# Tecnologías emergentes Algunos conceptos

### Sistemas de Información

Preparada por: Elizabeth E. Grandón, María Antonieta Soto, Pedro Campos

### AGENDA

- 1. Machine Learning
- 2. Big Data
- 3. Base de Datos Minería de Datos (conceptos)
- 4. Internet of Things (IoT)
- 5. Otras tecnologías emergentes

1. Machine Learning

# 1. Machine Learning

Aprende de la experiencia



Aprende de **los datos** 



Sigue instrucciones



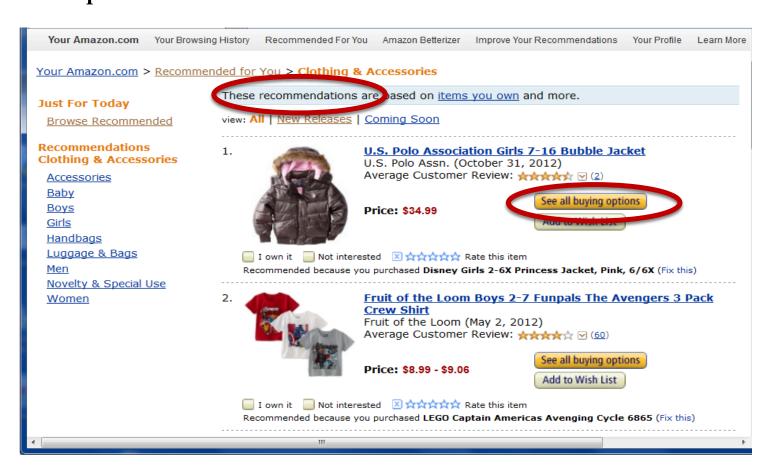
# ¿Qué es Machine Learning?

- Machine Learning (aprendizaje automático) es una rama de la Inteligencia Artificial que explora la construcción y estudio de algoritmos que sean capaces de aprender a partir de un conjunto de datos.
- ¿Qué se entiende por "aprender"?
  - Existen diferentes perspectivas filosóficas al respecto...
  - En general, en esta área se utiliza una perspectiva operacional:
    - Aprender implica mejorar un comportamiento (criterio de rendimiento) utilizando la experiencia pasada (datos).
- También llamado reconocimiento de patrones (pattern recognition)

# ¿Qué es Machine Learning?

- Un sistema de aprendizaje automático (ML), por tanto, involucra:
  - Un problema a optimizar (modelo)
  - Datos existentes (de los cuales extraer patrones → aprender)
  - Un algoritmo -eficiente- para optimizar el modelo
- ¿Para qué queremos que las máquinas sean capaces de aprender?

Permite crear sistemas que se adapten automáticamente a usuarios individuales.
 Retail personalizado



Permite crear sistemas que se adapten automáticamente a usuarios individuales. Retail personalizado

### Frequently Bought Together



Price For All Three: \$120.44



Show availability and shipping details

- This item: Data Mining: Practical Machine Learning Tools and Techniques, Third Edition (The Morgan Kaufm) Ian H. Witten Paperback \$42.09
- Data Mining: Concepts and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Sy
- ☑ Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites by Ma

• Permite crear sistemas que se adapten automáticamente a usuarios individuales





 Permite crear sistemas que se adapten automáticamente a usuarios individuales











#### Personalízalas y ordénalas a tu gusto

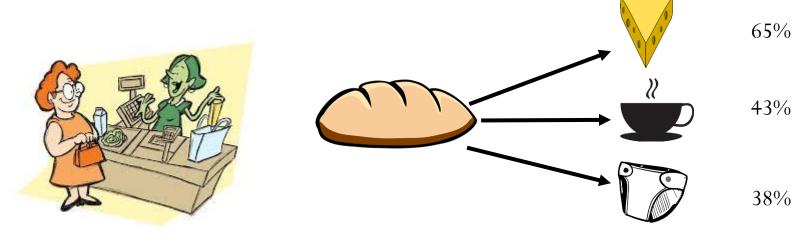
Sólo arrástralas y colócalas según el orden que quieres que se muestren.



#### **GUARDAR FAVORITAS**

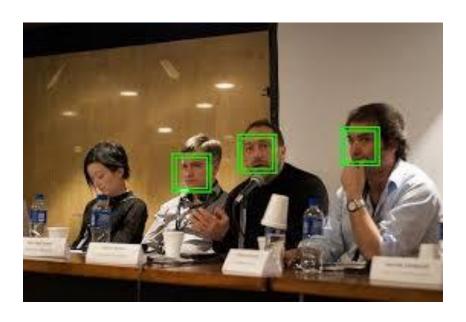


- Permite descubrir nuevo conocimiento a partir de grandes bases de datos (data mining)
  - Basket market analysis



- De las compras realizadas por mujeres adultas que incluyen pan:
  - 65% incluyen queso; 43% incluyen té; 38% incluyen pañales

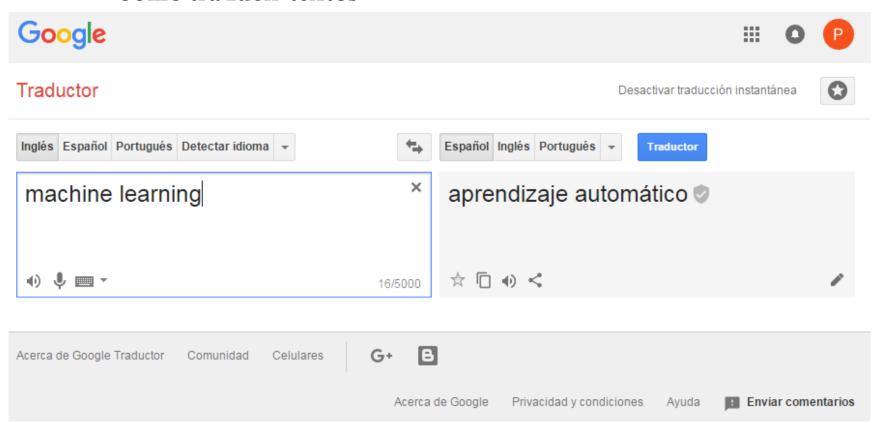
- Permite imitar ciertos comportamientos humanos y reemplazar tareas monótonas que requieren de cierta inteligencia
  - Como reconocer formas







- Permite imitar ciertos comportamientos humanos y reemplazar ciertas tareas monótonas que requieren de cierta inteligencia
  - Como traducir textos



- Permite imitar ciertos comportamientos humanos y reemplazar tareas monótonas que requieren de cierta inteligencia
  - Como manejar vehículos





### Interacción



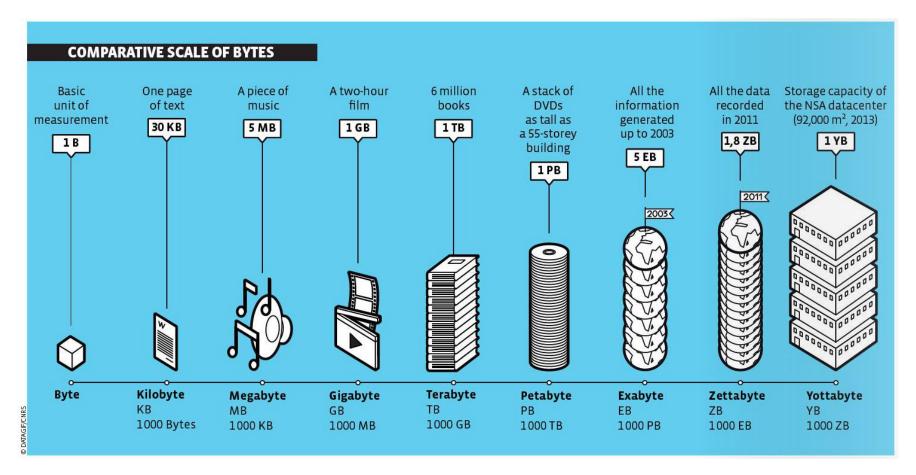
- Comente en qué tipo de industrias/empresas se podría utilizar la tecnología de Machine Learning?
- ¿Qué características deberían tener esas empresas?

# 2. Big Data

### 2. Big Data

- "Grandes volúmenes de datos" o "datos masivos" y todas las tareas que se pueden realizar con estos datos pero a gran escala
- Big Data no es solo datos: se reparte entre "los datos" y todos los procesos que se agrupan alrededor de estos datos, desde la recogida de los mismos, el almacenamiento y su posterior procesamiento o análisis, todo ello con el objetivo de extraer valor de los mismos
- Información que no puede ser procesada o analizada mediante procesos tradicionales. Se dispone de más datos de los que es posible manejar adecuadamente
- Los algoritmos "tradicionales" requieren ser modificados para dar tiempos de respuesta razonables

### Tamaño de los datos almacenados





( 43 TRILLION GIGABYTES ) of data will be created by 2020, an increase of 300

times from 2005



Volume

SCALE OF DATA



( 2.3 TRILLION GIGABYTES ) of data are created each day The

of Big

Data

FOUR V's

break big data into four dimensions: Volume,

Velocity, Variety and Veracity

4.4 MILLION IT JOBS



have cell



Most companies in the U.S. have at least

**100 TERABYTES** 

100,000 DIGABYTES I

The New York Stock Exchange

#### 1 TB OF TRADE INFORMATION

captures

during each trading session



**Velocity** 

ANALYSIS OF STREAMING DATA

By 2016, it is projected there will be

### 18.9 BILLION NETWORK CONNECTIONS

- almost 2.5 connections per person on earth



of data stored

#### Modern cars have close to 100 SENSORS

that monitor items such as

fuel level and tire pressure

As of 2011, the global size of data in healthcare was estimated to be

#### 150 EXABYTES

I 151 BILLION GIGABYTES I



**Variety** 

DIFFERENT FORMS OF DATA

### 4 BILLION+ HOURS OF VIDEO

By 2014, it's anticipated

WEARABLE, WIRELESS

**HEALTH MONITORS** 

there will be

are watched on-YouTube each month



are sent per day by about 200 million monthly active users

#### 30 BILLION PIECES OF CONTENT

are shared on Facebook every month



1 IN 3 BUSINESS

don't trust the information they use to make decisions



Poor data quality costs the US economy around

\$3.1 TRILLION A YEAR

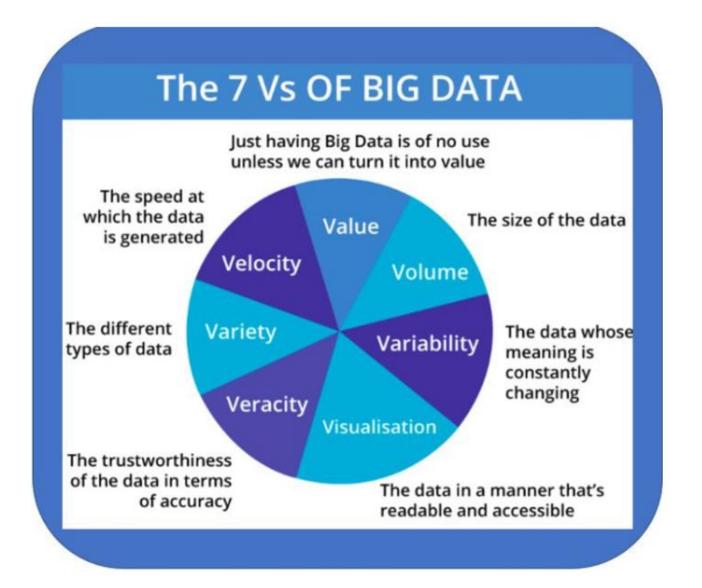


27% OF

in one survey were unsure of how much of their data was inaccurate

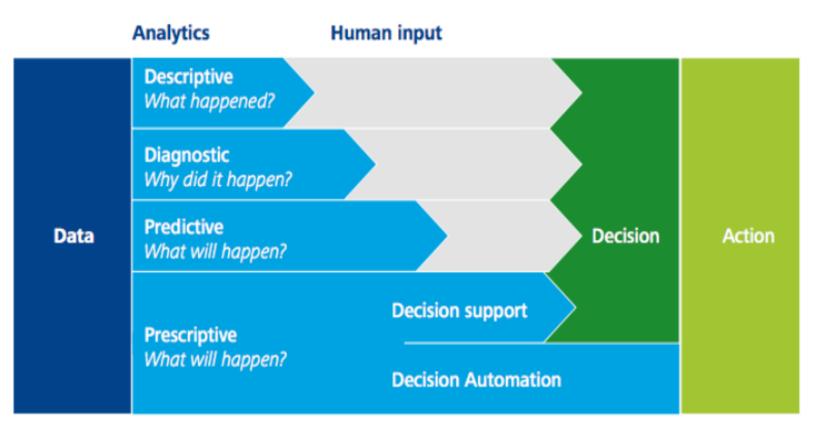
Veracity UNCERTAINTY

OF DATA



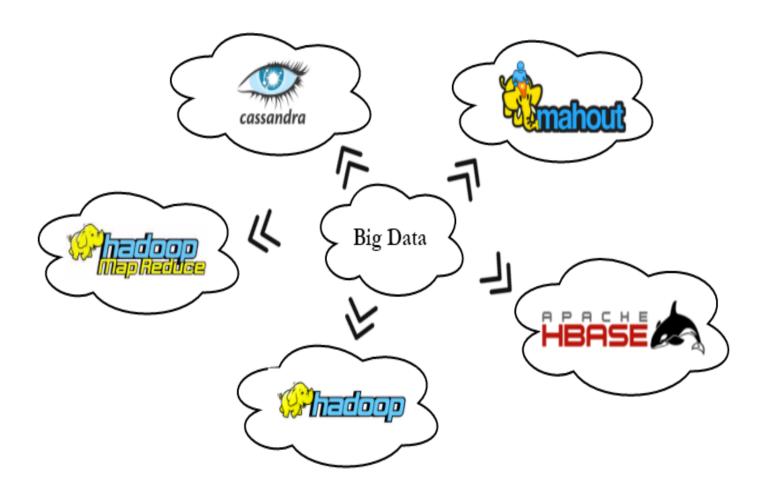
	1.Volumen	¿Cuántas personas están ahora mismo conectadas a internet? ¿Cuáles de ellas están viendo vídeos? ¿Y cuántas imágenes? ¿Cuántos millones de personas están en redes sociales en este preciso instante? la tecnología big data es capaz de manejar ese enorme volumen de datos
	2. Velocidad	Velocidad en cuanto al procesamiento de datos. Las interconexiones son constantes y esta tecnología es capaz de crear, almacenar y procesar datos en tiempo real.
	3. Variedad	Un comentario en Instagram, una visita a una tienda online, un formulario para descargar un ebook, una película vista en <i>streaming</i> No todos los métodos de análisis valen para todo tipo de datos, por lo que estos deben adecuarse a la naturaleza de cada información.
	4. Viabilidad	No todos los datos recopilados son útiles y aún menos si no se contextualizan. Se trata de darles veracidad y, derivado de esta, obtener información fiable.
	5. Veracidad	La veracidad trata de asegurar que los datos sean correctos, y por tanto, eficaces para su análisis y para la toma de decisiones. Una buena estrategia y una correcta planificación acercan al objetivo.
	6. Visualización	Tablas, hojas de cálculo, gráficas, flujos La manera en la que se muestran los datos también influye, y hace más fácil su comprensión.
	7. Valor	Seguro que ha oído aquella famosa frase que dice que "la información es poder". En el caso de la tecnología <i>big data</i> , podríamos decir que "la información es valor". Después de haber captado un sinfín de datos, haberlos almacenado, clasificado por tipología y analizado su veracidad, llega el momento aplicar los resultados a la estrategia
\		

## ¿Qué permite hacer Big Data?



Fuente: "Big Data Analytics for Policy Making Report", Comisión Europea, 2016

# Algunas Tecnologías para procesar Big Data



### Información en los datos

- Al contar con un importante volumen de datos, puede ser difícil extraer información.
  - Problema conocido como sobrecarga de información.



• Sin embargo, existen técnicas que permiten detectar patrones o tendencias, que pueden ser de gran utilidad.

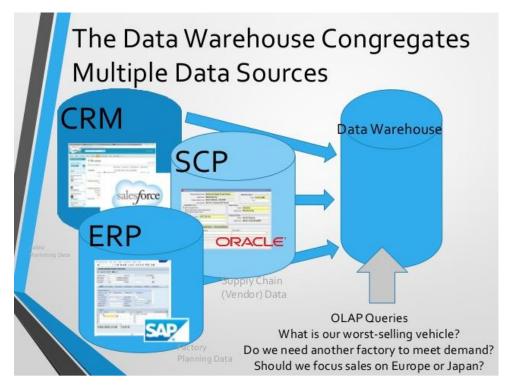


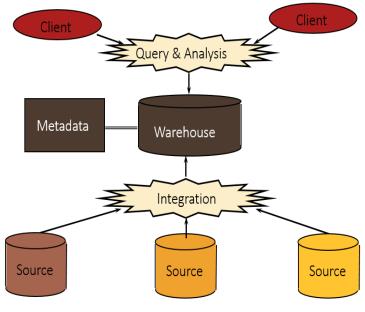
3. Conceptos BD y Minería de Datos

### 3. Bases de Datos y Minería de Datos

- BD es un conjunto de datos estructurado y relacionados entre sí
- Los datos son hechos conocidos que pueden registrarse y tienen un significado implícito.
- Ejemplos
  - Nombres
  - Números de teléfono
  - Direcciones
- Una base de datos se puede crear manualmente (ej: libreta de teléfonos) o mediante software (sistema de información y/o gestor de bases de datos).
  - La segunda opción permite una gestión y procesamiento mucho más rápido!

### Data Warehouse





Base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta.

### Minería de Datos

- Consiste en hallar tendencias o pautas interesantes en conjuntos de datos de gran tamaño para orientar las decisiones sobre actividades futuras
- Búsqueda de relaciones y patrones existentes en grandes bases de datos pero que se encuentran "ocultas" en **grandes volúmenes de datos**

(Frawley, Piatetsky-Shapiro & Matheus, 1992)

• Estas relaciones representan un conocimiento valioso acerca de la BD y los objetos que residen en ellas



## ¿Por qué usar Minería de Datos?

- Básicamente, porque las habilidades humanas no son suficientes para analizar grandes volúmenes de datos
- Hoy tenemos:
  - Crecimiento explosivo de datos
  - Medios de almacenamiento
  - Poder computacional
  - Automatización de la captura de datos

- Scanner en supermercados
- Tarjetas de clientes
- Transacciones de tarjetas de crédito
- Correo electrónico
- Registro de centros de llamados
- Datos demográficos
- Redes
- Cámaras
- Log de servidores de WEB
- etc.

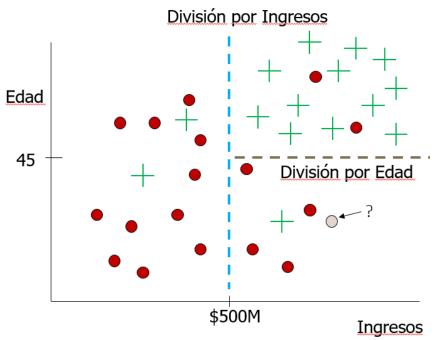
## Objetivo de la Minería de Datos

- La minería de datos ayuda en la extracción de patrones que no se pueden obtener directamente con consultas a la BD (SQL)
- La minería de datos puede ser usada en combinación con un DW para ayudar en la toma de decisiones de una organización

# Procesamiento de Bases de Datos v/s Minería de Datos

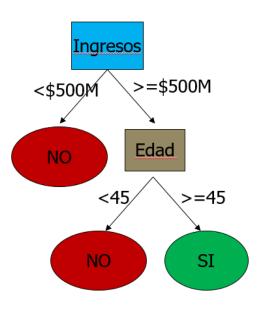
	Bases de Datos	Minería de Datos
Consultas	<ul> <li>Bien definidas</li> <li>Utiliza un lenguaje de consulta (SQL)</li> </ul>	<ul> <li>Pobremente definidas</li> <li>No necesita lenguaje de consulta</li> </ul>
Salida	<ul><li>Precisa</li><li>Un subconjunto de la BD</li></ul>	<ul><li>No precisa</li><li>No es un subconjunto de la BD</li></ul>

## Ejemplo: Segmentación de clientes



- No compró el producto
- Compró el producto

Árbol de Clasificación

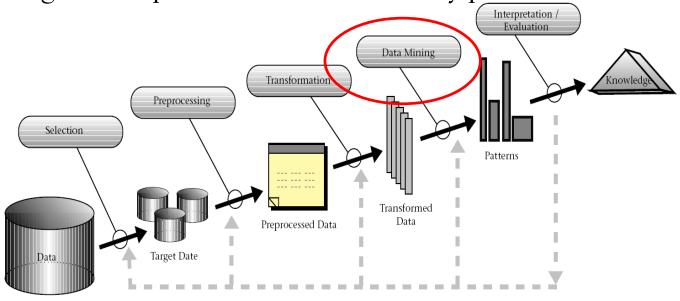


Interesado? NO

### El Proceso KDD

• Knowledge Discovery in Databases (KDD) es el **proceso** de encontrar información útil y patrones en los datos (proceso de descubrimiento de conocimiento)

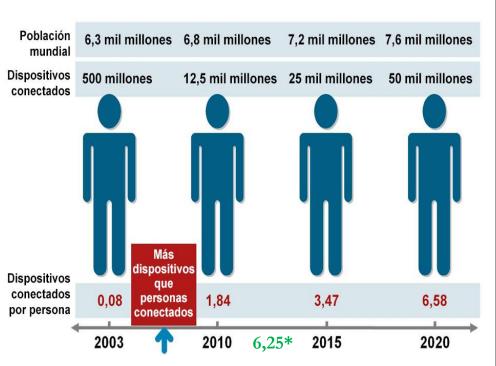
• La **minería de datos** es una fase central en el KDD y corresponde al uso de algoritmos para extraer información y patrones



4. Internet de las Cosas

## 4. Internet of Things (IoT)

- En 2009, un equipo de investigadores de China estudió los datos de routing de Internet en intervalos semestrales, desde diciembre de 2001 hasta diciembre de 2006.
- Los resultados mostraron que Internet duplica su tamaño cada 5,32 años.
- CISCO calculó el número de dispositivos conectados a Internet por persona en base a la población mundial (calculo no considera los avances en tecnología de Internet o dispositivos).



Nace The Internet of Things

Fuente: Cisco IBSG, abril de 2011; Forrester Research

<sup>\*</sup> en base a personas conectadas

# Ejemplo de objetos cotidianos conectados a Internet

- Refrigeradores inteligentes
  - Si cada alimento que compramos pudiese ser identificado (por ej. mediante dispositivos RFID) y nuestro refrigerador tuviese los sensores necesarios
  - Podría avisarnos cuando falte algún alimento
  - Podríamos revisar en línea lo que tenemos en el refrigerador, mientras compramos
  - O podría conectarse con un servicio de venta de alimentos online, y realizar automáticamente la compra!



## Entornos y aplicaciones inteligentes



Smart Planet
Green Environment

- Environmental sensors
- · Water, power leak detection
- · Pollution, weather monitoring



Smart Buildings Buildings, Smart Homes

- Thermostats, HVAC, lighting
- · Presence sensors, lockers, actuators
- Meters, smart-plugs, HEC



Smart Industry

#### Industrial Environments

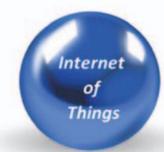
- Lightning, security, actuators
- · Production control
- Robotics



**Smart Cities** 

Connected Communities

- · Lighting, water management
- . Monitoring & security
- · Traffic control





Smart Health Healthcare System

- People monitoring
- · Bio sensors, probes
- · Remote health



Smart Energy Electric Grid

- Voltage and power sensors
- Meters and breakers
- Fault detection



Smart Transport

- Electric Mobility, EVs and HEVs
- High Speed Trains
- · Infrastructure, V2I, V2V, V2I+I



Smart Living Entertaining, Leisure

- · Independence through technology
- · Information when you need it
- · Connected when you need it

### **SMART BUILDING** SANTIAGO DE CHILE AIRPORT



Ensuring quality in the management and operation of airport infrastructures by monitoring environmental conditions



"We have made valuable information available for the airport managers to guarantee the comfort level

of the facilities and to increase passengers satisfaction in their passing through the airport", has arqued Alfonso Sardón, product manager of Ikusi.

### Sensors

Ambient conditions: humidity, temperature and luminosity.









The information is visualized in the platform Spider which is an integration software tool to control and optimize processes in airports. The platform centralizes, processes and exploits data in a multi-modal manner to enable transport planners, operators and authorities to visualize the data in a simplified, orderly and easy way.



### Benefits

- Establishing monitoring and auditing measures to make sure that the concessionaire company is carrying out its activity in accordance with the quality parameters agreeded.
- Consuming just the energy needed by maintaining luminosity, humidity and temperature within quality levels established by the airport operator.
- Passengers spending more time in shopping areas getting better commercial experience as they are more comfortable and relaxed.

### Libelium Insight

Global smart airports market is expected to reach \$13 billion by 2019 with an annual growth rate of 6%. Organizations are aware of the benefits of applying the latest technology to improve any kind of process in airports. It represents the idea of aerotropolis that integrates airports' features with metropolis functionalities





### **SMART ENVIRONMENT PERU RAINFOREST**



Monitoring weather and water conditions to control climate change in the National Park of Manú in Peru



"The sensors installed in this area will monitor the environment of the biosphere, both flora and fauna. There was a great need to know the different parameters because they reflect nature's behaviour. This scientific data will have a significant useful to the worldwide researcher and environmental communities",

### Sensors

- Ambient conditions: humidity, temperature, atmospheric pressure and weather station WS-300.
- Emission levels: carbon dioxide (CO<sub>2</sub>), nitrogen dioxide (NO<sub>2</sub>), oxygen (Oa), carbon monoxide (CO), methane (CH,), ozone (Oa) and air pollutants.







### Cloud

Information gathered by the sensors is visualized in an own Cloud platform from RFID Radical Solutions.

### Benefits

- Giving real-time and open information about the rainforest environment to current researchers.
- Combating against climate change.

### Libelium Insight

Libelium is proud to contribute to help scientists and researchers to figth against climate change as it is one of the priorities of the Corporate Social Responsibility of the company. Alicia Asín, CEO of Libelium, states that "this project is interesting to show the capability of Libelium Wireless Sensor Platform to provide tools to monitor changes in biological conditions".





affirms Jorge Camus, engineer of the project.







### Interacción



- ¿Qué tipo de empresas se podrían beneficiar de la tecnología IoT?
- ¿Qué otro tipo de solución se podría aplicar para aliviar el problema de Covid-19 utilizando IoT?



### Discover 10 Smart Tourism Solution Post-COVID:

- 1 » Temperature control for fever screening
- 2 » Indoor occupancy control
- 3 » Crowd and traffic tracking
- 4 » Parking space detection
- 5 » Air quality and pollution detection
- 6 » Fire prevention
- 7 » Weather conditions prediction and UV radiation measuring
- 8 » Water quality monitoring (pools and spas)
- 9 » Sea water quality control (beaches and off-shore)
- 10 » Water level measuring (streams and rivers)



### Libelium-Aridea Fever Screening Kit





An individual simply places their forehead five inches from the sensor and, using **infrared technology**, the microcontroller takes three separate temperature measurements and compares them to acceptable temperatures ranges.

A LED will turn red to indicate a fever or green to indicate no fever.





Screen 500 Employees for fever in under 30 minutes with only two kits!



### Monitor people and vehicle movements during alarm state

Meshlium Scanner is a tool capable of knowing the movements of people and vehicles in different parts of the city. In this way, it will be possible to evaluate compliance with the measures adopted for the COVID-19 crisis and implement improvements, if necessary. This system could serve as a deterrent strategy for uncivil behavior that does not respect the restrictions of confinement.

#### How it works √

The solution is composed of a mesh of smartphone detection scanners (Meshlium Scanner) that are installed in different points of the city. The scanners are capable of detecting WiFi and Bluetooth signals emitted by mobile devices of people and cars.



5. Otras tecnologías emergentes?

### Otras tecnologías emergentes: Caso de Estudio 4

- Cloud Computing
- Bolckchain
- Vision Artificial
- Impresoras 3D
- Realidad Aumentada
- Análisis de Sentimiento
- Inteligencia Artificial
- Industria 4.0
- etc, etc.