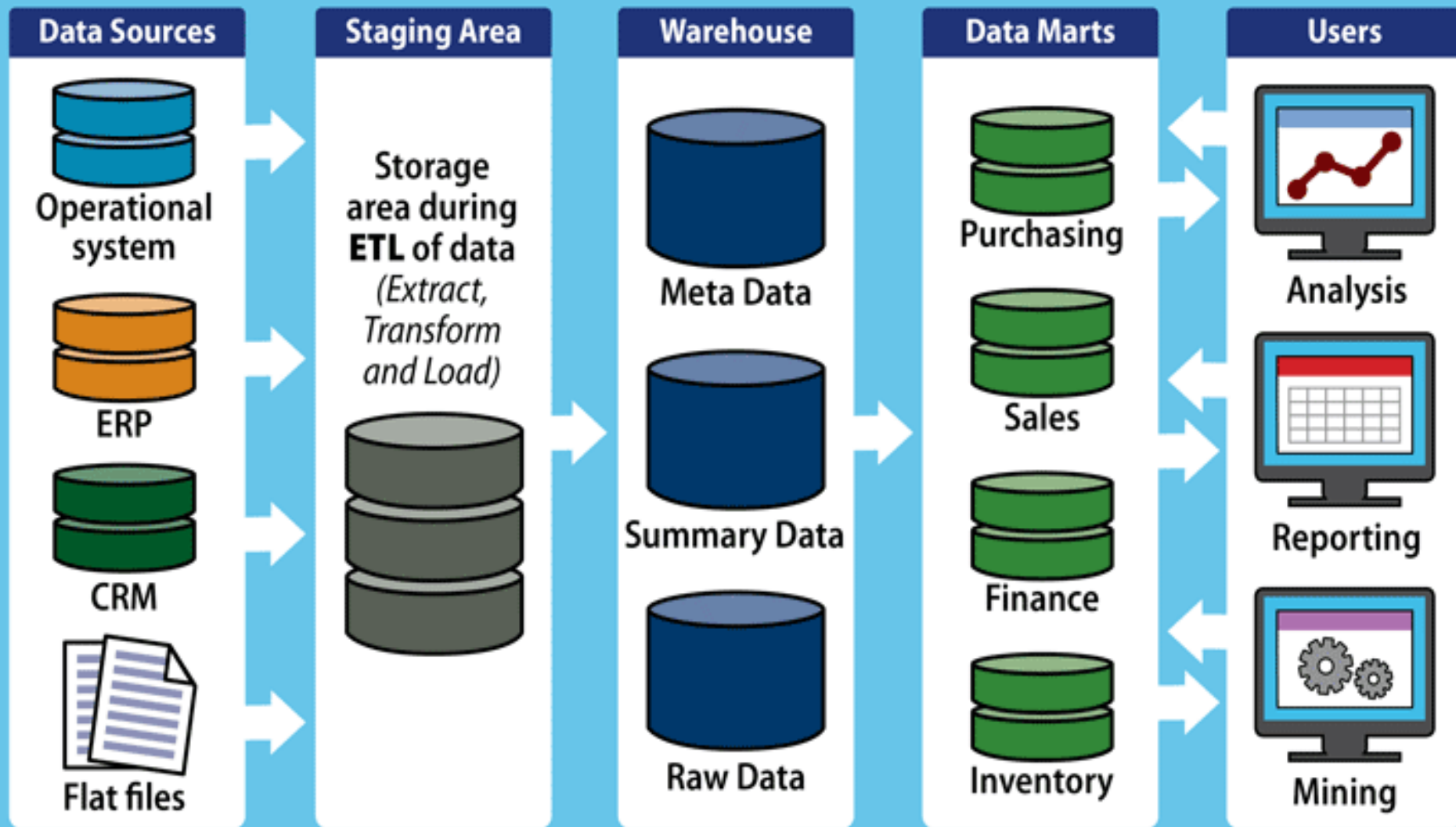


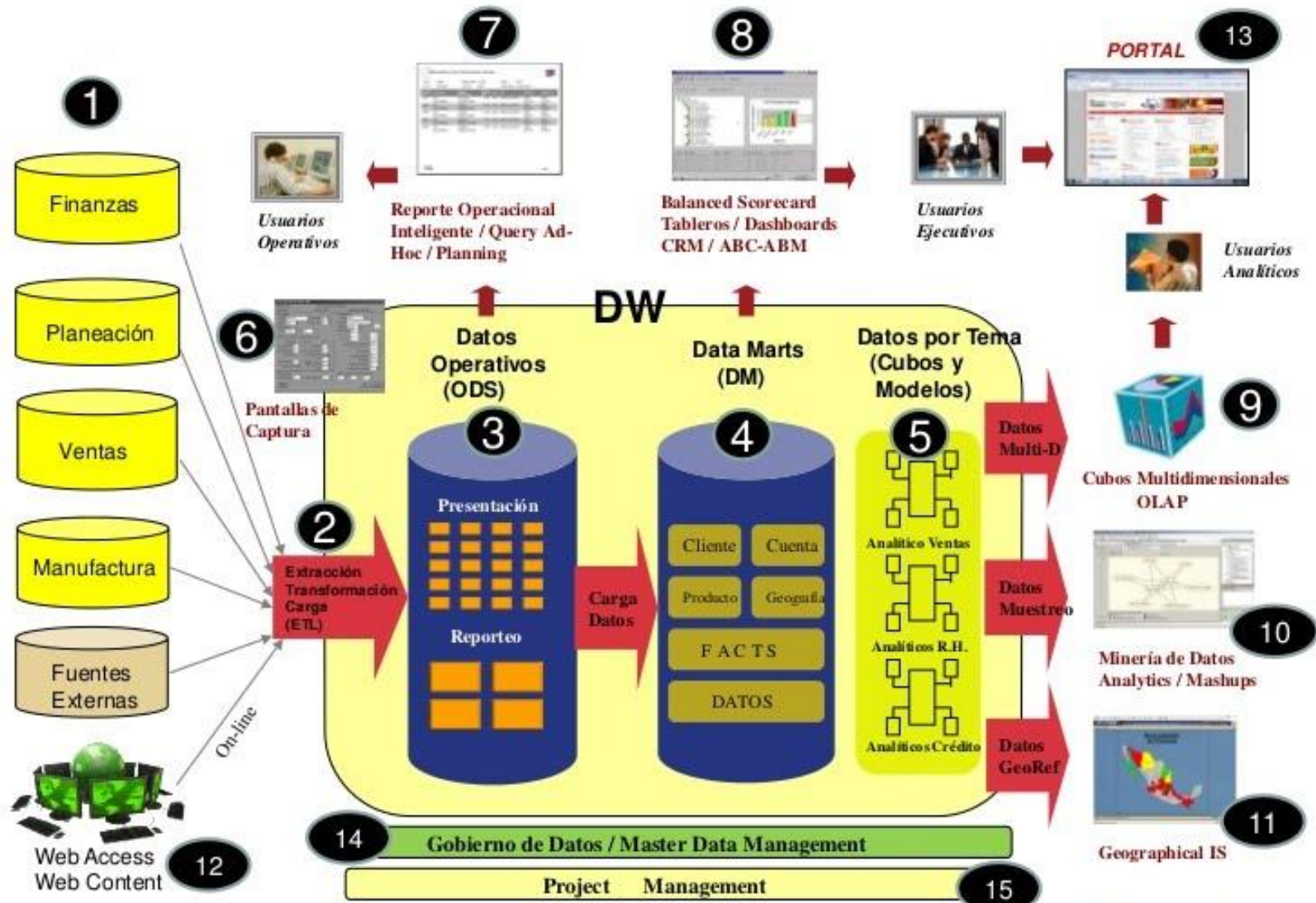
Elementos

Generalmente se comienza
con un Warehouse, para
responder al BI



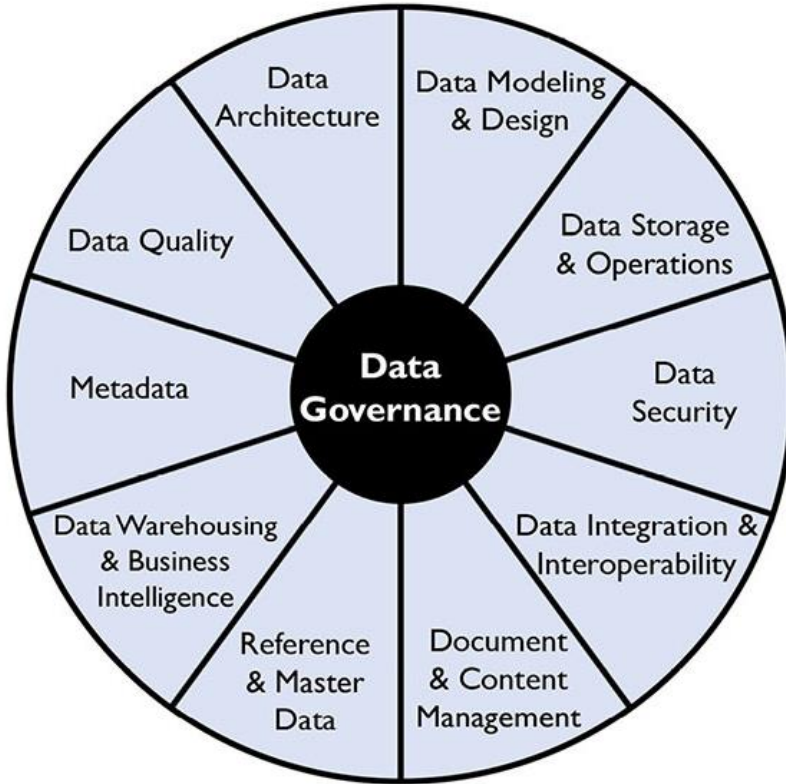
Data Warehouse Architecture





Pero, ojo, siempre hay un
problema central...

Data Governance



DAMA-DMBOK2 Data Management Framework

Copyright © 2017 by DAMA International



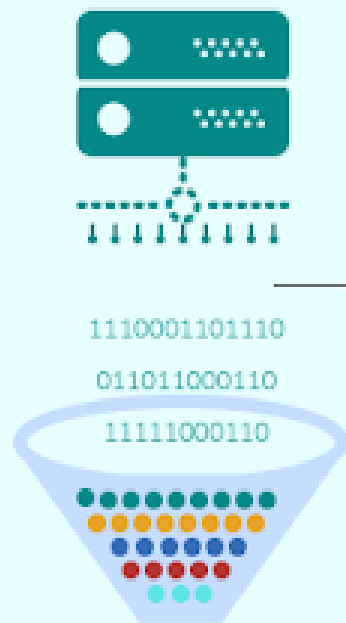
DAMA Environmental Factors Hexagon

Y, claro, luego del Data Warehouse se puede seguir con un Data Lake, para ir avanzando hacia el BA

DATA WAREHOUSE

VS

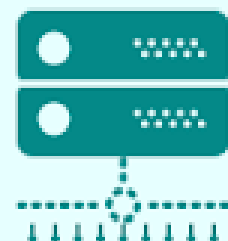
DATA LAKE



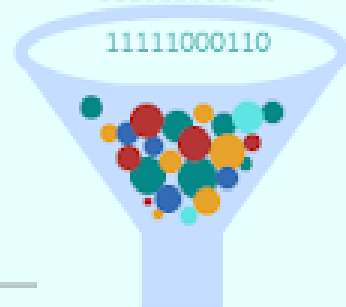
• Data is processed and organized into a single schema before being put into the warehouse



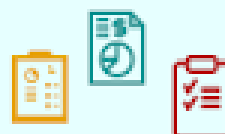
• The analysis is done on the cleansed data in the warehouse



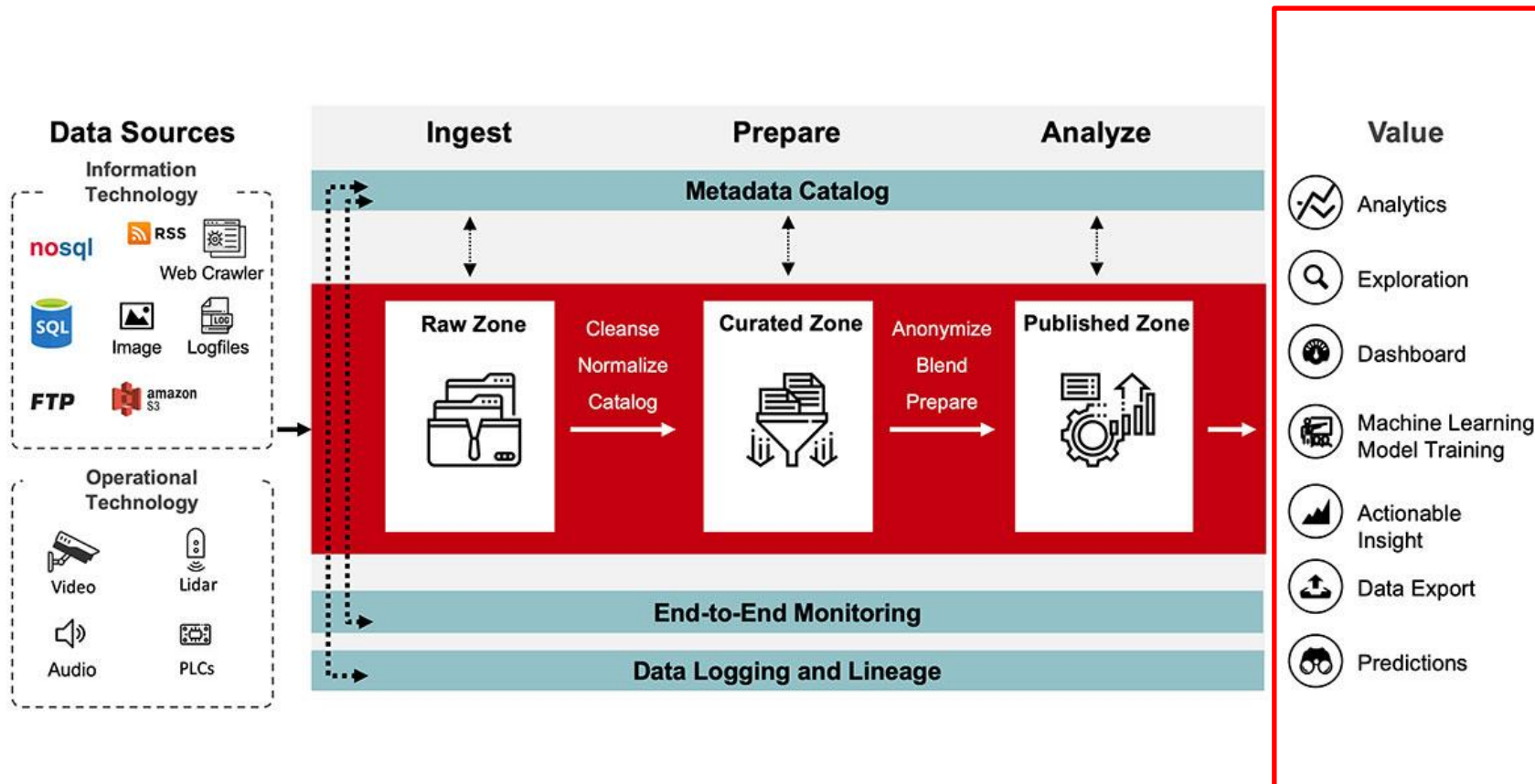
• Raw and unstructured data goes into a data lake



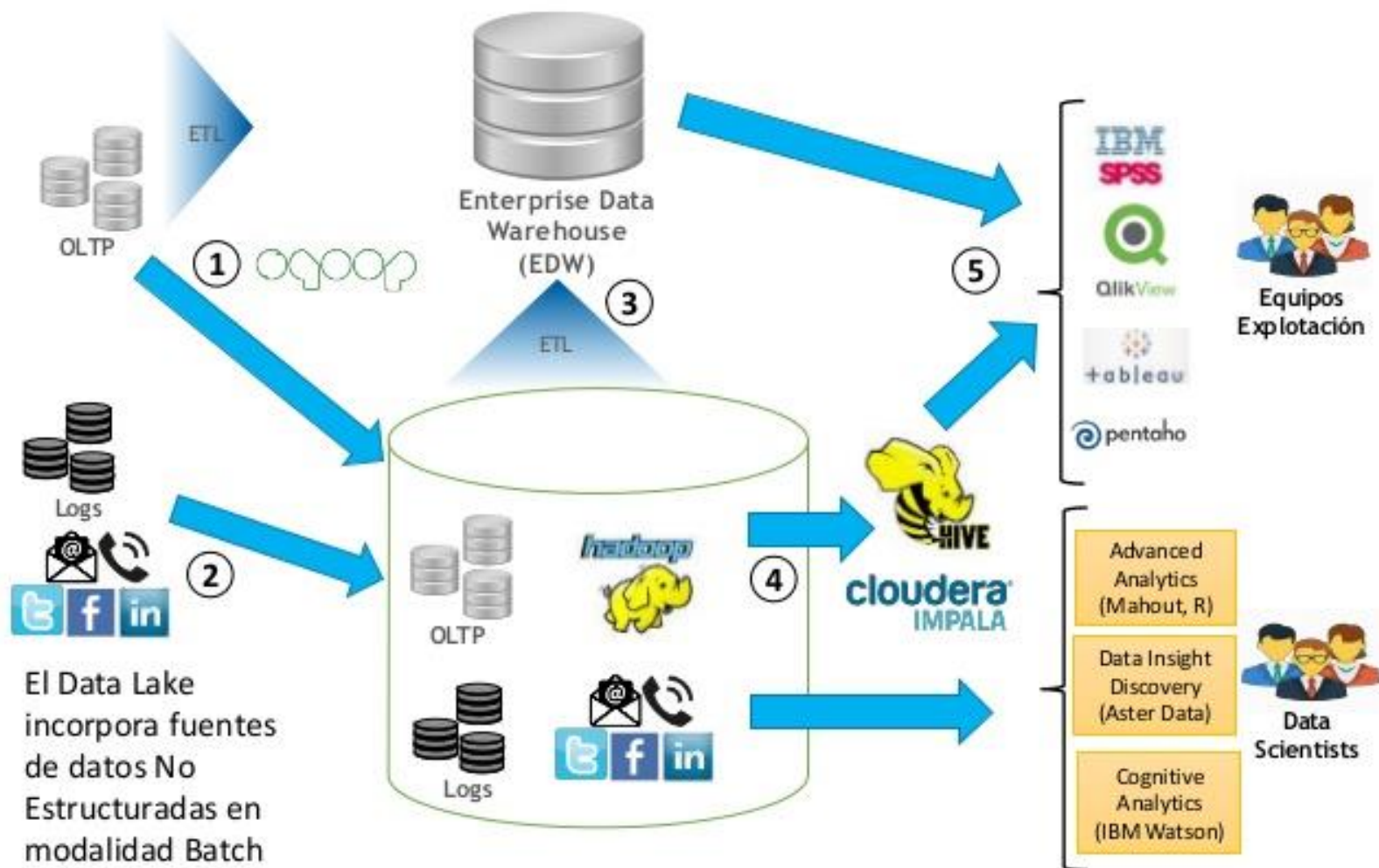
• Data is selected and organized as and when needed



Para agregar valor



Enterprise Data Lake + Nuevas Fuentes

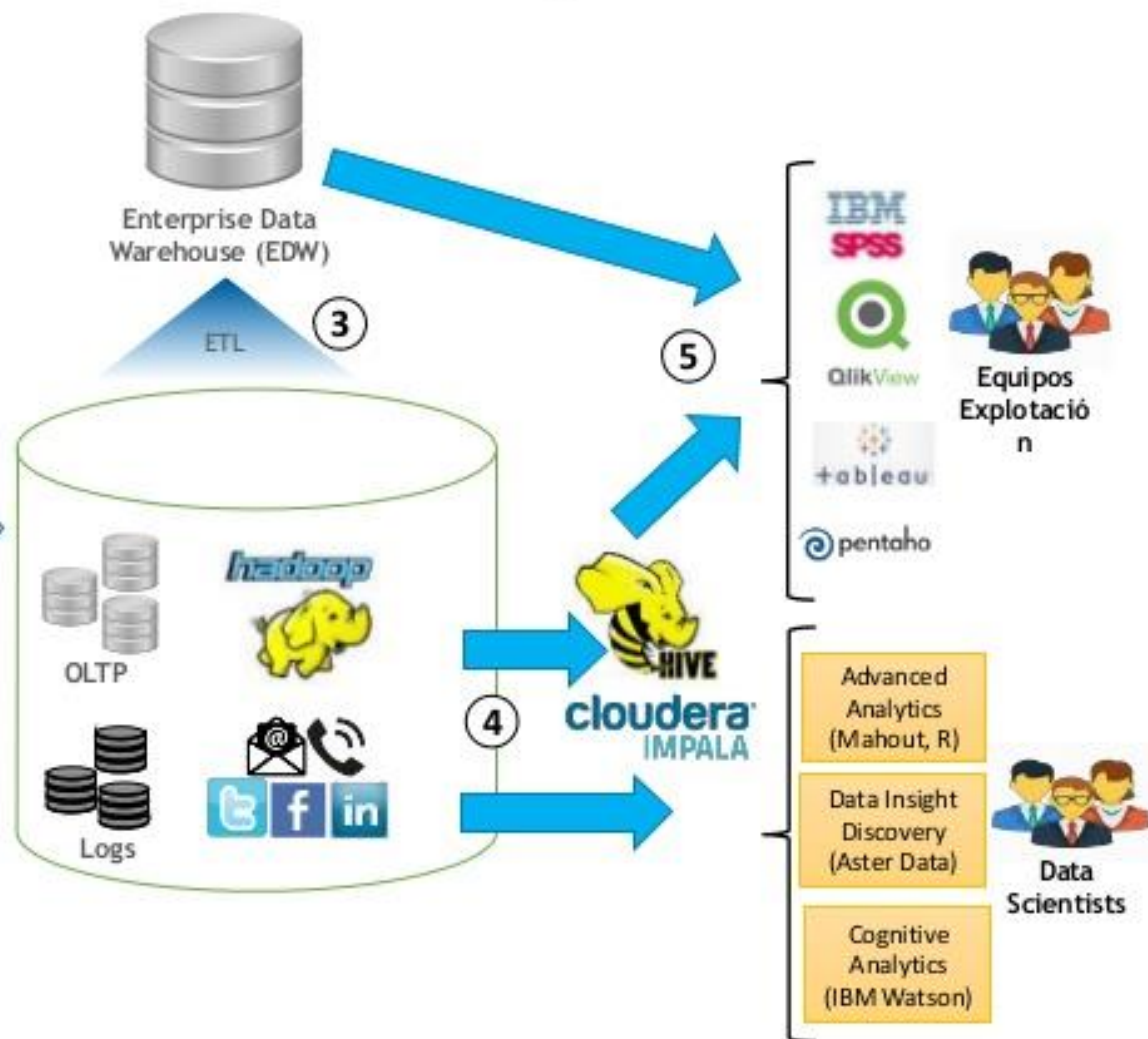


Near-Real-Time Data Ingestion

Los agentes de Flume hacen streaming de la información generada



Los Interceptors de Flume permiten enriquecer la ingesta de datos para finalmente guardarla en HDFS (Hadoop)



Escenario futuro (Big Data)



ETL



Enterprise Data Warehouse

ETL



Big Data Enterprise Platform (Data Lake)

EL

Carga incremental (en near-real-time) de todas las fuentes de información relevantes.

Carga de fuentes no estructuradas como logs, correos e información de redes sociales.

Capacidad analítica potenciada por nuevas fuentes de información

Soporta la implementación de servicios real-time y la práctica de Discovery de datos

Modelos Predictivos

Campañas Push

Modelos Riesgo

Reportes

Real-time Analytics API

Machine Learning based models

Campañas Real-time

Data Insight Discovery



Inteligencia de Client



Riesgo



Canales



Operacione



Web/Móvil



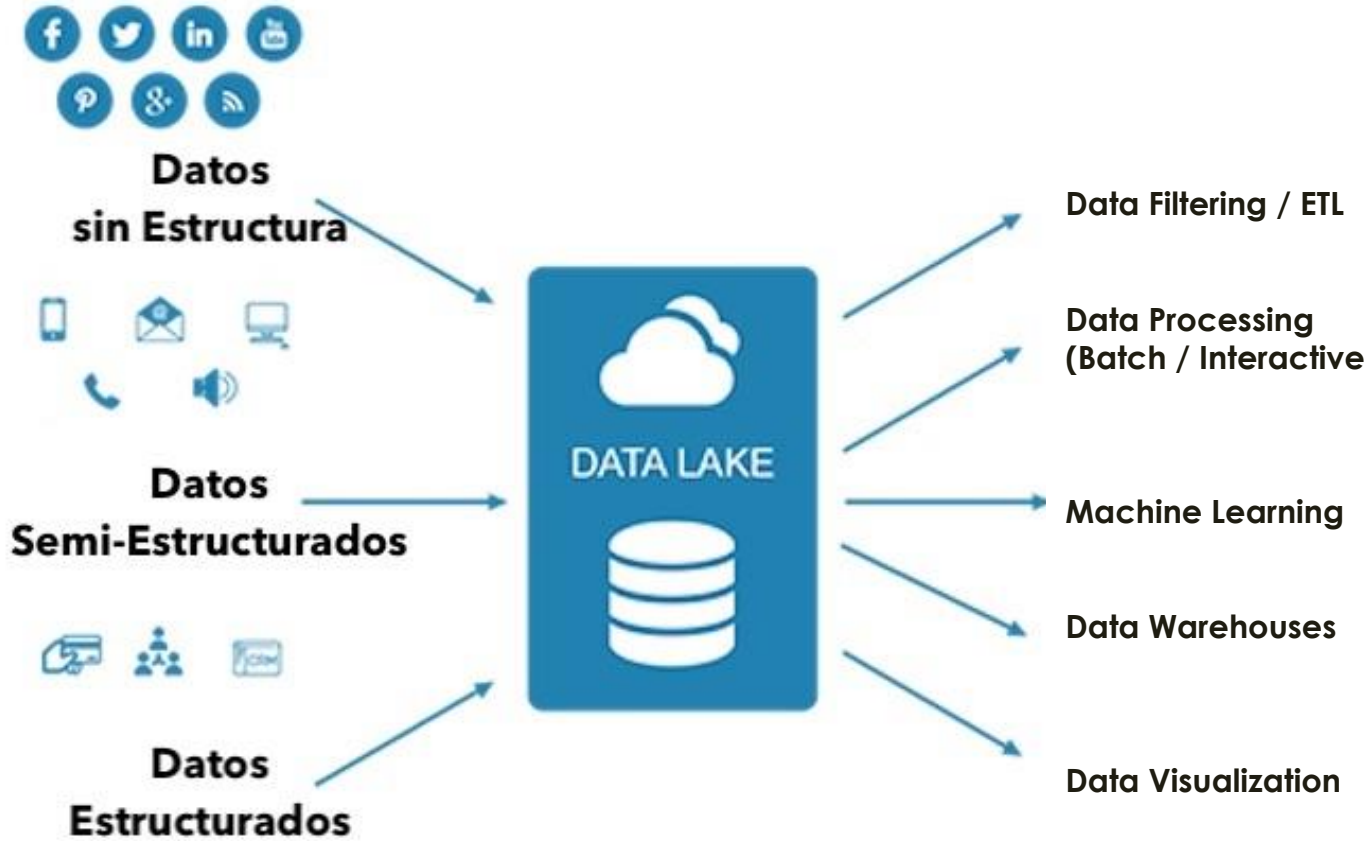
Sucursales



Callcenter

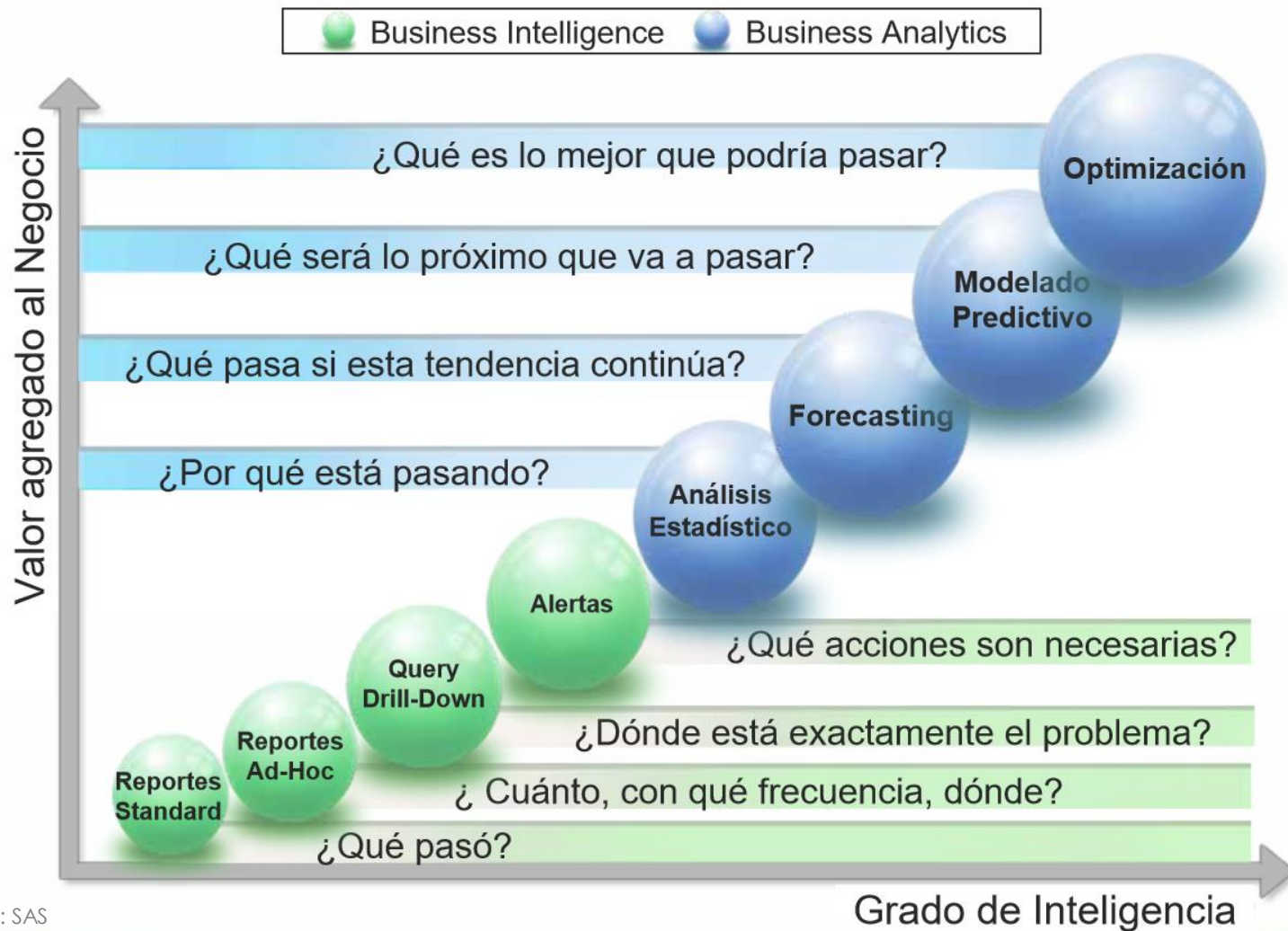


Data Scientists



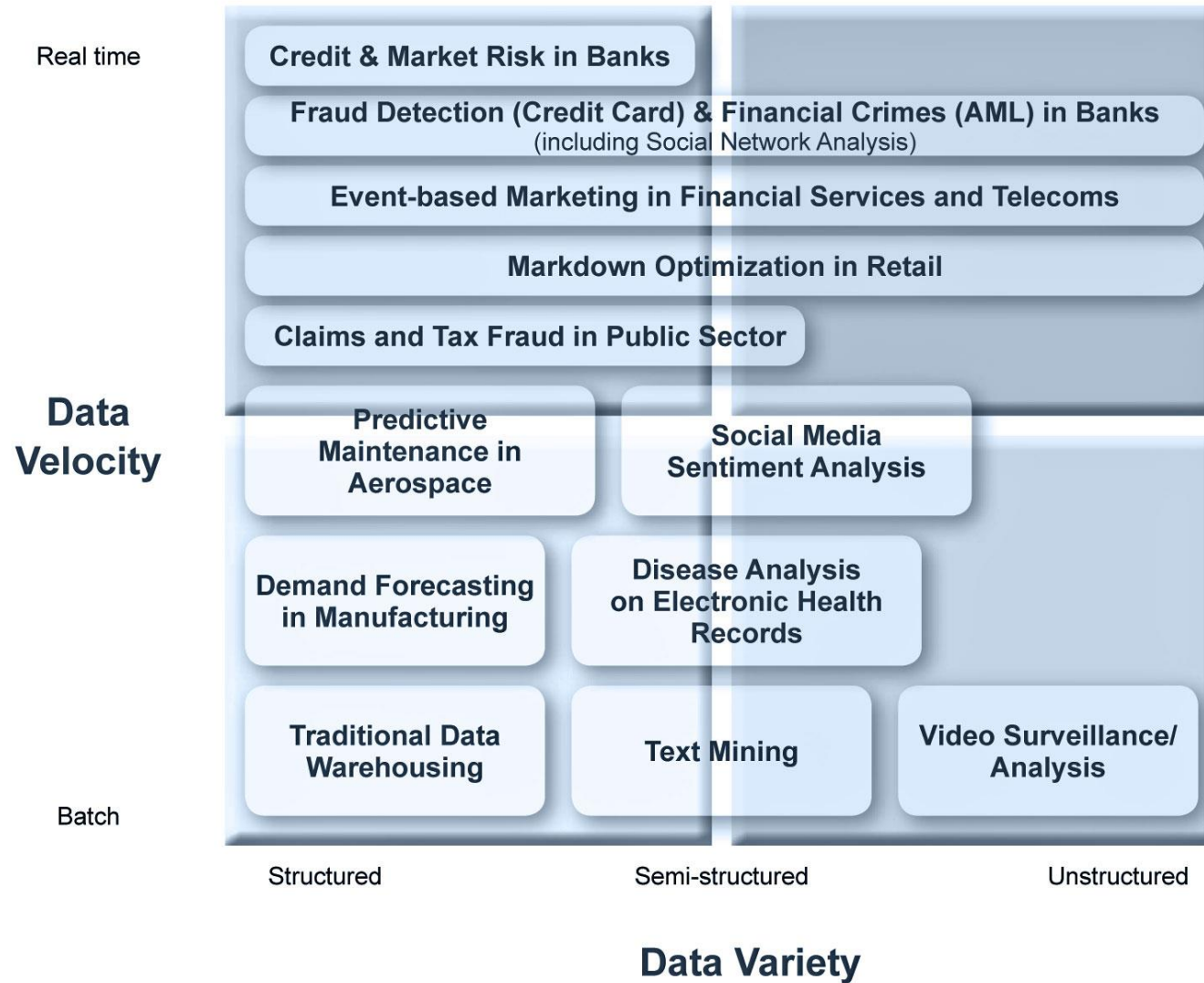
The background features a glowing blue digital environment. A laptop is positioned in the lower center, its screen displaying a complex dashboard with multiple data visualizations. Above the laptop, several floating panels show different types of charts: a bar chart at the top left, a line graph with a red peak in the center, and several circular progress indicators or donut charts on the right side. These charts contain various percentages like 25%, 50%, 75%, and 100%. The overall aesthetic is high-tech and data-driven, with a color palette dominated by blues, purples, and oranges.

Casos de Uso



La pregunta es, ¿y en qué
áreas se utiliza?

Potential Use Cases for Big Data Analytics



Mirando transversalmente



Operations

- Predictive maintenance
- Demand forecasting
- Operational efficiency
- Inventory optimization
- Operations anomaly insights
- Quality assurance
- Connected devices and smart buildings
- Supplier and spend insights



Marketing

- Personalization
- Customer insights
- Churn analytics
- Dynamic pricing
- Product innovation
- Marketing optimization
- Product recommendation



Finance

- Finance forecasting
- Fraud management
- Risk management



Workforce

- Employee insights
- HR insights
- Resource matching and planning



Service

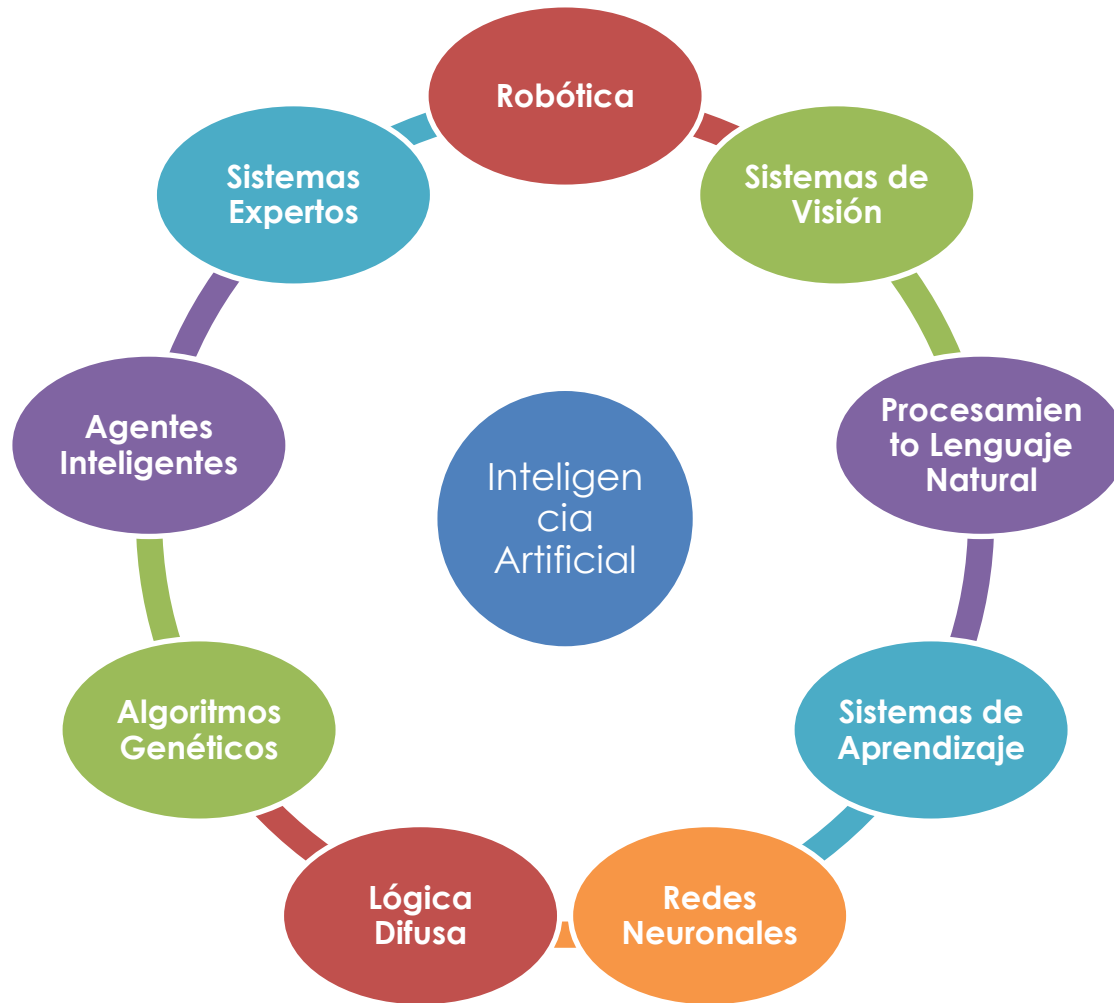
- Intelligent contact center
- Patient care and healthcare analytics



Sales

- Lead and opportunity scoring
- Sales insights

Lo interesante es que son las
herramientas de la IA nos
ayudan a BI/BA, pero
principalmente en BA



Artificial Intelligence

mimic human
behavior

sense,
reason,
adapt!

highly
autonomous!

Machine Learning

statistical
learning
techniques

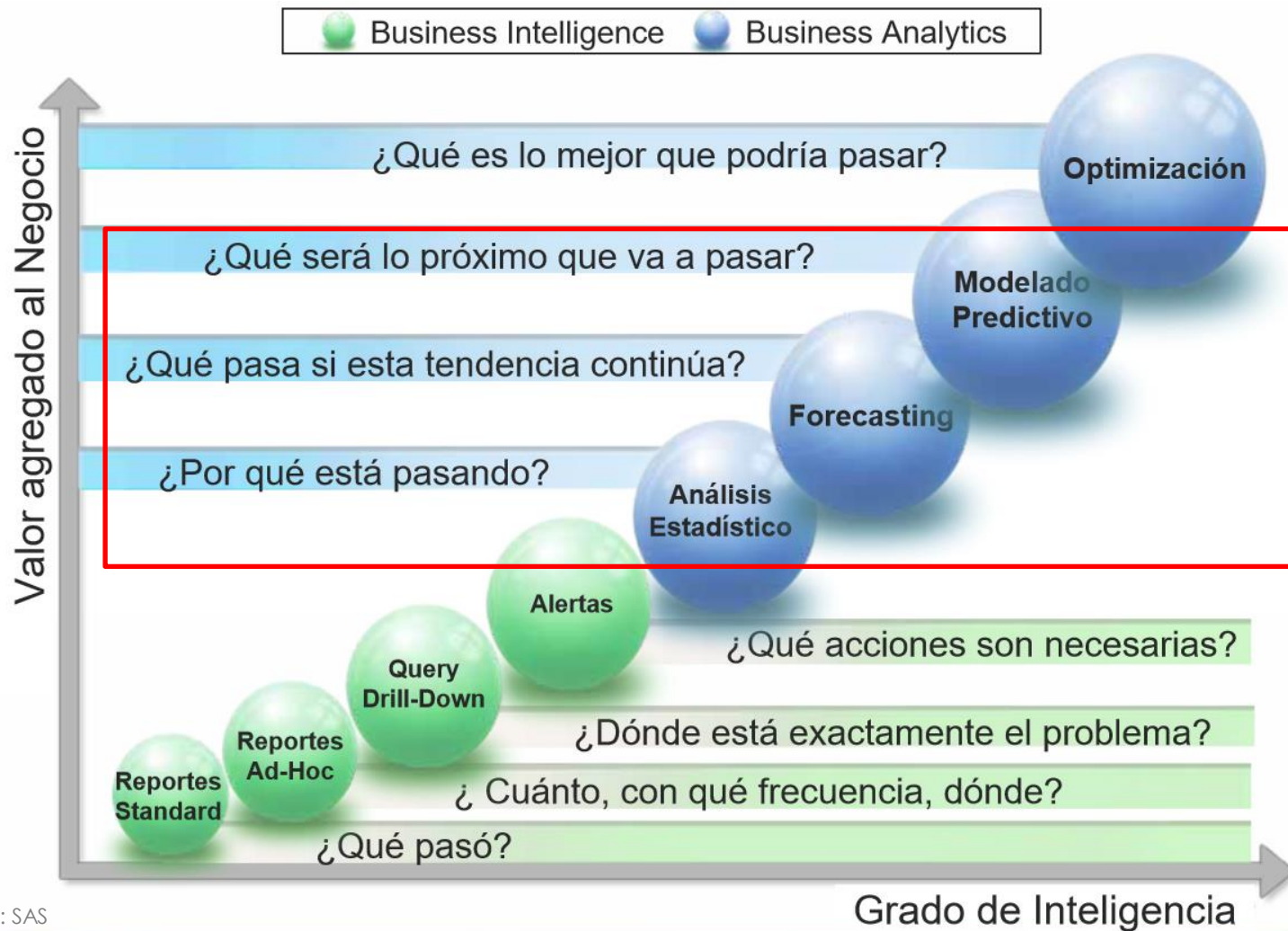
improve
learning with
example data!

Deep Learning

Subset of ML

Specific type of
learning algorithms

¿Dónde?



Predictive Analytics Use Cases



Analytics USE Case

Table 1.1 Example Analytics Applications

Marketing	Risk Management	Government	Web	Logistics	Other
Response modeling	Credit risk modeling	Tax avoidance	Web analytics	Demand forecasting	Text analytics
Net lift modeling	Market risk modeling	Social security fraud	Social media analytics	Supply chain analytics	Business process analytics
Retention modeling	Operational risk modeling	Money laundering	Multivariate testing		
Market basket analysis	Fraud detection	Terrorism detection			
Recommender systems					
Customer segmentation					

Los tipos clásicos de
aprendizaje que se utilizan son



Classical Machine Learning

Task Driven

Data Driven

Supervised Learning

(Pre Categorized Data)

Unsupervised Learning

(Unlabelled Data)

Classification

(Divide the socks by Color)

Eg. Identity
Fraud Detection

Regression

(Divide the Ties by Length)

Eg. Market
Forecasting

Clustering

(Divide by Similarity)

Eg. Targeted
Marketing

Association

(Identify Sequences)

Eg. Customer
Recommendation

Dimensionality Reduction

(Wider Dependencies)

Eg. Big Data
Visualization

Obj: Predications & Predictive Models

Pattern/ Structure Recognition



Qualitative

Like Easy
Awkward Slow
Squirrel
Efficient
Ambiguous How
Confusing

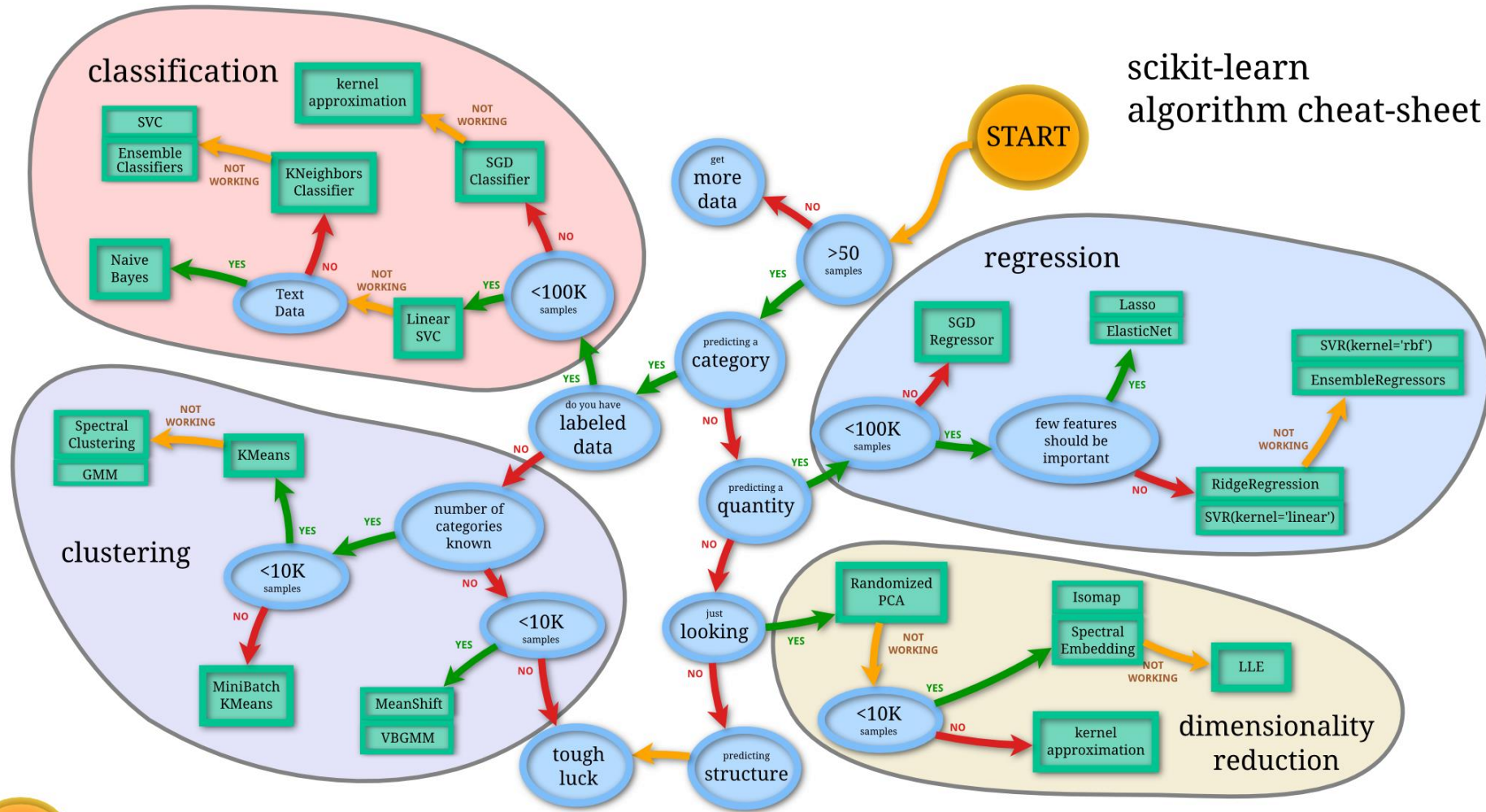
Quantitative

23,406 4.3
2m32s
76.8%
\$45,849
1,127 3.76%
€12.75

	Categorical values	Continuous values
Supervised learning	Two-class/Binary classification • Decision Forest • Decision Tree • Naive Bayes/Bayes Classifier • (Deep) Neural Networks • Support Vector Machines Multi-class classification • Decision Forest • Logistic Regression • (Deep) Neural Networks	Regression • Bayesian Linear Regression • Decision Forest Regression • Decision Tree Regression • Linear Regression • Neural Network Regression • Ordinary Least Squares Regression
Unsupervised learning	Association Rule Learning • Apriori algorithm • Frequent Pattern Growth Classification • Autoencoders • Restricted Boltzmann Machines	Clustering • Density-based Clustering • Hierarchical Clustering • Partitional Clustering (incl. K-means) Dimensionality reduction • Principal Component Analysis (PCA)

Una guía práctica para los
usuarios más avanzados

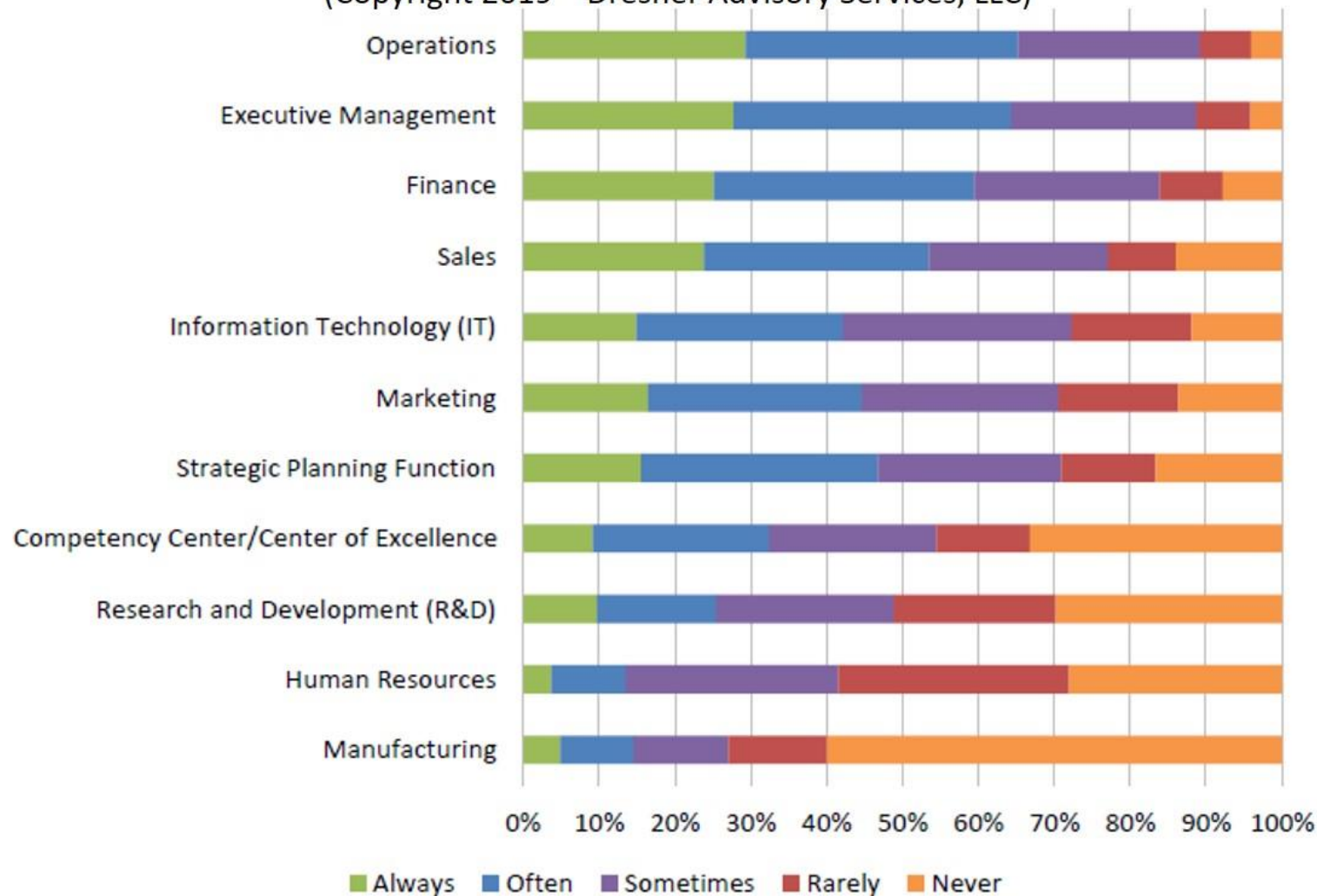
scikit-learn algorithm cheat-sheet



Principalmente en los ámbitos
de...

Functions Driving Business Intelligence

(Copyright 2019 – Dresner Advisory Services, LLC)



Y, principalmente realizando
los siguientes tipos de BI/BA

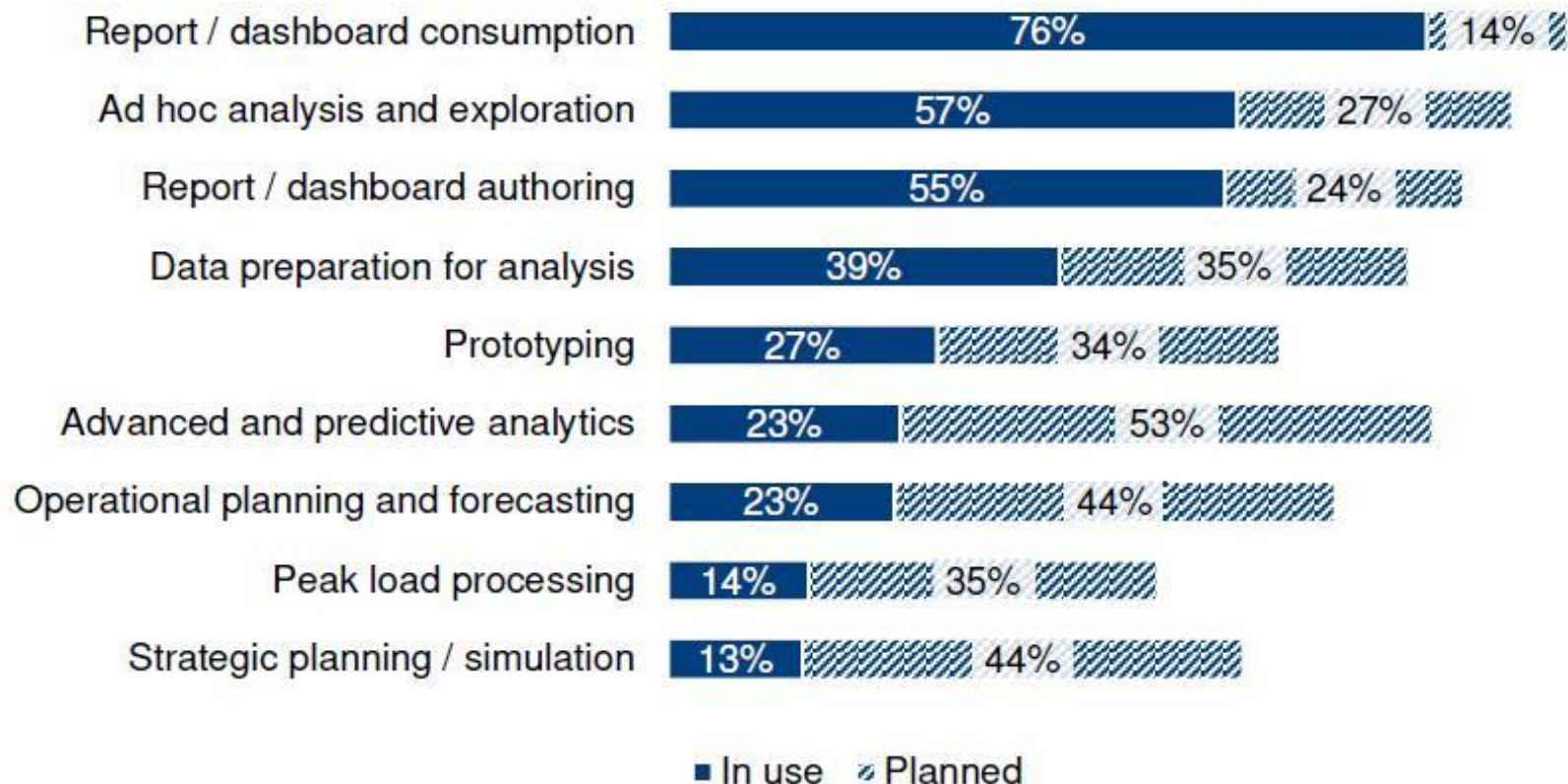
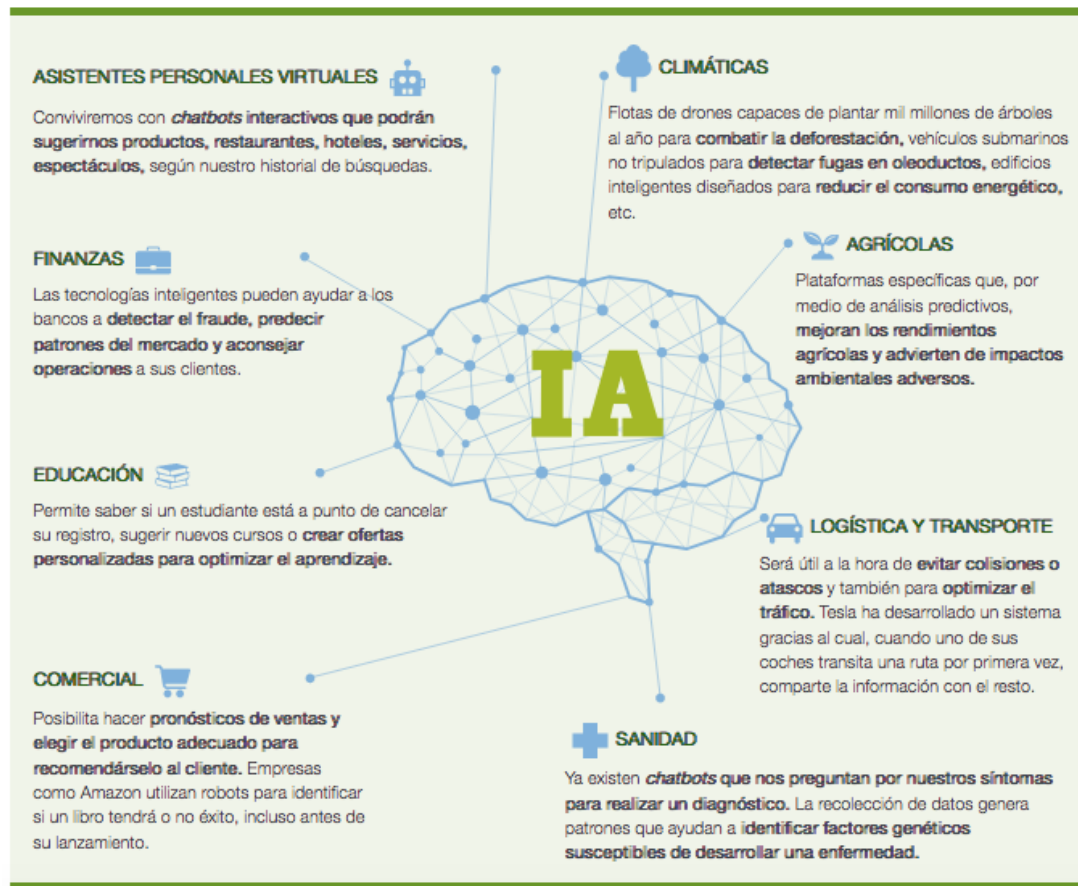


Figure 7: What are the primary use cases for cloud BI in your organization? (n=164)

Casos de Uso Clásicos más de BA



Analytics Use Cases

<https://www.oreilly.com/library/view/analytics-across-the/9780133835922/app01.html>



Las Grandes Preguntas



Las grandes preguntas

#QUIÉN_ME_VA_A_COMPRAR

#CÓMO LLEGO A TIEMPO

#QUÉ PRECIO ASIGNO

#QUÉ RUTA ES MÁS PELIGROSA

#CUÁNTO INVIERTO EN MKT

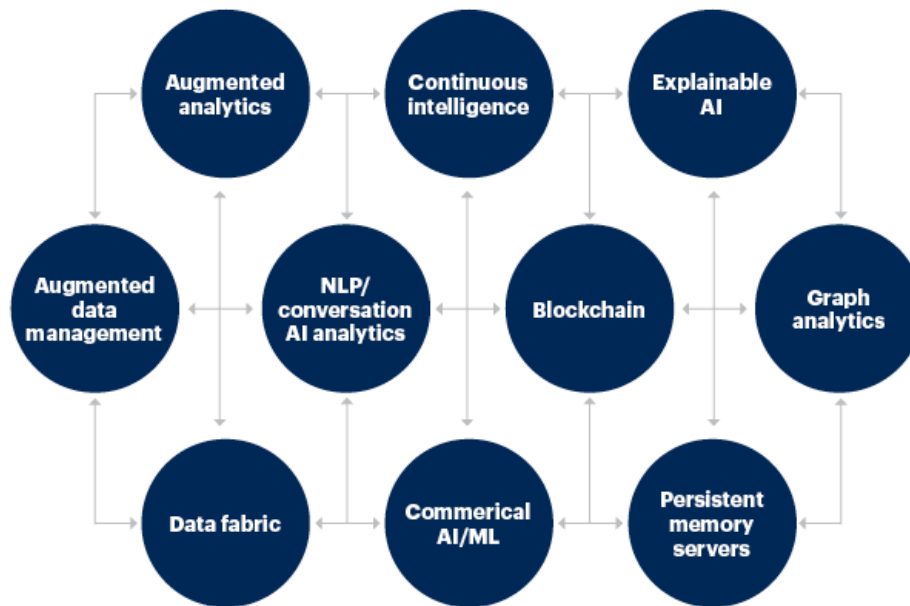
#QUÉ COMPRO

#QUÉ PRODUZCO



Herramientas Utilizadas en BI

Top 10 technology trends in data and analytics



Intelligent



Emergent



Scalable



gartner.com/SmarterWithGartner

Source: Gartner

© 2019 Gartner, Inc. All rights reserved. PR_781030

Gartner

Magic Quadrant for Analytics and Business Intelligence Platforms.



Source: Gartner (Feb 2019 and 2020)

Magic Quadrant for Data Science and Machine Learning Platforms

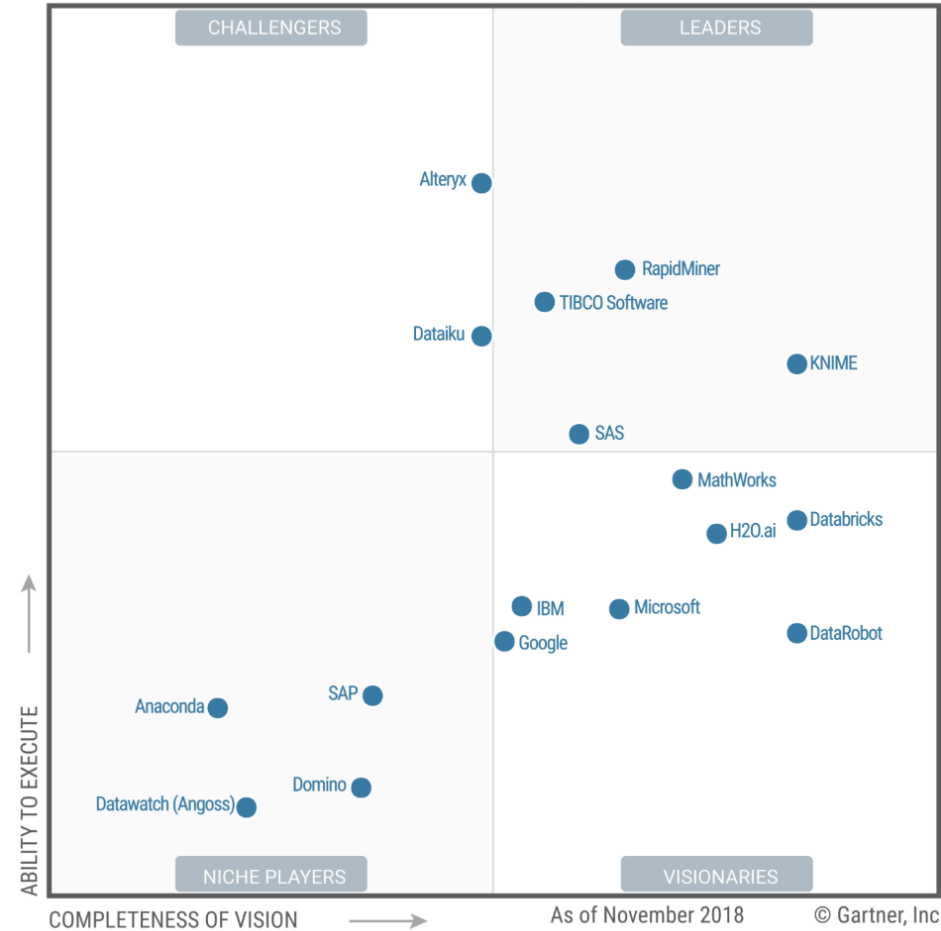


Figure 1. Magic Quadrant for Data Integration Tools



Source: Gartner (August 2020)

Figure 1. Magic Quadrant for Data and Analytics Service Providers



Source: Gartner (February 2020)

Figure 1. Magic Quadrant for Data Integration Tools



Source: Gartner (August 2020)

Ejemplo de Herramientas Utilizadas en BI/BA



Best BI Tools for 2020



Updated on 29th May, 20

<https://intellipaate.com/blog/top-bi-tools/>

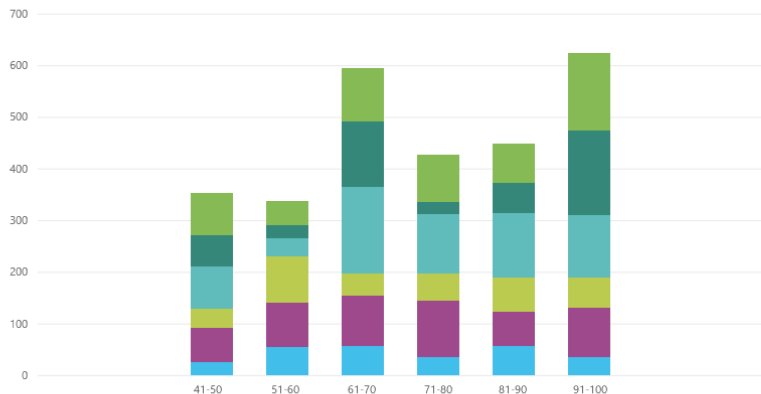


Travel Analysis

Ask a question about your data

Regional Load

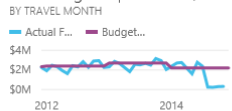
BY TERRITORY
Pacific Mountain Plains Central Northeast Southeast



Delayed Flights



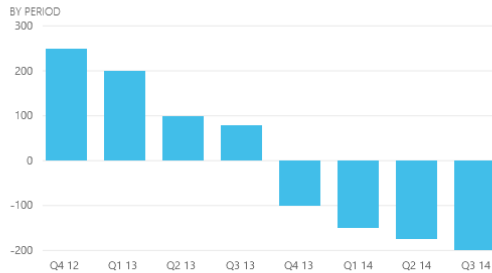
Actual Flight Expenditures, Bu...



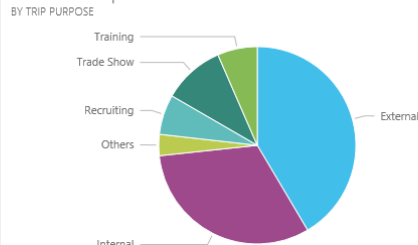
Budget Remaining



Sales

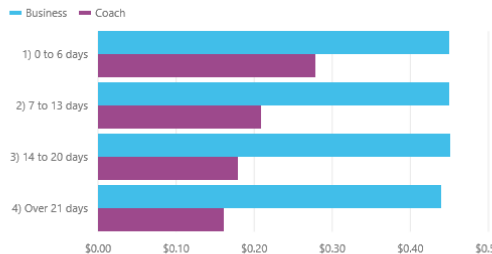


Number of Trips

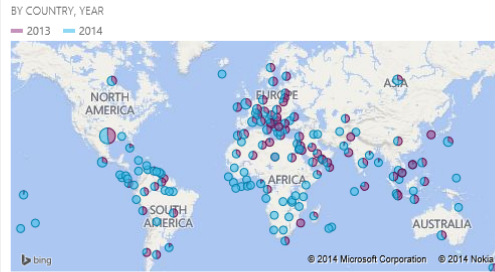


Average Cost Per Mile

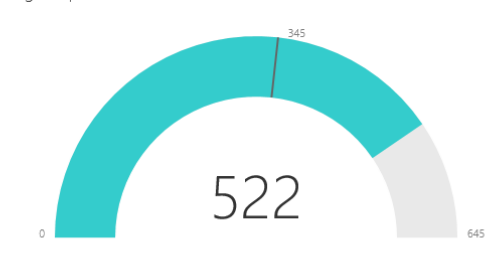
BY ADVANCED BOOKING CATEGORY, TRIP CLASS



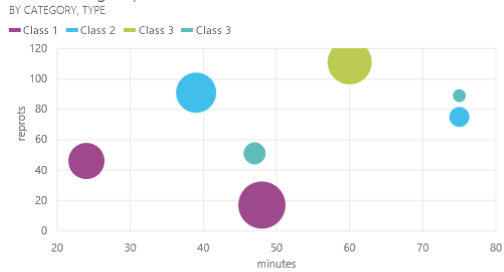
On Time Arrival



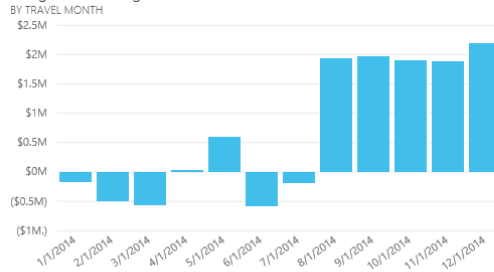
Flight Expense



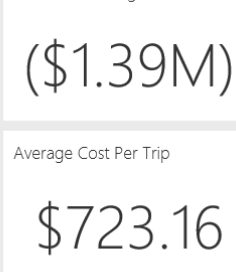
Rush Booking, Trip Class, Season



Budget Remaining



Variance to Budget



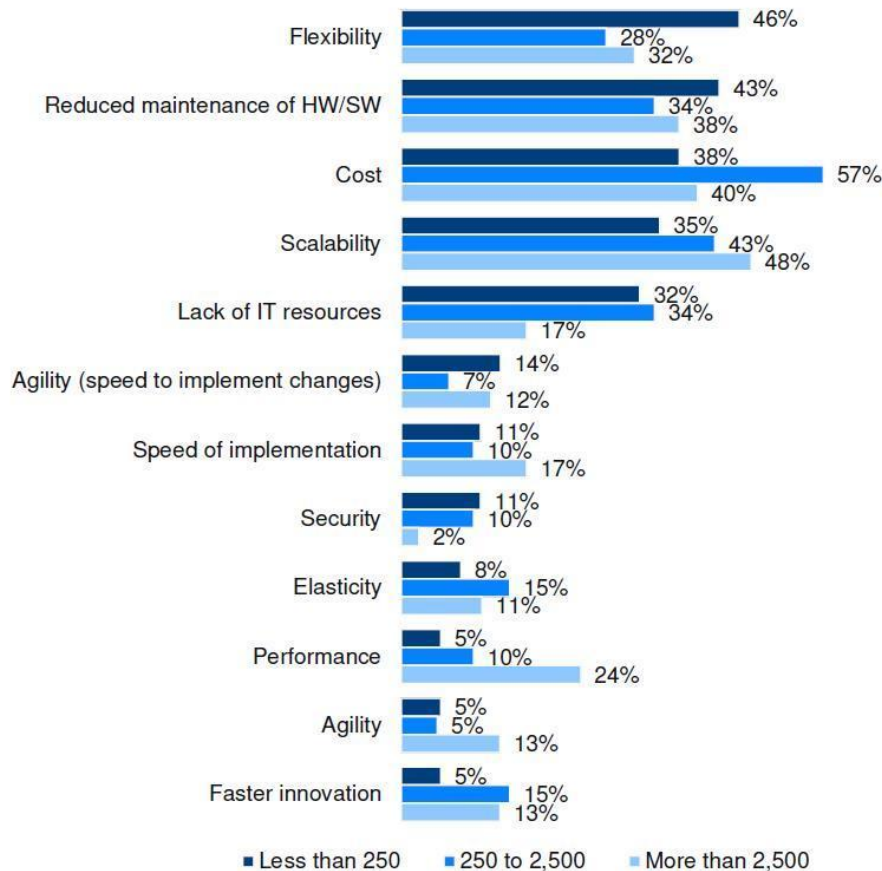
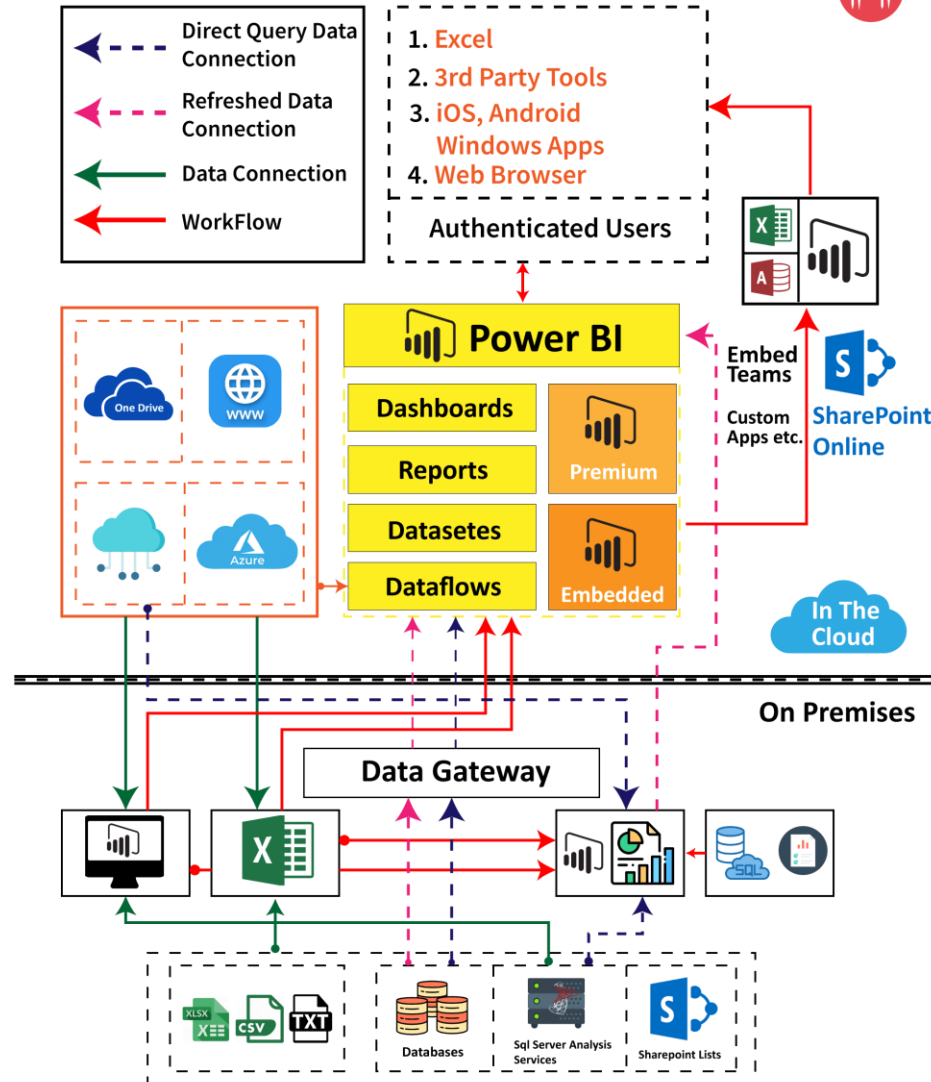
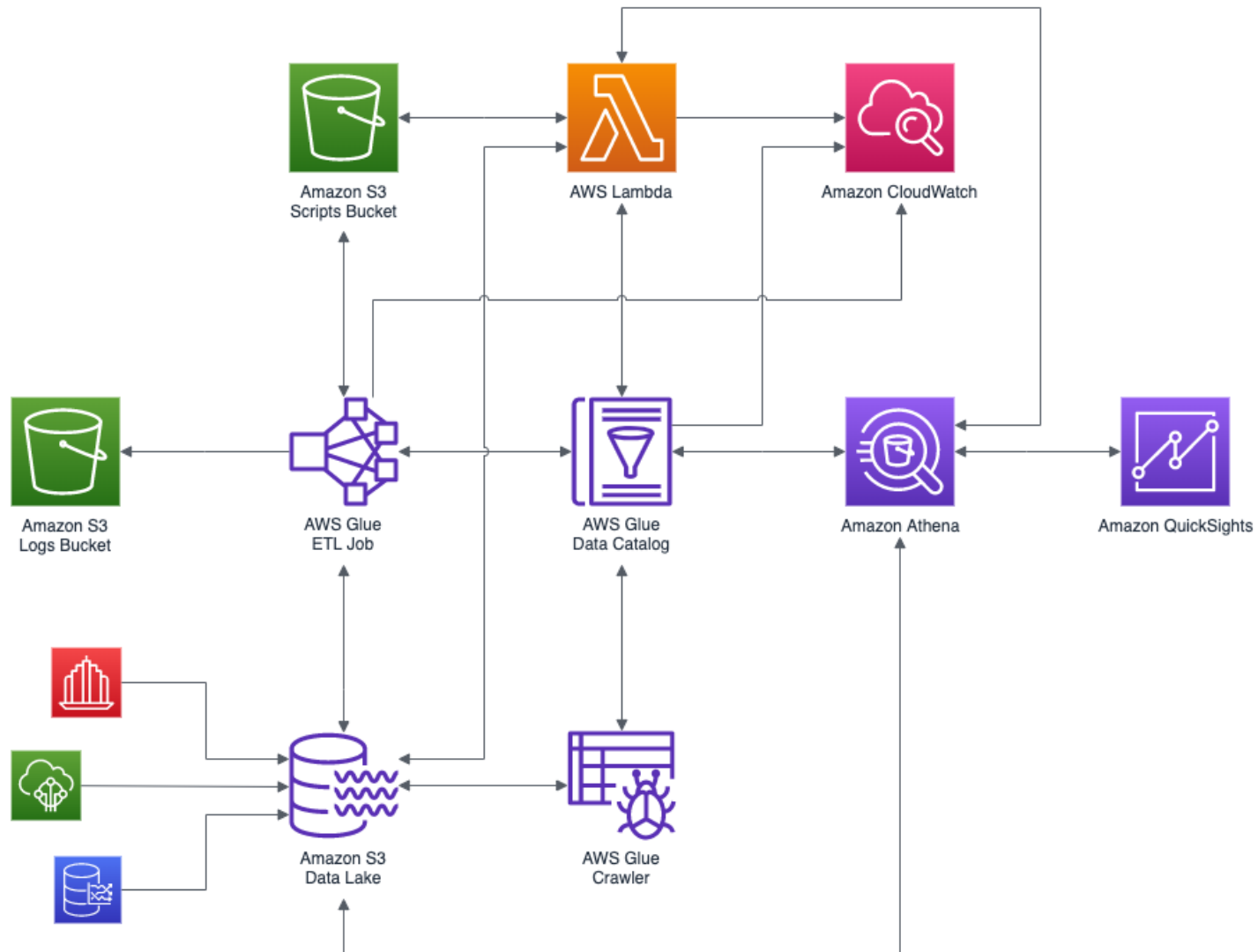


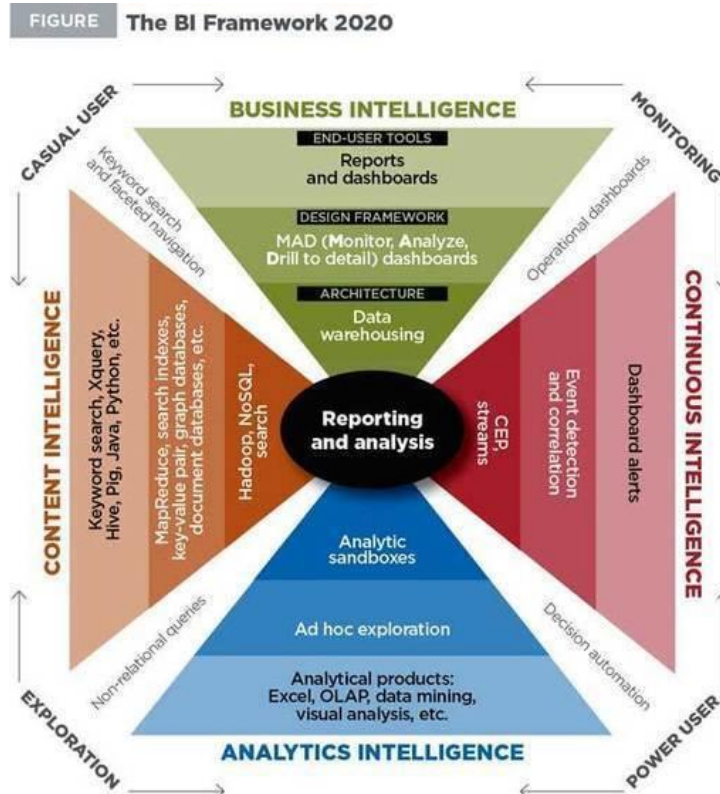
Figure 23: Most important reasons for implementing BI and data management in the cloud by company size (n=162)

Working of Power BI Architecture





Finalmente, tenemos que ir construyendo en nuestras compañías, una mirada sistémica



Cuidado, que hay muchas
soluciones y mezclas que se
pueden hacer

Presented by enliteAI

The diagram displays 18 startup sectors, each with a collection of company logos:

- AgriTech:** SMARTBOW, FARMDOX, METOS, smaXtec
- Automotive & Mobility:** AVL, virtual vehicle, EMOTION3D, INVENIUM, Tiftech, Parkbob
- Big Data & Enablers:** PANTRICE, dynatrace, EXOSCALE, mostly, econob, KIVU, JAROONA, TRICENTIS, SAILLABS, The Machine Company
- Enterprise:** Microsoft, Leftshift One, servicenow, SAP, 2021.AI, BOTBASE
- FinTech:** abacus.ac, BLU DELTA, COMMUNITY, DATA SCIENCE, kompany, NECCTON, SPECTRAPHORE, 42, Rebellion, SmartStream, Rabar Solutions
- AI Hardware:** AVISYSTEMS, EVOLVE
- Health & MedTech:** allcyte, contextualia, Diagnosis, Innophore, cogvis, eMurmur, MAIA, KMLVISION, mindcode, ares genetics, SCARLET, snics, Symptoma
- Image & Video Processing:** ANYLINE, CarCutter, forlan, KALEIDO, MoonVision, Trayzi, MAPPING EYE GmbH, Mapillary, AVISYSTEMS
- Industry & Robotics:** AGILOM, Catalysts, craftworks, D'ARIA, senzoro, FIDAG, humai, IIVISO, ORIGINAL+, robotart, FIDAG, Simple-AI, ToolSense, WorkHeld, QUOMATIC.AI
- InsurTech:** dolphin, twingz, prewave, riskine
- Language & Text:** ALPHAVY, cortical.io, ITELL SOLUTIONS, FACT AI, DeepOpinion, SIGNTIME, iTranslate, SEMANTIC WEB COMPANY, CALISTA
- Media & Social Analytics:** APOLLO.AI, tailorpost, eMentalist, STORYCLASH, UPDATEMI.
- Productivity:** Butleroy, DEEPSEARCH, grape, ONDEWO, ONLIM, xamoom, yodel
- Retail & eCommerce:** App Radar, FIDAG, 7LYTIX, ash Sophie, smartengine, RATEBOARD, Cozylo, Danube, ZOÖVÜ, DETEGO
- Sales, Marketing & PR:** emarsys, Flowtap, FRUX, intellyo, smec, kjero.com, TUNNEL32

The diagram illustrates the AI ecosystem stakeholders, categorized into six main groups:

- Universities & Educational Institutions:** Includes logos for Medizinische Universität Wien, FH Joanneum, FH Kärnten, FH Vorarlberg, TU Wien, FH Technikum Wien, Fachhochschule Salzburg, FH Technikum Wien, Universität Salzburg, and Universität Wien.
- Research:** Includes logos for Fraunhofer Austria, RISC Software GmbH, Salzburgresearch, dwh, sccch, vrv, AIT, IST Austria, and Amazon Prime Air.
- Organizations & Communities:** Includes logos for AI Austria, AI Styria, Vienna Deep Learning Meetup, Digital Factory Vorarlberg, Machine Learning Graz, and WDSG.
- Public:** Includes logos for Advantage Austria, Bundesministerium für Wirtschaft, Innovation und Technologie, itSV, BRZ TIP, and Stadt Wien.
- Accelerators, Incubators & Investors:** Includes logos for ELEVATE, FACTORY, eQventure, and various other incubators.
- Media/Conferences:** Includes logos for der brutkasten, derStandard.at, futurezone, and various conferences like APPLIED ARTIFICIAL INTELLIGENCE CONFERENCE.

European Machine Intelligence Landscape

Privately Held & Recently Exited Companies

Fundamentals

Artificial General Intelligence 	Natural Language Processing 	Computer Vision 	Speech Processing 	Optimisation & Simulation 	Training & Learning
--	--	----------------------------	------------------------------	--	------------------------------------

Autonomous Systems

Factory 2.0 	Robotics 	Autonomous Vehicles & Drones 	Smart Cities 	Smart Home 	Logistics, Mapping & Telematics
------------------------	---------------------	---	-------------------------	-----------------------	--

Human Machine Interaction

Smart UX 	Conversational Agents 	Visual Perception 	Touch & Gesture 	VR/AR 	Generative AI & Animation
---------------------	----------------------------------	------------------------------	--------------------------------	------------------	--

Search & Discovery

Search 	Visual & Audio Search 	Discovery 	Match Making 	Science 	Business Intelligence
-------------------	--------------------------------------	----------------------	-------------------------	--------------------	----------------------------------

Industries

Health 	Finance 	Adtech & Marketing 	Retail 	Media, Publishing & Education 	Security
-------------------	--------------------	-----------------------------------	-------------------	--	---------------------

Picks & Shovels

Hardware 	Data Sources 	Processing 	Deployment 	Visualisation 	Platforms
---------------------	-------------------------	-----------------------	-----------------------	--------------------------	----------------------

Big Data Landscape 2016

Infrastructure

Analytics

The collage displays logos for various data science and analytics platforms, organized into four main categories:

- Analyst Platforms:** Includes Palantir, AYASDI, Quid, and others.
- Analytics Platforms:** Includes Microsoft, Tableau, and others.
- Data Science Platforms:** Includes Continuum Analytics, Databricks, and others.
- Visualization:** Includes Tableau, Google Cloud, and others.

Applications

Sales & Marketing	Customer Service	Human Capital	Legal
 RADIUS Gainsight  bloomreach Zeta  livefyre blueyonder  Lattice  SALESFORCE  persado winter  AVINO  ACTIONIQ	 MEDALLIA  ATtenUnity  STELLA Service  NGDATA  Sift Science  Pipedrive  Tobias  eSight	 gild  Bartle  cruce  leverian  Brendo	 RAVEL  leverian  Brendo

NoSQL Databases

- Oracle Database
- Google Cloud Platform
- Microsoft Azure
- MongoDB
- Kerospine
- Cassandra
- Redis
- Aerospike

NewSQL Databases

- SAP HANA
- Clustrix
- Pivotal
- memsql
- paradigm4
- nuodb
- MariaDB
- VOLTDB
- cloudata
- others

The collage displays logos for various data science and analytics platforms, organized into four categories:

- BI Platforms:** Includes logos for Power BI, Amazon Redshift, Tableau, QlikView, Alteryx, and others.
- Statistical Computing:** Includes logos for SAS, SPSS, and MATLAB.
- Log Analytics:** Includes logos for Splunk, Sumologic, Logstash, and others.
- Social Analytics:** Includes logos for NetScout, DataSift, and others.

Ad Optimization
MediaMath
Integral AI
OpenX
rockwell
Adaptive
TheTradeDesk
Luminate
Discovery

Security
CYCLANCE
CounterTrack
ThreatMetric
AFRA SECURITY
PivotalOne
Remonday
CyberTrust
CyberTrust
CyberTrust

Vertical AI Applications
Easynote
Clara
KASIST

The collage displays logos for the following tools and categories:

- Graph Databases:** Neo4j, Gremlin, OrientDB.
- MPP Databases:** Teradata, Vertica, SAP HANA, Kognitio, Greenplum.
- Cloud EDW:** Amazon Redshift, Microsoft Azure Pivotal, SAP HANA, SAP BW, SAP S/4HANA.
- Data Transformation:** Alteryx, Informatica, Talend, SAP Data Services, SAP Business Objects.
- Data Integration:** SAP Data Services, SAP Business Objects, SAP HANA, SAP BW, SAP S/4HANA.

The collage is organized into four columns, each with a title and a collection of company logos:

- Real-Time:** Includes logos for Amazon, NCS Markets, OpenFlair, DataTorrent, and Alteryx.
- Machine Learning:** Includes logos for Intel, H2O, Skytree, and others.
- Speech & NLP:** Includes logos for Narrative Science, Capital, Huddle, and others.
- Horizontal AI:** Includes logos for IBM Watson, Clarifai, VIV, and others.

Data20 Oppor Tools Publisher Tools - Outbrain - mixpanel - Chartbeat	Govt/ Regulation - Socrata - OPENDV - Fiscal Note - emigna	Finance - Affirm - OnDeck - Kreditech - Lendip - Kabbage - Insikt - uora - Datamin
--	---	---

The image displays a collection of 20 startup logos, organized into five distinct categories, each with a title and a grid of logos. The categories and their respective logos are as follows:

- Management / Monitoring:** New Relic, NFDYNAMICS, Amazon CloudWatch, Octio, Splunk, and Tropicana.
- Security:** TANIUM, Okta, Cymru, Data Gravity, CipherCloud, VECTRA, and Bluebird.
- Storage:** Amazon S3, Google Cloud Storage, Microsoft Azure, Pure Storage, EMC, and Curology.
- App Dev:** Apigee, Google App Engine, IBM Cloud, Typecase, and CloudElements.
- Crowdsourcing:** Amazon Mechanical Turk, oDesk, CrowdFlower, and WorldVox.

The image displays a collection of logos for various companies in the data science and business analytics space, organized into four distinct categories:

- Search:** Includes logos for Amazon, KIWICLOUD, and Lucidworks.
- Data Services:** Includes logos for LRI, OPERA, and KAGGLE.
- For Business Analysts:** Includes logos for ClearStory, CIRRO, and Import.
- SM / Commerce:** Includes logos for Google, RUMetric, SUMALL, and others.

Education/Learning

- Yieldmo
- mark43
- KENSHO
- ADVA
- iSENTIUM

Life Sciences

- OpenDataSoft
- Quantopian

Industries

- OPOWER
- eHarmony
- RetailNext
- STITCH FIX
- TACHYUS

Cross-Infrastructure/Analytics

Open Source

[illegible]

Data Sources & APIs

The banner displays logos for various companies and organizations, categorized into six groups:

- Health:** Apple, Jawbone, Garmin, Fitbit, Withings, Validic, Human API.
- IOT:** Uptake, ThingWorx, InMobi, Somnify.
- Financial & Economic Data:** Bloomberg, Dow Jones, Yodlee, Premise, S&P Capital IQ, Quant, Xignite, CS Insights, iMacros, Goodhouse, Plume.
- Air / Space / Sea:** Planet Labs, Spire, Windows, Aircrea, DroneDeploy, SkyCatcher.
- Location/People/Entities:** Garmin, FourSquare, InsideView, Esri, Streetline, Captiva, Factual Places, Common Hubs, GlanceMeter, Basis, Sagent.
- Other:** Qualtrics, GA, DataCamp, Insight, Data Elite, HETP, Panjiva, Data.gov, The Data Incubator.

BIG DATA LANDSCAPE 2017

INFRASTRUCTURE

The diagram illustrates the evolution of data storage and processing technologies, categorized into three main areas:

- HADOOP ON-PREMISE:** Includes logos for Cloudera, Hortonworks, MapR, Pivotal, IBM InfoSphere, and BlueData Jethro.
- HADOOP IN THE CLOUD:** Includes logos for Amazon EMR, Microsoft Azure, Google Cloud Platform, IBM InfoSphere BigInsights, Treasure Data, Altimetric, Bole, Cazena, and CenturyLink.
- STREAMING / IN-MEMORY:** Includes logos for Amazon Kinesis, Databricks, Confluent, Strim, GridGain, MetaMarkets, DataTorrent, DataArtisans, Hazelcast, and Terracotta.

NOSQL DATABASES

- Google Cloud Platform
- ORACLE
- Microsoft Azure
- MongoDB
- redislabs
- Couchbase

NEWSQL DATABASES

- SAP
- Clustrix
- Pivotal
- Cockroach Labs
- memsql
- Aerospike
- VoltDB
- CytusData
- Truflow

GRAPH DBS

- Neo4j
- IBM
- ORACLE

CLOUD EDW

- Tera Data
- Vertica
- HETZGER
- Sicon
- Kognitio
- DSQL
- Dremio
- Amazon Redshift
- Google Cloud Platform
- Microsoft Azure
- Snowflake
- Informatica

DATA TRANSFORMATION

- talend
- pentaho
- alteryx
- TRIFRICA
- tomr
- Paxata
- StreamSets
- UNIFI

DATA INTEGRATION

- informatica
- synlogic
- Segment
- TEALUM
- IBM
- original
- alibaba
- skyhigh
- podium data
- ZALONI
- colibra
- glintly
- Import.io
- Stitch

DATA GOVERNANCE

- informatica
- IBM
- Alation
- Waterline data

MGMT / MONITORING

- amazon
- New Relic
- amazon dynamodb
- acinfo
- WAVEFORM
- unicon
- plunkr
- Yrrocano
- paperkit
- Numberly

The image displays a collection of logos for various cloud and data services, organized into five categories with orange headers. The logos are arranged in a grid-like fashion, with some categories having more logos than others.

- STORAGE** (4 logos):
 - Amazon S3
 - Google Cloud Platform
 - Microsoft Azure
 - ALLUXIO
 - minio.io
 - IBM Cloud Storage
 - Cloudinary
 - COHO
 - pantheon.io
- CLUSTER SERVICES** (4 logos):
 - Amazon EMR
 - kubernet.es
 - elastic
 - docker
 - MESOSPHERE
 - Core OS
 - datastax.com
 - psqlm.com
- APP DEV** (4 logos):
 - lightbend
 - Google App Engine
 - IBM Bluemix
 - rainforest
 - CRACK
- CROWDSOURCING** (4 logos):
 - amazon.mechanicalturk.com
 - upwork
 - WorkFusion
 - Cloud Bee
- HARDWARE** (4 logos):
 - Google TPU
 - ARM
 - PERVIA SYSTEMS
 - Graphcore
 - MYTHIC
 - CONEXA
 - NVIDIA
 - SCORTEX
 - Movidius

ANALYTICS

DATA ANALYST PLATFORMS

- Microsoft
- Pentaho
- Alteryx
- Digital Reasoning
- QlikView
- AYASDI
- MATIVVO
- Datameer
- Quid
- ClearStory
- OrigamiLogic
- interana
- Bottlenose
- ARIMO
- ENDOR
- MODE

DATA SCIENCE PLATFORMS

- IBM
- KNIME
- dataiku
- DOMINO
- yhat
- rapidminer
- CONTINUUM ANALYTICS
- Alpine
- Angeos

BI PLATFORMS	VISUALIZATION	VERTICAL ANALYTICS	STATISTICAL COMPUTING	DATA SERVICES
 Microsoft  Domo  Amazon Web Services  Looker  Alteryx  Arc4Data  GoodData	 Tableau  Google Cloud Platform  Qlik  Celonis  Pentaho  ZEPH  Chartio  Kogniton	 Predix  Gigamon  CAPE  Uptake  Orbit Insight  Tachyus  Allium  DeloRama	 SAS  SPSS  MATLAB	 Palantir  Verintec  Operant  Kaggle  Finical  DataKind

MACHINE LEARNING

- IBM Watson
- Google Cloud Platform
- H2O
- DataRobot
- connected relevant
- visenze
- bonsai
- dataREPA
- deepgenetic
- dataLion

HORIZONTAL AI

- IBM Watson Cortana
- Face++
- Voyager
- Affective
- Chronon
- OSARO
- curious AI
- sentient
- clarifai
- Service-ITG
- Cognitive Scale
- Numenta
- ETETUM
- curious AI
- VISION

SPEECH & NLP

- Google Cloud Platform
- twilio
- Narrative Science
- semantic search
- NaturalCloud
- ARRA
- Talkia
- maibao
- sniips
- Gridspace
- SoundCloud Inc

SEARCH

- Jaumate
- Elastic
- Oracle
- SumoLogic
- Loggly
- Algolia
- Logstash
- Swiftify
- MAANA
- AlphaSense

LOG ANALYTICS

- Elastic
- Oracle
- SumoLogic
- Loggly
- Algolia
- Logstash
- Swiftify
- MAANA
- AlphaSense

SOCIAL ANALYTICS

- Hootsuite
- Sprinklr
- NetBase
- DataSift
- Truex
- Synthesio
- Smilerach
- Bitly

WEB / MOBILE / COMMERCE ANALYTICS

- Google Analytics
- Mixpanel
- Amplitude
- Sumall
- Airta
- Retention
- Sisgop
- Granify
- Cusora

APPLICATIONS – ENTERPRISE

HUMAN CAPITAL
 HumanVue entelo
 hiQ GIGSTER
 flexipay
 WorkdayWISD
 Quaderno
 Stellia
 pymetrics

LEGAL
 RAVEL
 Seal
 GIGSTER
 WISD
 Brevia
 RISS

FINANCE
 anaplan
 Urovo
 bdmart
 HANA
 TRADESHIFF

ENTERPRISE PRODUCTIVITY
 slack
 Facebook
 ORACLE
 sun, at, diffbot
 Claria talia
 buttrick KASIST

BACK OFFICE AUTOMATION
 CyberScience
 apricity
 AppZen

SECURITY
 TANIUM
 CYRANCE
 StackPath
 DARTTRACE
 illumio
 C00642
 Vectra
 ThreatMetrix
 DataGravity
 CyberCloud
 Guardian
 ANOMALI
 sift science
 Gossamer
 OneKode
 Recorded Future
 BlueTalon
 Securonix
 Trend Micro
 Feedzai

APPLICATIONS – INDUSTRY

ADVERTISING AppNexus critico mexcoad fund Xid OpenMOAT distillery L3 algorithms L3 Leventhal TAPPS DYNAMIC VUE Chopra DataSift Pumpkin Pulsebots Vindex	EDUCATION K12 Clever Declara kidaptive PANDORA knowU Synapse	GOVERNMENT Socrata OPENGOV mark43 FiscalNote OpenDataSoft	FINANCE - LENDING OnDeck Kredtech ANVANT TALA MoneyLion Trueaccord cignifit Active AI	FINANCE - INVESTING Dataminr KENSIC INVEST NUMERA ISENTUM clarifymoney Adverse ASISA Rebulet	REAL ESTATE Opendoor VTS CREDIO reonomy COMPASTK	INSURANCE Intonimile Ironnode CYENCE Shift Technology Tractable
---	--	---	--	---	--	---

HEALTHCARE

- FLATIRON
- HealthTap
- Gingerio
- COTA
- ACure
- 23andMe
- enlitic
- aventus
- mySugr
- IMAGEN
- Transcarent
- AYRUS
- MFABIO
- color
- oora
- zymogen
- Benzypher
- REALiTE
- Phosphorix
- citrine
- TwoX
- Alaris
- slap
- mydental

LIFE SCIENCES

- Standard
- color
- oora
- zymogen
- Benzypher
- REALiTE
- Phosphorix
- citrine
- TwoX
- Alaris
- slap
- mydental

TRANSPORTATION

- UBER
- TESLA
- Zipcar
- CLEARPATH
- drive.ai
- pitmotus
- hexar
- comma.ai
- GO
- OTTO
- CivilMaps
- NIO
- tesla
- prospero

AGRICULTURE

- FARMERS
- farmstead
- FarmLogs
- AGRODRIVER
- mavrx

COMMERCE

- instacart
- STITCH FIX
- RetailNext
- HowGood

OTHER

- eHarmony
- stem
- refurb
- robotics
- RECYCLE
- collect
- springer
- duetto
- Umbel
- Second
- Structures
- Joystick

CROSS-INFRASTRUCTURE/ANALYTICS



OPEN SOURCE

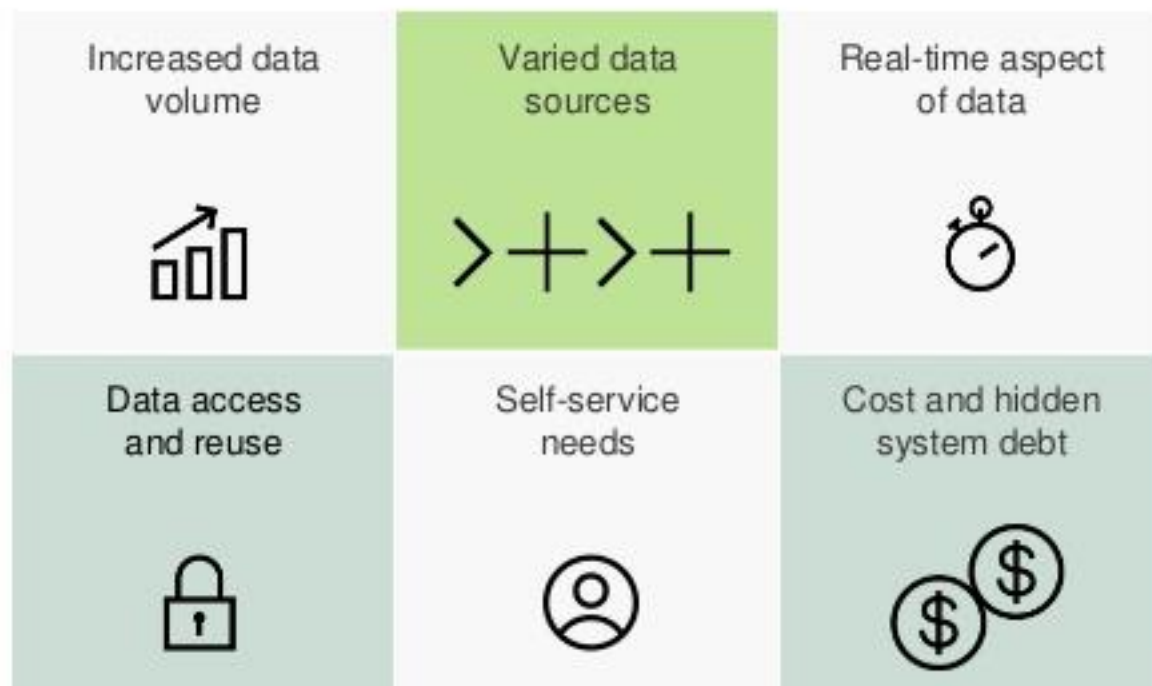
The banner displays logos for various technologies, organized into functional categories:

- FRAMEWORK**: Hadoop, MapReduce, Tez, Flink, YARN, Mesos, Spark, CDAP.
- QUERY / DATA FLOW**: Spark SQL, Presto, SLAMDATA, Google Cloud Dataflow.
- DATA ACCESS**: Cassandra, Nifi, MongoDB, Aerospike, Redis, CouchDB, Greenplum, Riak, HBase, Spanner, Accumulo.
- COORDINATION**: Talend, Apache Zookeeper, Ambari.
- STREAMING**: Spark Streaming, Flink, Kafka, Beam, Druid, Storm.
- STAT TOOLS**: Python, ScalaLab, Neo4j, FeatureFu, D3.js, SciPy.
- AI / MACHINE LEARNING / DEEP LEARNING**: TensorFlow, Caffe, CNTK, DM TK, Keras, PyTorch, VEEVA DIMSUM, DL4J, Aerolve.
- SEARCH**: Elasticsearch, Solr.
- LOG ANALYSIS**: kibana, logstash.
- VISUALIZATION**: BEAKER, Rodeo.
- COLLABORATION**: jupyter, Anaconda.
- SECURITY**: Knox, Sentry.

DATA SOURCES & APIS

DATA RESOURCES

A challenging big data landscape



MACHINE INTELLIGENCE 2.0

AGENTS

PROFESSIONAL	PERSONAL	OS INTERFACES

AUTONOMOUS SYSTEMS

AIR	GROUND	SEA	INDUSTRIAL

ENTERPRISE

SECURITY / FRAUD	HR / RECRUITING	SALES	MARKETING	CUSTOMER SUPPORT	INTERNAL INTEL	MARKET INTEL

PLATFORMS

RESEARCH / AGI	FULL STACK	MACHINE LEARNING	INDUSTRIAL IOT	AUDIO	VISION	DATA ENRICHMENT

INDUSTRIES

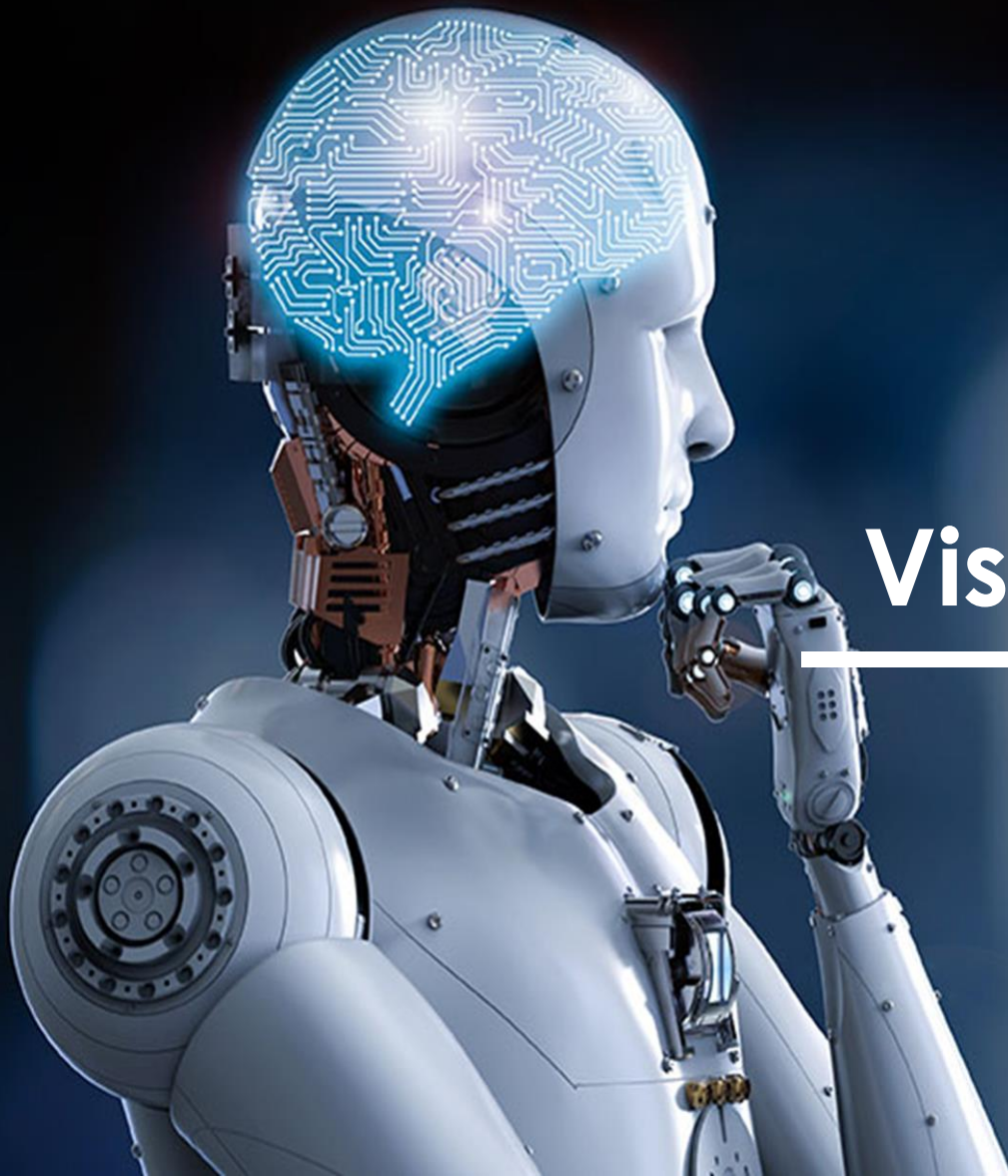
ADTECH	AGRICULTURE	FOR GOOD	RETAIL FINANCE	LEGAL	MATERIALS & MFG	HEALTHCARE

INDUSTRIES (CONT'D)

EDUCATION	TRANSPORT & LOGISTICS	INVESTMENT FINANCE

TECH USER TOOLS

DATA SCIENCE	MACHINE LEARNING	OPEN SOURCE



Visión de Futuro

<https://www.youtube.com/watch?v=NrmMk1Myrxc>



La **integración de tecnologías** como: visión artificial, fusión de varios tipos de sensores, algoritmos de deep learning y la nube, todo esto hace posible un **negocio totalmente distinto**.



TIPs

#tips

#ENFOQUE_NEGOCIOS_Y_VALOR

#EQUIPO_HUMANO_MULTIDISCIPLINARIO

#PROYECTOS_CORTOS_E_ITERATIVOS

#ENFOQUE_FLEXIBLE

#COMUNICAR_RESULTADOS_E_ITERAR

#ENCONTRAR_LOS_DATOS_APROPIADOS

#GRACIAS

WWW.

Select Q

