Calo Challenge: An overview on how ML is used in Industry

Google Cloud Partner

HOVERTURE
A SYMPHONIE PRIME COMPANY

SYMPHONIE PRIME
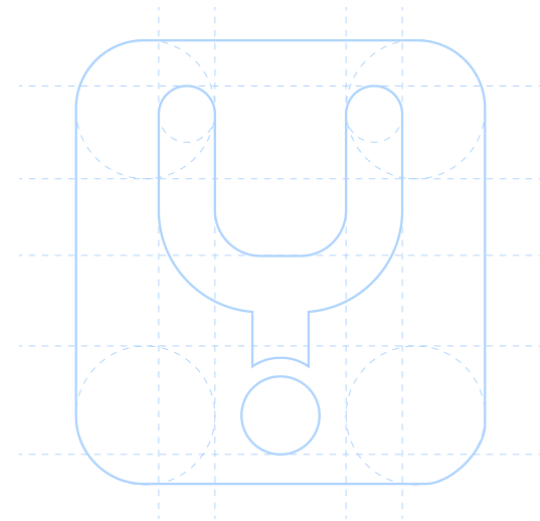
ML in industry
Who I am

- Serena Palazzo, PhD in particle physics, ex postdoctoral researcher, ex CERN

What I have done during my research activities:
- Working for the ATLAS experiment at CERN
  - MicroMegas test for ATLAS upgrade
  - Top quark physics. ex-convener of Top reconstruction, ex Monte Carlo production Manager, ex Derivation framework Manager, ex Muon liaison
  - Fast Calorimeter simulation: GANs implementation, test etc

What I am doing:
- Working for a Consulting company as:
  - Team Leader of Data Science and AI group within the Big Data area
  - Architectures definition, projects idea and implementation of AI and data related projects

Symphonie Group | **OUR FIVE ELEMENTS**

HOVERTURE

HOVERTURE

System integration
Cloud Transformation
Big-data & Smart Analytics
Applied Intelligence
Security & Compliance

RAPSODOO odoo

Business Cloud Platform
Vertical Business Application
Interoperability Framework

BIT2WIN

Omnichannel Sales Suite
Customer Engagement Platform
Post-Sales & Loyalty
Voice of Customers

SEEDBLE

Corporate Innovation
Organization Design
Ecosystem Strategy
Startup Acceleration
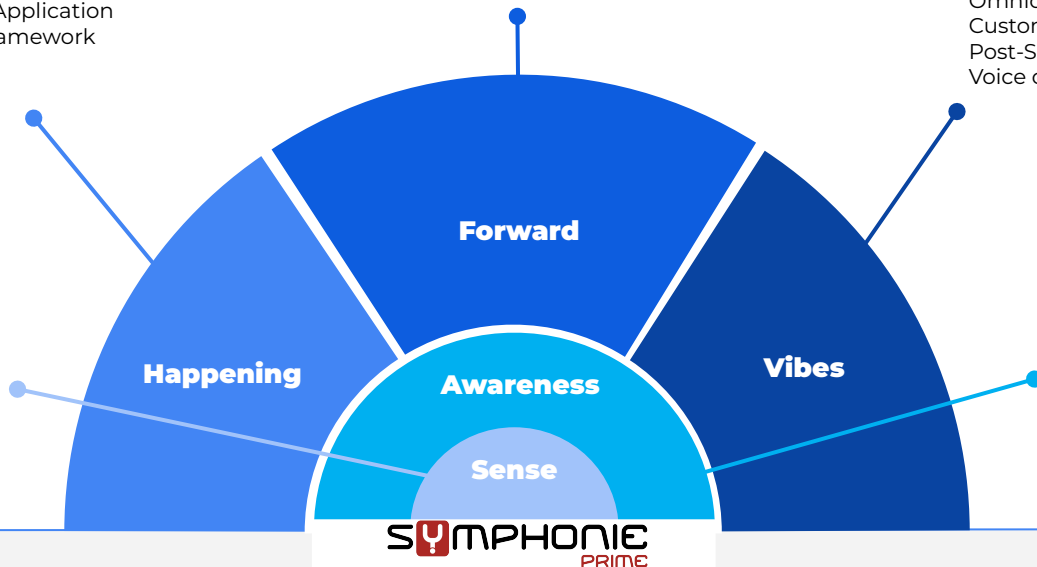
Brand Strategy
Customer Experience Design
Digital Customer Engagement
CMS Development
Digital & Social Marketing

Forward

Happening

Awareness

Vibes

Sense

SYMPHONIE PRIME

Italy   Germany   Iberia   Switzerland

Our Distinctive Factors

❏ **15+ years of experience** in Cloud Transformation and Data modernization

❏ Business Driven Approach encouraging **measurable results**

❏ Multi-Country, Multi-Language, Multi-Cultural

❏ Cross Industry Experience

HOVERTURE

We Are Hoverture
Our value proposition

CLOUD TRANSFORMATION

- Cloud Strategy & Architecture Design
- Infrastructure modernization
- DB & DWH Modernization
- Application Modernization & Cloud native development
- SAP Migration to Cloud

SECURITY & COMPLIANCE

- Cybersecurity assessment, remediation planning and execution)
- Cloud security
- Data security
- Digital identity
- Governance & GDPR Compliance
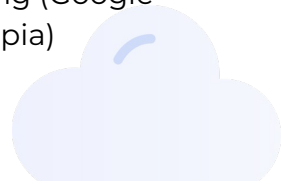
Our Services

BIG DATA, SMART ANALYTICS & AI

- Big data architecture set up and optimization
- Data quality & Enrichment
- Analytics modeling development (leveraging ML and AI)
- Data Visualization
- Analytics Lab

ENTERPRISE COLLABORATION EXPERIENCE

- Collaboration innovation through Google WS
- Employee experience integrated offering (Google ws, Lumapps, Itopia)

Contents

Big Data and AI in industry

⬡ **Big Data**

Big data is mostly characterised by large volumes of different types of data

These data are so large that traditional data processing software cannot handle them

But these huge volumes of data can be used to tackle business problems you might not have faced before

Cloud migration make it possible to have more simplicity, flexibility, scalability and velocity in every geographic area

Big data: What caracterize Big Data
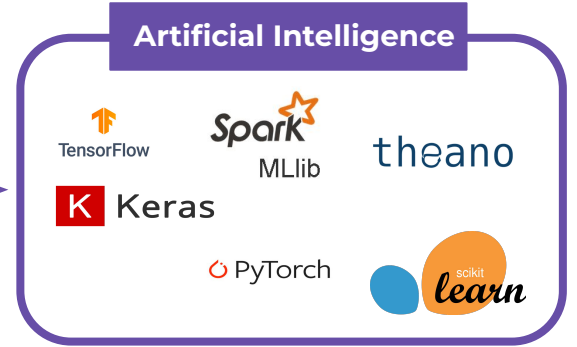
HOVERTURE

3Vs of Big data

1. **Volume:** quantity of generated Big data. Order of magnitude of **Terabytes** or **Zettabytes**
2. **Variety:** diversity of formats and structure
   a. Big data include not only **structured** data but also **unstructured** data like images, video, blog posts etc
   b. Variety refers also to the **sources**: some data are generated by users, other automatically from machines
3. **Velocity:** it refers to the velocity at which data are generated and processed. Big data are often available **real-time**

SYMPHONIE PRIME

Big data & AI: Relation between Big Data and AI

Big data & AI: Three main roles in the Big data & AI realm

DATA ENGINEER	DATA SCIENTIST	DATA ANALYST
Proper understanding of different data warehousing concepts & ETL (Extract, Transform, and Load)	Proficiency in statistical & analytical skills	Sound knowledge of analyzing granular data for better insights
This profile requires preparing data through Big Data technologies like Hadoop, Spark, etc.	This profile requires an intensive understanding of Machine Learning models, Deep Learning principles, and data science concepts.	This profile requires intensive programming knowledge to extract granular data.
They are responsible for building data architecture and pipelining.	**They are responsible for creating advanced data analytics to help executives make data-driven decisions.**	**They are responsible for developing data insights and reporting them through data visualization.**
They deploy statistical models through large-scale data processing tools like Hadoop and Kubernetes for further data operations.	They optimize the data and then deploy predictive ML models involving strategic planning and ad-hoc analysis.	They deploy program-oriented analysis of data and use data visualization tool like Tableau.

SYMPHONIE PRIME

Big data & AI: Cloud at the heart of the world

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction

Main characteristics of a Cloud

- High flexibility and adapt services to match your ever-changing goals and needs

- Cost effectiveness
  - Pay-as-you-go model
  - IT infrastructure and Data Security managed by the cloud provider

- High availability, fault tolerance and disaster recovery

Cloud provider leaders

Google Cloud Platform
aws
Microsoft Azure

Cloud Service Providers

Big data & AI: Why did we choose Google Cloud

- **Google** made massive investments on important aspects of Big Data & Analytics:
  - Data Lakes & Data Warehouses
  - Real-Time (stream) and batch processing
  - Advanced Analytics
  - Artificial Intelligence
  - Hybrid and multi-cloud environments
    - Scalable infrastructure to merge DWH and DL into few powerful tool (BigQuery/Big Table)

- Has one of the most efficient and reliable **storage** solution (Google Cloud Storage) able to perform significant saving processes
  - Google, by developing MapReduce and HDFS components spawn the Apache Hadoop Framework (the first open-source software utilities aiming to solve problem involving massive amounts of a data)

- Committed to comply **General Data Protection Regulation** (GDPR)

- Support development of **AI Algorithms** in low-code mode with advanced tools

- Is constantly on the **cutting edge** of technology looking for new ways to disrupt and innovate

Big data & AI: Our technology offering

COLLECTION

STORAGE

QUALITY / MNGMT/ ENRICHMENT

ANALYSIS

MODELLING & INSIGHTS GENERATION

VISUALIZATION

Google Cloud Platform

COLLECTION
- Kafka
- Apache Spark
- Apache Pig
- Python
- Pentaho Kettle
- Talend
- presto

STORAGE
- MongoDB
- PostgreSQL
- ORACLE
- hadoop
- Hive
- Hbase
- elastic
- cloudera IMPALA

QUALITY / MNGMT/ ENRICHMENT
- Talend
- Pentaho
- elastic

ANALYSIS
- Python
- Pandas
- Jupiter
- R
- elastic

MODELLING & INSIGHTS GENERATION
- scikit learn
- TensorFlow
- Keras
- Spark MLlib

VISUALIZATION
- Tableau
- Pentaho
- ORACLE BUSINESS INTELLIGENCE
- kibana

Big data & AI: GCP technology offering

COLLECTION	STORAGE	QUALITY / MNGMT/ ENRICHMENT	ANALYSIS	MODELLING & INSIGHTS GENERATION	VISUALIZATION
Pub/Sub	Big Query	Data Prep	Auto ML	Big Query	Data Studio
Migration Service	Data Proc	Data Flow	Tensor Flow	Looker	Looker
Data Transfer	Cloud BigTable	Cloud Spanner	Big Query	Tensor Flow	
Data Catalog	Data Fusion	Cloud Sql	CloudML		
		Data Fusion			

ML in industry: Overview

Going more in detail about Artificial Intelligence in Industry

- Companies are mainly divided into:
    - Product companies
    - Consulting companies

- The way they use ML is different:
    - Product companies use ML for their own product processes
    - Consulting companies use ML to deliver projects or sell employee expertise to other companies

AI in industry ho do we use it?

ML in industry: What are the most common applications

AI in industry is used for several applications

- Healthcare
- Manufacturing
- Marketing
- Editorial companies
- …

**But How AI and ML are used in the industry?**
- There are several techniques that are used for this purpose
  - Computer vision
  - Natural Language Processing
  - Forecasting predictions
  - Marketing analytics

HOVERTURE

SYMPHONIE PRIME

ML in industry: Transfer learning overview

➤ Consists of retrain a Network that has been already trained (pre-trained) with a huge dataset to perform a fine-tuning of the model by using our own dataset
  - ◆ The type input is the same (if a Network has been trained on images the new input has to be image dataset)
  - ◆ The idea is to leverage already learned features without train from scratch with a huge amount of data

Why use transfer learning?

➤ Save time and resource
➤ Have more accurate results by leveraging on the pretrained features learned
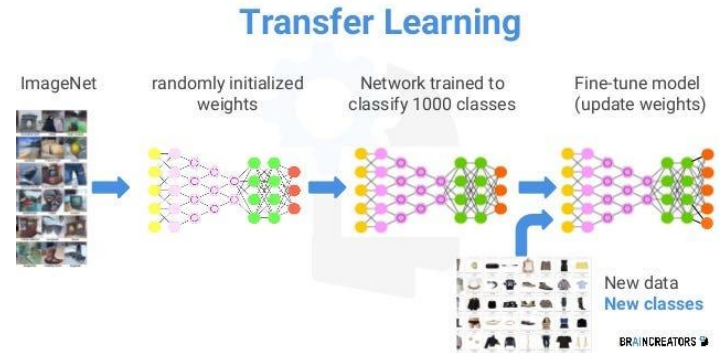
ML in industry: How to use transfer learning

➔ **Hugging Face** is an AI community, a large open-source community
➔ It provides state-of-the-art models for different tasks, mostly based on **Transformers**
◆ It has a vast number of pre-trained models, in particular for Image, Text and audio tasks

Computer Vision

What is Computer vision?

- Is a field of **Artificial Intelligence** that deal with how computer can gain high-level understanding from images and video

What do Computer vision tasks include?

- Methods for **acquiring**, **processing**, **analysing** and **understanding** images and frames of video

AI and Computer vision

- The development of sophisticated **Deep learning** techniques allowed to achieve challenging results on computer vision complex problems

ML in industry: How to get business value from computer vision

Tasks	People	Objects
Classification	• Facial recognition • Demographic characterization • Clothing color • Attributes, such as backpack or head coverings	• License plates • Color and attributes of vehicles • Logos • Identify objects in broad set of categories
Detect & track	• Track people entering in specific areas • Track people moving in-store • Behaviour interactions • Detect emotions	• Track vehicles in traffic • Objects detection • Extract text from images with OCR functionality • Explicit content moderation
Count	• Count people sitting, standing or walking • Count people waiting in a queue	• Count number of vehicles in a area
Heatmaps	• People movement • People crowding	• Vehicles movement • Vehicles crowding

SYMPHONIE PRIME

Model approaches:

- **Image recognition** approach within Computer vision field
  - Use of Deep Learning algorithms, in particular, **Convolutional neural networks** (CNN)
    - Most common architectures used to understand single images.
- **People recognition** within a frame and counting them
- **Crowding areas identification** through the analysis of a **heatmaps** obtained from people recognition

**People recognition**

Crowding recognition
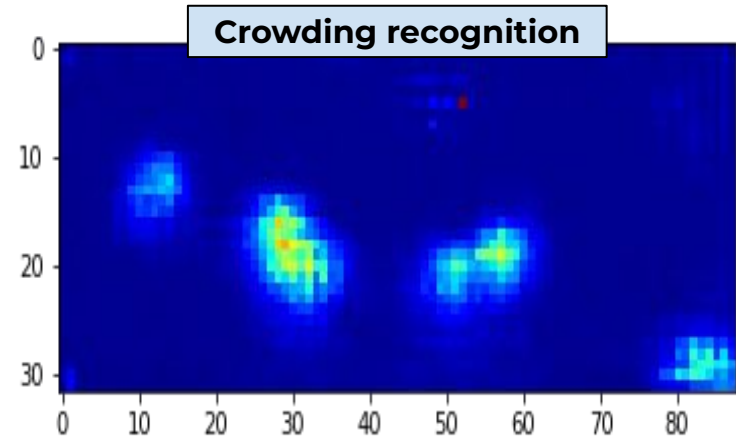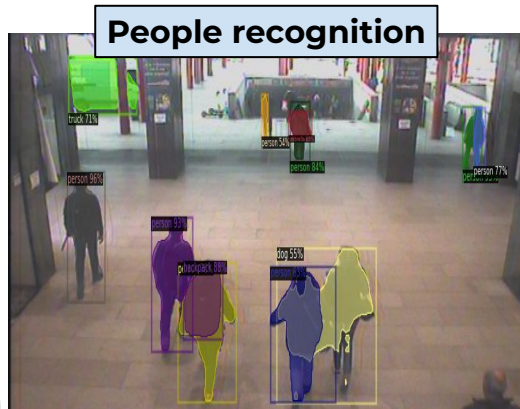
Model approach:

- **Video analytics** approach:
  - Use of artificial intelligence to complete various tasks
  - Apply Computer Vision and deep learning to video footage or live video streams
  - Use of deep learning algorithms **Recurrent Neural Network** (RNN) architecture
   - More suitable for series of frames to understand how are them related to each other
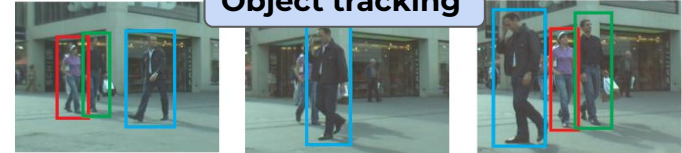
Video Analysis:

- **Objects detection**
  - Detect an object in a frame and classify it
- **Objects tracking**
  - Multiple object tracking: uniquely identify each object in a frame and track these objects across the entire video.
- **Action classification**
  - Recognize actions within a video

Object detection and labelling
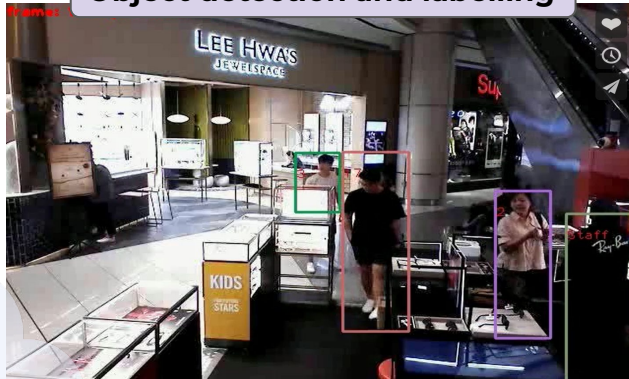
Object tracking

Natural Language processing
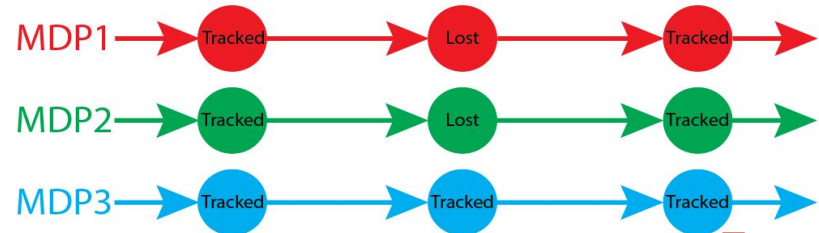
Natural Language processing

What is Natural Language Processing?

- **Natural Language processing** (NLP) refers to the branch of AI concerned with the ability of computers to understand **text** and **spoken words** in the same way as human beings can.

What do Natural Language Processing tasks include?

- Analysis of different aspects of text and speech like syntax, semantics, pragmatics, and morphology.

AI and Natural Language processing

- NLP combines computational linguistics (rule-based modeling of human language) with statistical, machine learning and deep learning models.

SYMPHONIE PRIME

How to get business value from NLP

Tasks	Use case
Classification	• Text classification • Content classification
Sentiment analysis	• Understand the overall sentiment expressed in a block of text. • Understand the overall sentiment expressed in social media post
Entity recognition	• Identify entities within documents (receipts, bills) • Identify entities and label by types such as person, organization, location, events, products and media. • Speech/Voice recognition
Conversational AI	• Virtual assistants, chatbots, search engines

Use case #1: Text classification

Model approach:

- **Natural Language processing** modelling approach
  - Allows to **classify text** by categorizing text into group of words
  - Automatically analyze text and then assign a set of predefined tags or categories based on its context.

Deep Learning Algorithms:

- **Long-short Term Memory** (LSTM) that is a **Recurrent Neural Networks** (RNN) architecture
  - These algorithms are suitable for text analysis

Text Classification

Acme Article

Acme Article → Technology

Foo Article → Sports

Bar Article → Entertainment
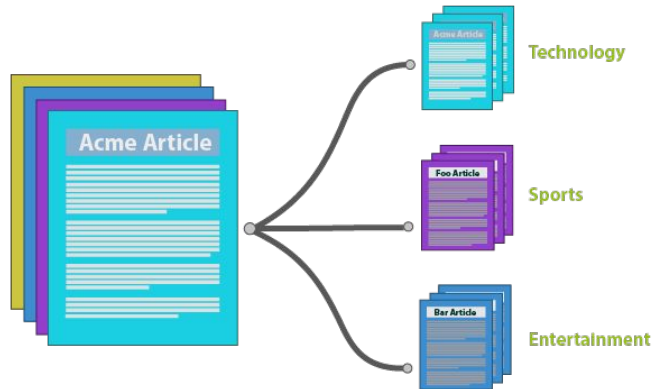
What is Sentiment Analysis?

- **Sentiment analysis** also referred to as opinion mining is an approach to Natural Language processing (NLP). It is the automated process that analyze a text and interpret the sentiments behind it.
- Through AI, algorithms can classify statements as positive, negative or neutral

What is Sentiment Analysis used for?

- It is used to analyze social media posts and online product reviews, as a way to track opinions, reactions, and ultimately improve customer service and experience.
- These tools are not only used for analysis purposes, but also for predictions.

My experience so far has been fantastic!

POSITIVE

The product is ok I guess

NEUTRAL

Your support team is useless

NEGATIVE

Use case #2: Sentiment analysis

Model approach:

- **Natural Language processing** modelling approach
  - Analyse text of social media posts to understand sentiment in it
  - Categorize sentiment within text as positive, negative or neutral
  - Use of Deep Learning algorithms, in particular, **Long-short Term Memory** (LSTM) that is a **Recurrent Neural Networks** (RNN) architecture
    - These algorithms are suitable for text analysis
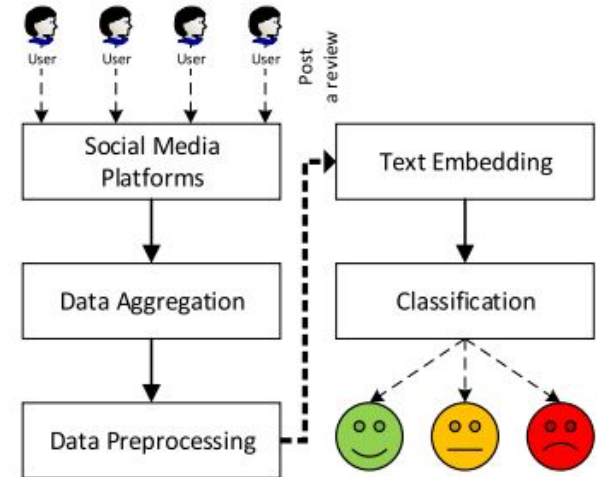
Use case #3: Chatbot with NLP

- A chatbot is a computer program that simulates and processes human conversation (either written or spoken), allowing humans to interact with digital devices as if they were communicating with a real person.

- AI + automated rules + NLP + ML → chatbot process data and deliver responses to request of all kinds.

Task-oriented (declarative) chatbots	Data-driven and predictive (conversational) chatbots
Single-purpose programs that focus on performing one function	often referred to as virtual assistants or digital assistant and they are much more sophisticated, interactive, and personalized
Using rules, NLP, and very little ML	contextually aware and leverage natural-language understanding (NLU), NLP, and ML to learn as they go
Handle common questions: queries about hours of business, simple transactions, etc..	Digital assistants can learn a user's preferences over time, provide recommendations, and even anticipate needs.
Engage end users in a conversational way	Can initiate conversations

https://www.oracle.com/chatbots/what-is-a-chatbot/

SYMPHONIE PRIME

Forecasting predictions

HOVERTURE
A SYMPHONIE PRIME COMPANY

Use case #2: Price forecasting

Predict natural gas price trends to make optimal short- mid- and long term purchase decisions, also to guarantee to customers the most stable and fair prices

Model approaches:

- **Forecasting model** approach, to predict future values based on previously observed values.
- Different algorithms considered for the training, to compare performances and choose the best one.
- Models suitable for short/mid or long-term price forecasting (based also on the granularity of data availability).
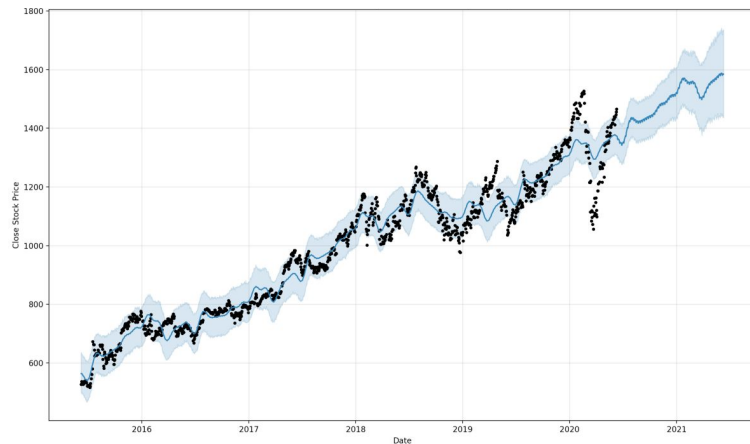
Marketing analytics

CDP overview: What is a CDP and how it helps

A tool that speeds up and simplifies access to **customer data**. Considered to be a pillar of **Marketing Technology**, with an estimated CAGR of 34% between 2020 and 2025
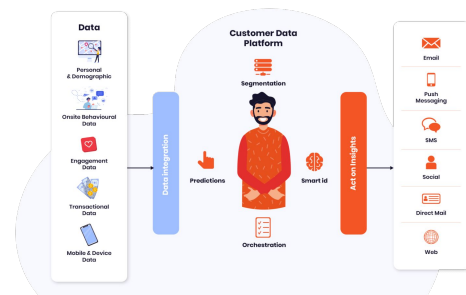
➢ It **unifies first party data** from different sources and allow **data enrichment** to create a **single customer view**

➢ A single centralized customer database accessible by **marketers** to manage interactions with customers

➢ A tool to build **personalized customer experiences**

Key benefits

➢ **Real time update** of customer profiles to capture changes in preferences and behaviors

➢ Valorize your data with **advanced functionalities** for customer analytics, AI and Machine Learning

➢ Easier to regulate consumers **privacy** and **consent** with identity governance tools (CPMP, CIAM..)

➢ Easier to trace and analyze **customer behaviors**, and create omnichannel **marketing campaigns**

➢ More efficient **communication** with customers: perform **personalized actions** through data activation

Customer Data Enrichment:

Provide a complete view of the customer with a unique, accurate and complete information

CUSTOMER DATA ENRICHMENT

Improve customer comprehension
their behaviours, their habits and preferences

Improve performances
of marketing, sales and business overall

Optimize Cost
to-acquire, to-serve, to-sell

KEY POINTS

◆ **BETTER INSIGHT** Thanks to information enrichment socio/demo, browsing, social behaviour, social, commuting/moving...

◆ **BETTER ACTIONABILITY** thanks to a better understanding of their client, their behaviour and their preferences

IS THE ENABLER TO IMPROVE PERFORMANCE OF BOTH CONSUMER AND TRADE MARKETING

WHAT KIND OF ENRICHMENT WE DO? some examples:

Socio Demo iwe integrate socio-demographic data to understand the behaviour of a customer's cluster compared to the various census variables

Social Presence to identify any correlation between customer's cluster and look-a-like cluster in terms of interaction, interests, Social Graph

Digital Behaviour to verify the existence of any correlation between cluster and digital initiatives (in terms of ads/performances/devices used/online preferences)

Mobile Behaviour* to enrich in an exponential way data used while, at the same time, analyze not only people's commuting/travelling/moving but also the conversion between online and not only offilne but also about physical point of sales**

*partnership con Telco Operator
** GCP Beta

Make the most of your data by **transforming** insights into Action

Breakdown Data Silos

Breakdown Data Silos to connect sources, leverage native ingestions and gain a **holistic** view of the customer

1 **Collect & trasform**

Analyze & visualize

Visualize and share real-time, **AI and ML-enriched insights** across the organization to empower data-driven business decisions and customer segments

2 **Analyze & visualize**

Activate in real-time

Personalize marketing and customer experiences by **activating data** in your marketing & customer service platforms to increase customer engagement and grow sales

3 **Activate in real-time**

HOVERTURE

SYMPHONIE PRIME

A step by step approach to enhance data and capabilities over time

➤ Customer Data Platform use cases give value at different levels
➤ The combination of use case outcomes, at different levels, gives a comprehensive view of customer needs

Level 3 Capabilities
- Demand Sensing
- Enhanced Customer Understanding
- Accelerating Qualification & Innovation

Level 2 Capabilities
- LTV/pLTV
- Predict Next Best Action
- Cluster Modeling

Level 1 Capabilities
- Conversion/Purchase Propensity
- Churn Propensity
- Self Service Marketing Analytics

Marketing use cases

01 — Marketing
Enhanced Consumer Understanding
Understand, conceptualize, and construct consumer profiles, including behavioral, attitudinal and other attributes.

02 — Marketing
Smart Consumer Segmentation
Create look-a-like segments for analytics and activation based on 1P data across multiple touchpoints

03 — Marketing
Personalized Marketing Communication
Action the insights on individual consumers by delivering them the next best action, in the right message, at the right stage of the consumer journey.

04 — Marketing
Attribution Modeling
Analyze performance of marketing and advertising efforts at a granular level (e.g., after completion of individual campaigns) to surface highest ROI channels and campaigns

05 — Marketing
Marketing Spend Optimization
Perform real time consumer marketing spend analytics to review and inform marketing budget allocation across marketing channels and audiences

06 — Marketing
Consumer Signals for Demand Sensing
Improve accuracy of sales and operation planning to help shape demand, optimize distribution and increase revenue.

07 — Omnichannel Experience
Financial Product Recommendations
Personalized product recommendations, 1:1 merchandising / rank assortment to cross-/up-sell, drive online revenue and reducing cost.

08 — Omnichannel Experience
Consumer Lifetime Value Optimization
Generate and action on prescriptive marketing insights aimed at extracting maximum per-consumer value over their lifetime and minimizing churn

09 — Omnichannel Experience
Propensity to Convert Modeling
Predict customers who are most likely to make a purchase, so that you can personalize communications with them.

10 — Innovation
Trend Spotting
Utilize consumer insights and analytics to understand trends to identify promising product innovation / renovation areas early.

11 — Innovation
Sentiment Analysis
Harness insights from consumer sentiment across social media platforms to inform your product innovation using AI/ ML.

12 — Innovation
Accelerating Qualification & Innovation
Accelerate analysis of consumer data for rapid qualification and speed to market.

Google Cloud

HOVERTURE

SYMPHONIE PRIME

And generative AI?

Generativa AI in industry

**Yes also generative AI is entering in industry**

- The most famous application is of course ChatGPT that can be embedded to create applications useful for industry
- ChatGPT is not the unique tool in the generative landscape, also Google is offering some alternatives

There are several applications in industry:
- Content creation
- Text summarization
- Minute organization
- Video and image creation
- Product description
- —

BARD & PALM2: Google Cloud brings Generative AI to organizations

Generative AI has unleashed a new variety of digital assistants, content creation tools, and applications, changing how apps are built, who can build them, and the capabilities end users expect from them.
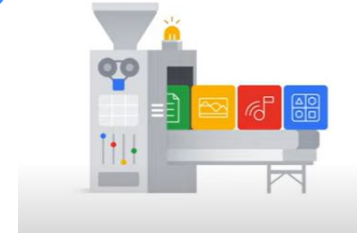
Vertex AI & Generative AI

Vertex AI and **Generative AI studio** to help building transformative applications powered by your data

Prompt design:
the quality of the input determines **the quality of the generated output**

New foundation models

3 new foundation models
- **Codey:** text-to-code model
- **Imagen:** text-to-image model
- **Chirp:** speech-to-text model

Prompt design & fine tuning

What Generative AI Studio provides:
- **Prompt design and tuning:** manually creating text inputs (or prompts) that inform the model without re-training it
- **Fine-tuning:** further training a pre-trained model on new data to have outputs with specialized results

Generative AI will transform

- **how marketers create content**
- **how marketers create campaigns**
- **how marketers generate original description of their products**
- **how marketers can provide a better customer support system**

ML in industry
Conclusions

- There is a wide range of applications of ML in industry
- Most of the applications in consulting are related to marketing analytics to help companies in understanding customer behaviour
- There are also applications in other fields as we have seen

- Still Not easy to bring the data-driven culture in the industry, it is a slow process but we are getting in there

THANK YOU

🌐 www.hoverture.com

in www.linkedin.com/company/hoverture

Backup

HOVERTURE
A SYMPHONIE PRIME COMPANY