

# Gestión Óptima de Consejerías, Servicios y Proyectos desde DGAD

## Día 4

Nueva plataforma de IA.  
Teoría, el proyecto de la plataforma  
y Estrategia

David Rodríguez  
davidr@jccm.es

## ¿Qué vamos a ver hoy?

- Introducción a la IA.

- Sistema de IA prototipo.

- Características sistemas IA

- La Plataforma IA

- Servicio IAA

■ ¿Qué actividad te resulta mas compleja?



## ■ Robotic grasping.

- <https://paperswithcode.com/task/robotic-grasping>

## ■ Ajedrez

- Deep Blue (1997): <https://www.catalunyavanguardista.com/kasparov-vs-deep-blue-la-maquina-vence-al-hombre/>

## ■ Go

- Alpha Go (2016): [https://es.wikipedia.org/wiki/AlphaGo\\_versus\\_Lee\\_Sedol](https://es.wikipedia.org/wiki/AlphaGo_versus_Lee_Sedol)
- Documental: <https://www.youtube.com/watch?v=WXuK6gekU1Y>

## Ejemplo sencillo: clasificación de imágenes

### ■ ¿Perro o gato?



➤ Demo: <https://huggingface.co/spaces/bencoman/cats-v-dogs>

### ■ Pero...

- ¿Sabe el sistema lo que es un perro y lo que es un gato?
- ¿Sabe que un perro ladra?
- ¿Conoce que los gatos cazan ratones?

# Introducción a la IA

## IA Fuerte vs IA Débil

### ■ IA fuerte o general (IAG).

- Intenta emular comportamiento y procesos de pensamiento humano.
- Orientación tomada inicialmente. Resultados decepcionantes.

### ■ IA Débil.

- Se centra en una única tarea. No tiene todo el contexto de la realidad.
- Ejemplo. Clasificación de perros y gatos

### ■ Sistemas actuales: IA Débil.

- Obtienen mejores resultados y resuelven problemas reales.

■ RGPD garantiza el derecho a no ser sometido a decisiones automatizadas, incluyendo la elaboración de perfiles, cuando:

- No hay intervención humana.
- Produzcan efectos jurídicos
- O afecte de forma similar y significativa al interesado

■ Uso IA:

- Evitar utilización para toma y ejecución de decisiones automatizada.
- Herramienta de soporte/ayuda al proceso de decisión, proporcionando una inferencia o perfil sobre un sujeto o una situación, para que un ser humano tome la decisión final.

## Clasificación IA

### Inteligencia Artificial

#### Aprendizaje Automático (Machine Learning)

#### Aprendizaje Profundo (Deep Learning)

Habilidad de las máquinas para imitar el comportamiento humano.  
Resolución de tareas complejas (percepción visual, reconocimiento de patrones, toma de decisiones, ...)

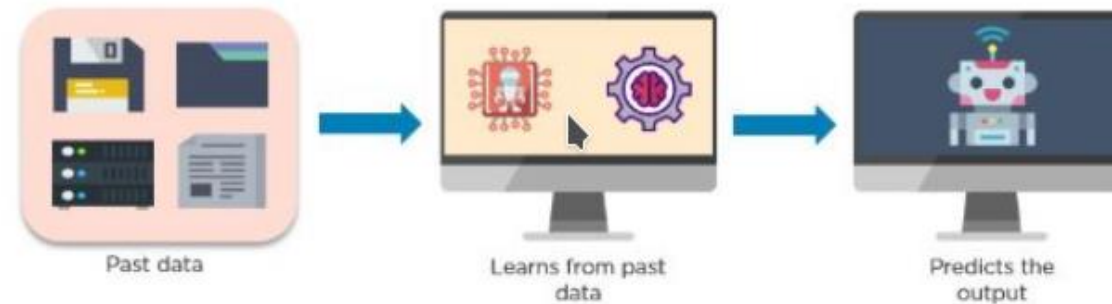
El sistema aprende y mejora a partir de la experiencia (los datos)

Utilización de modelos matemáticos basados en redes neuronales complejas y estructuradas en múltiples capas.



Sistema IA prototipo. Aprendizaje Automático (Machine Learning – ML)

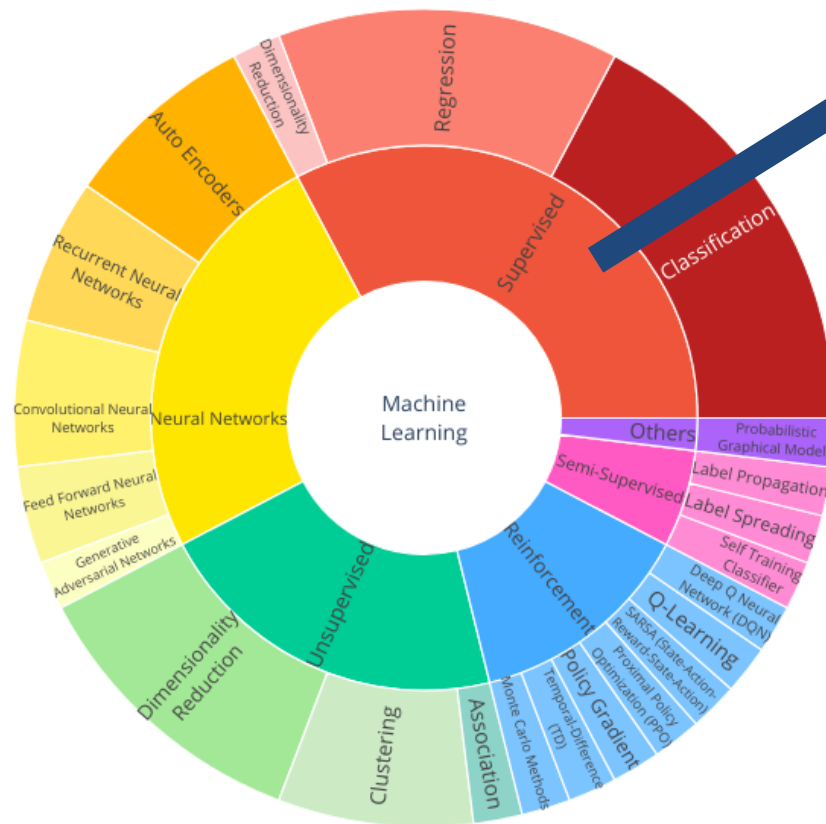
- Algoritmos que **aprenden** de los datos sin necesidad de programación de reglas.



- ¿Qué significa aprender?

➤ Ver <https://playground.tensorflow.org>

## Sistema de IA prototipo: Aprendizaje Automático Supervisado.



### ■ Aprendizaje supervisado

- Paradigma más utilizado en la actualidad.
- Normalmente cuando se habla de IA o de Aprendizaje Automático se trata de Aprendizaje supervisado.

Ver: <https://chart-studio.plotly.com/create/?fid=SolClover:53>

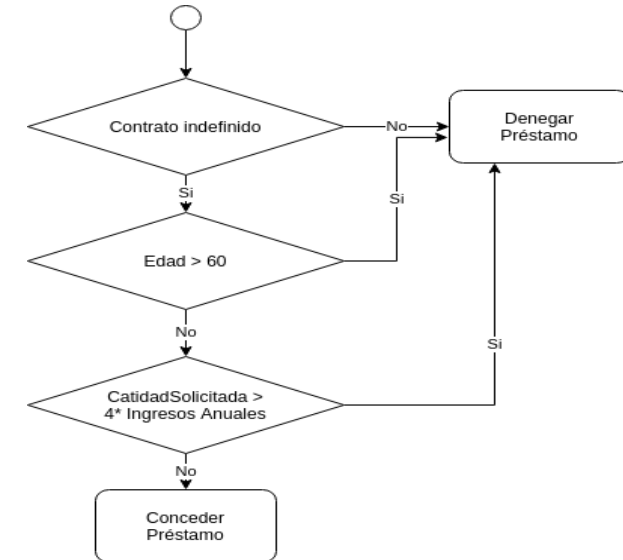
## Programación tradicional.



## Programación tradicional. Ejemplo de concesión de préstamo

### ■ Entrada: Información sobre el solicitante:

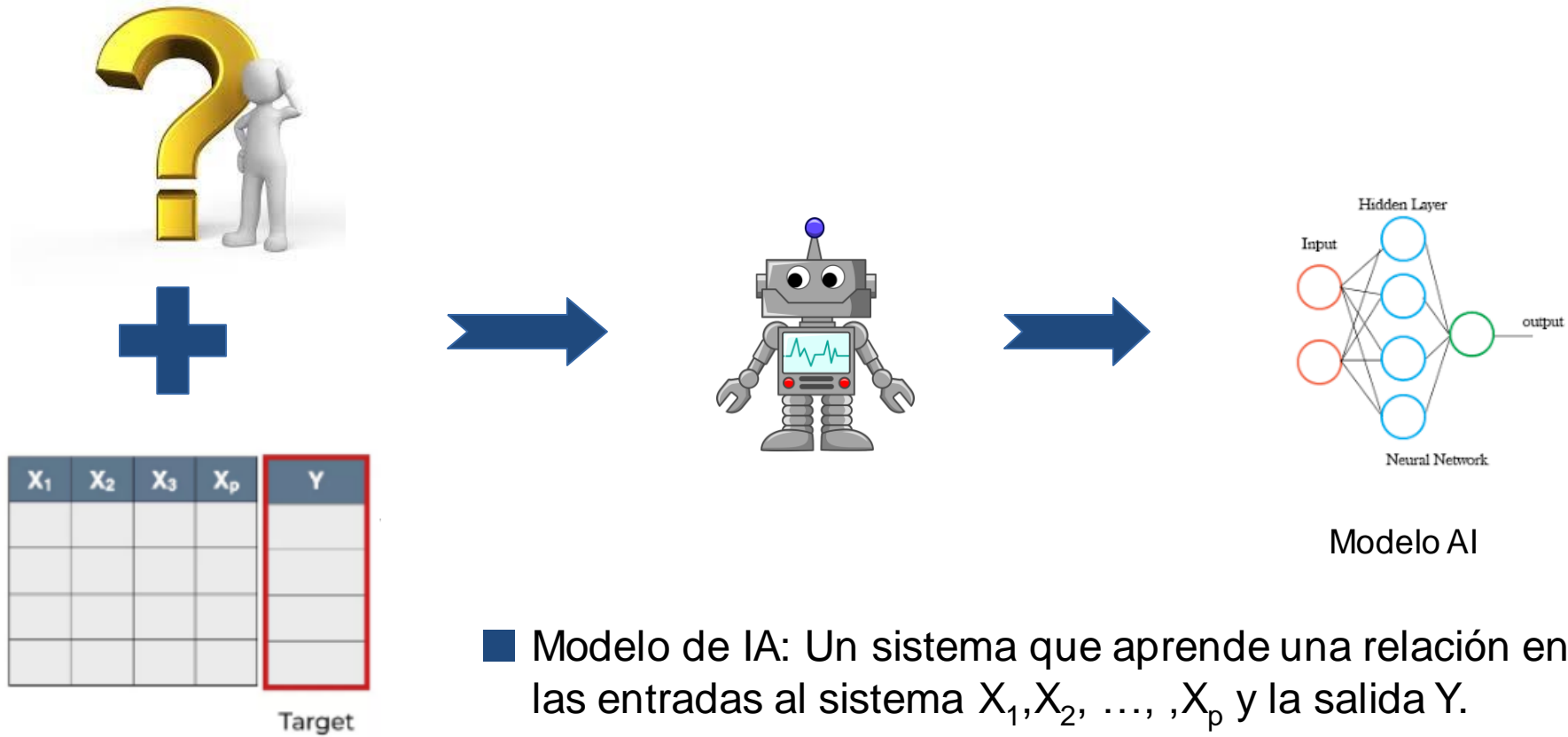
- $x_1$  Tipo de contrato
- $X_2$  Edad solicitante
- $X_3$  Ingresos anuales
- ...
- $X_p$  Cantidad solicitada



### ■ Resultado Deseado:

- Concesión o denegación del préstamo.

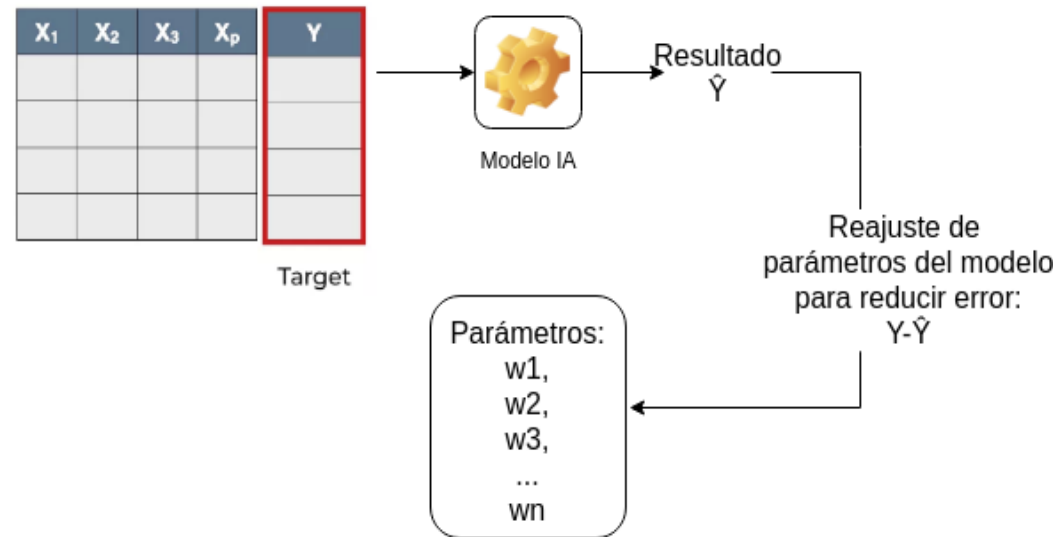
## Aprendizaje Supervisado (Basado en ejemplos)



- Modelo de IA: Un sistema que aprende una relación entre las entradas al sistema  $X_1, X_2, \dots, X_p$  y la salida  $Y$ .

## Aprendizaje Supervisado (Basado en ejemplos)

Fase de  
entrenamiento

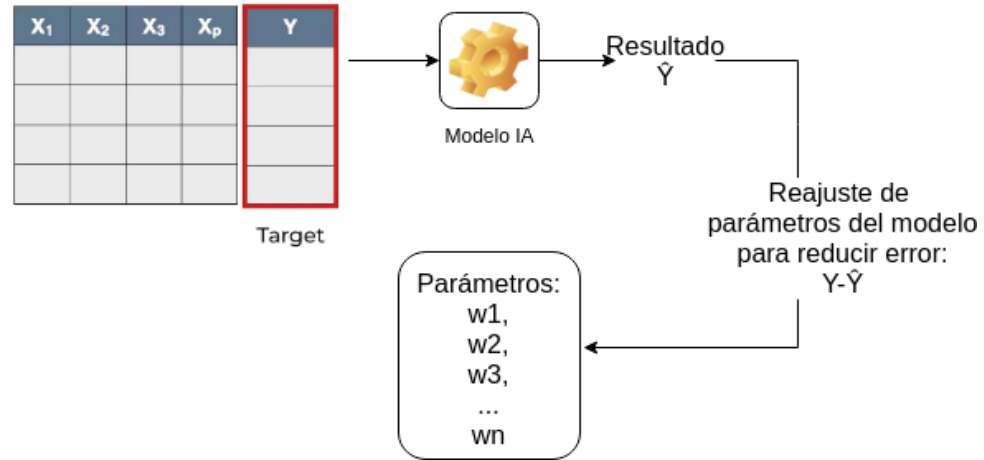


Ejemplo  
Datos Entrenamiento:

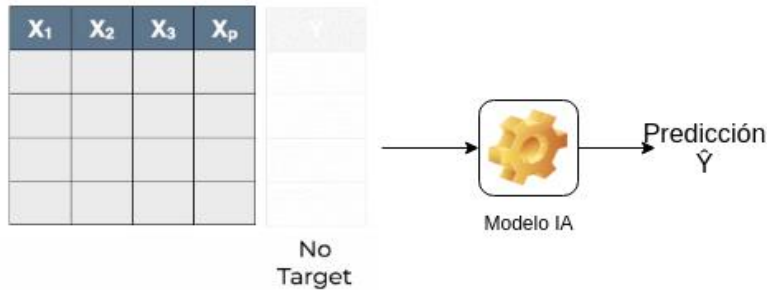
Tipo Contrato	Edad	Ingresos	Cantidad	Concedido S/N
Indefinido	41	30000	40000	S
Temporal	32	35000	28000	N
Indefinido	64	60000	45000	N
Indefinido	28	35000	90000	S

## Aprendizaje Supervisado (Basado en ejemplos)

Fase de  
entrenamiento



Fase de  
Utilización del  
Modelo  
(Inferencia)



➤ Ver: <https://www.aifunded.es/demos/demo.php>

## Aprendizaje Supervisado (Basado en ejemplos)

- ¿Cómo se mide la calidad del modelo resultante?
- ¿Cómo se sabe en qué momento hay que detener el entrenamiento?

Tipo Contrato	Edad	Ingresos	Cantidad	Concedido S/N	
Indefinido	41	30000	40000	S	Datos de Entrenamiento (70 %)
Temporal	32	35000	28000	N	
Indefinido	64	60000	45000	N	
Indefinido	28	35000	90000	S	
...	...	...	...	...	Datos de Test (30%)
...	...	...	...	...	
Temporal	36	25000	5000	S	

Durante el entrenamiento se reserva un conjunto de datos (datos de test) para obtener métricas.



## ■ Requiere gran cantidad de datos.

- Datos etiquetados para training y testing. En algunos casos muy costoso.

## ■ Datos no agregados.

- Datos históricos en bruto, con alto nivel de detalle.

## ■ Los resultados dependen de la calidad de los datos disponibles:

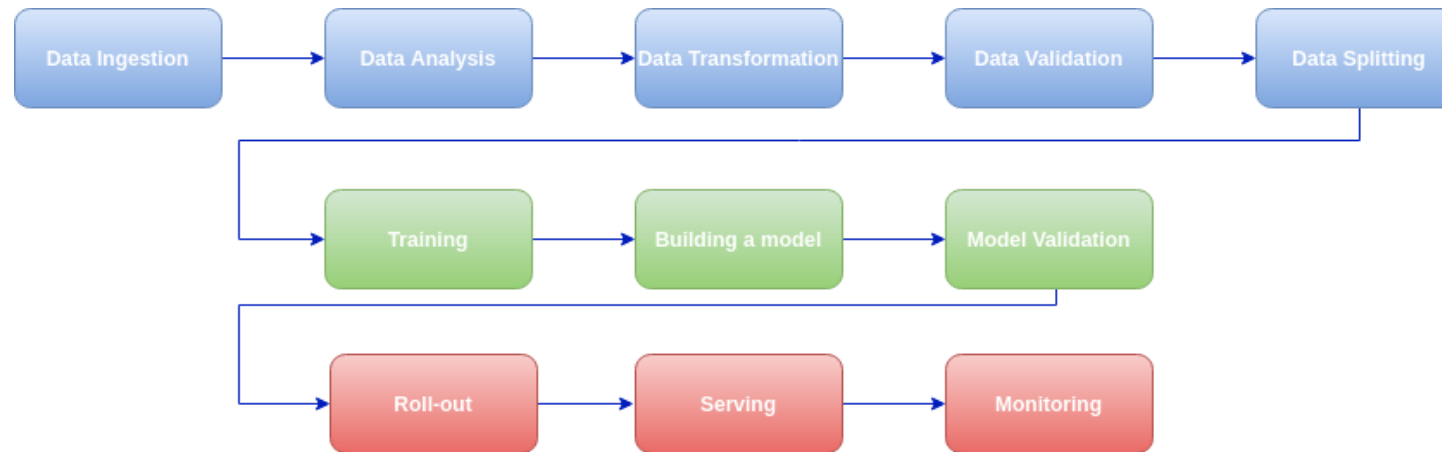
- “Garbage in, garbage out”

## Organizaciones Data Driven

- De datos aislados a gobierno y explotación de los datos.
  - Big Data
- Los datos se convierten en uno de los activos mas importantes de la organización.
- Toma de decisiones basada en datos
  - En lugar de observaciones, opiniones personales, instintos, ...
- Contamos con las herramientas necesarias para extraer valor de estos datos
  - Pero no se trata solo de tecnología, en gran medida es un cambio cultural

## ■ Premisas totalmente diferentes a los sistemas informáticos tradicionales.

- Proceso no determinista. Requiere de una fase de experimentación y optimización.
- Suele utilizar tecnología poco implantada en la organización: Python, R y sus librerías especializadas en ML.
- Los resultados tienen un nivel de precisión/acierto que no es el 100%



## ■ Fases diferentes a las empleadas en el desarrollo de Software "tradicional"

- Datos (Preparación, análisis, tratamiento).
- Entrenamiento.
  - Alto coste computacional. HW especializado (GPUs)
  - No tiene sentido entrenar con datos de desarrollo o preproducción.
- Evaluación del modelo.
- Empaquetado
- Entrega (Despliegue)
- Supervisión (Monitorización)

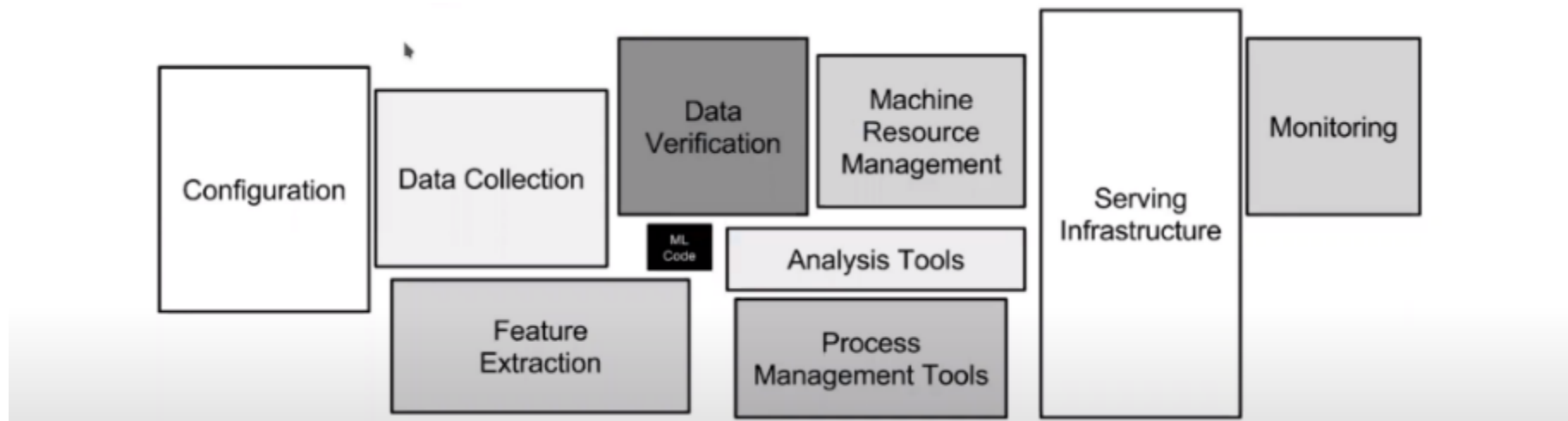
# Características sistemas IA

## MLOps

- Extensión de la metodología DevOps aplicada a gestionar la complejidad de los activos involucrados en el aprendizaje automático.
  - Permite adaptarse antes a los cambios
  - Gestionar ágilmente el elevado número de versiones de productos generados.
- Busca aspectos como:
  - Automatización de procesos de aprendizaje automático.
  - Versionado de datos, modelos y código.
  - Testing automático de los componentes del sistema.
  - Monitorización.
  - Reproductividad de las diferentes fases

## ■ En resumen

- Sistema complejo
- Muchas piezas interconectadas.
- Requerimientos diferentes a los sistemas tradicionales.



- Surgen herramientas específicas destinadas a:
  - Orquestar todos los activos involucrados en la puesta en funcionamiento de este tipo de sistemas.
  - Proveer facilidades que permitan acelerar la construcción de modelos.
  - Homogeneizar el proceso de construcción de estos modelos.
  - Aplicar el paradigma MLOps.
  
- El uso de este tipo de herramientas no es estrictamente imprescindible, pero si muy aconsejable para tener control sobre todo el proceso.

# Plataforma IA

Algunas características requeridas.

- Solución modular que permita ejecución sobre Openshift.

- "On premise" (contrato Sistemas).

- En "Cloud".

- Escalabilidad tanto vertical como horizontalmente.

- Debe dar cabida a todos los procesos involucrados en el ciclo de vida de los sistemas IA (solución extremo a extremo):

- Acceso a las fuentes originales de datos, su limpieza, transformación, catalogación, etc.

- Fase de experimentación IA, entrenamiento y optimización de modelos.

- Fase de productivización y monitorización de los modelos en producción.



La solución

## IBM Cloud Pak for Data

Magic Cuadrant for Data Quality Solutions



Figure 1: Magic Quadrant for Data Science and Machine Learning Platforms



Source: Gartner (March 2021)

La solución

## IBM Cloud Pak for Data

### ■ Documentación

➤ Knowledge Center:

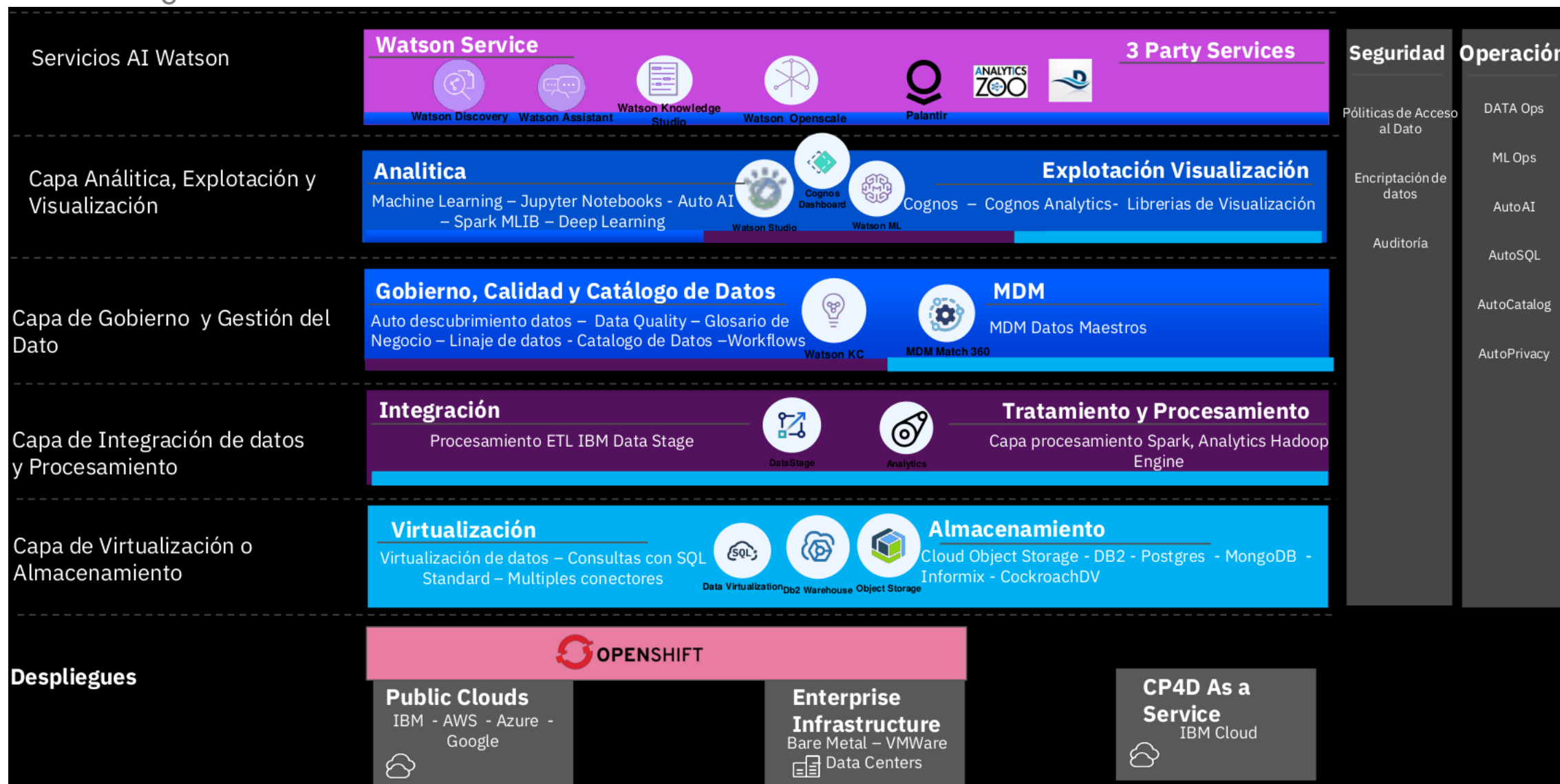
<https://www.ibm.com/support/knowledgecenter/en/SSQNUZ>

➤ Demos y videos: <https://www.ibm.com/demos/collection/Cloud-Pak-for-Data/?lc=en>

➤ Community:

<https://community.ibm.com/community/user/cloudpakfordata/home>

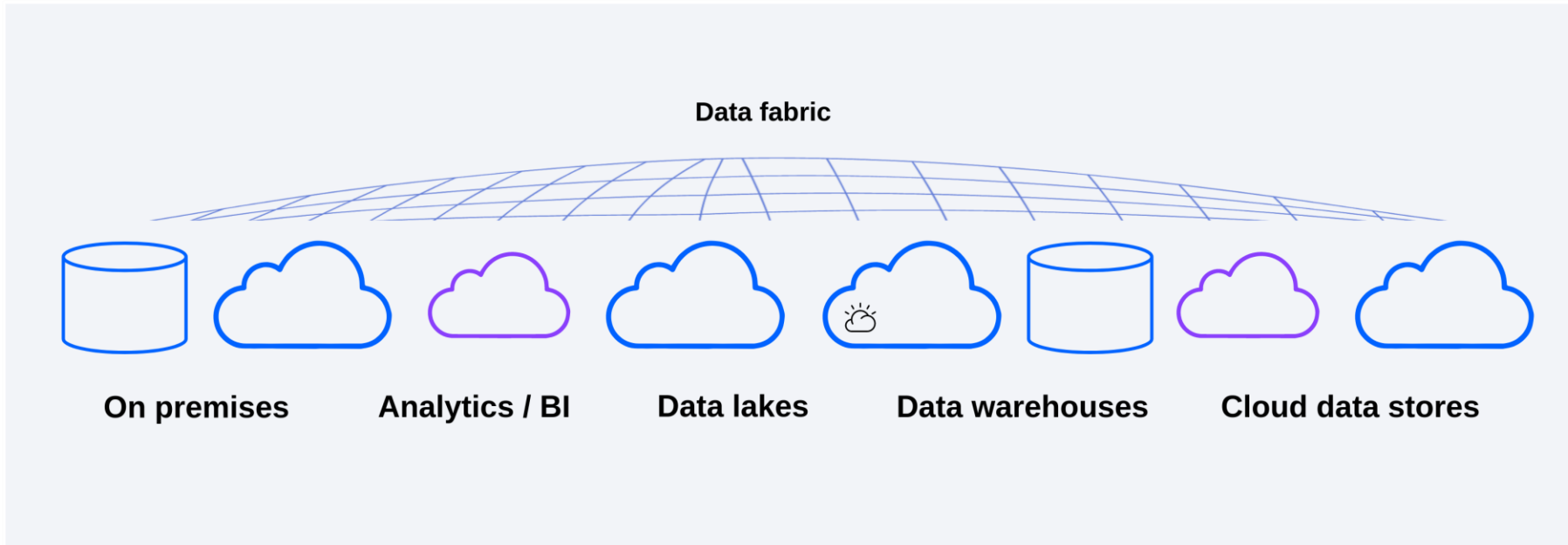
## ¿Qué ofrece?



¿Qué ofrece? En el ámbito del acceso a datos

## ■ Acceso a los datos. Data Fabric

- Acceso seguro y consistente a gran variedad de fuentes de datos.
- AutoServicio: Compartir y encontrar activos de datos agrupándolos en catálogos.



¿Qué ofrece? En el ámbito del acceso a datos

## ■ Data Fabric

- Consultas combinando diferentes orígenes de datos: Oracle, SQLServer, Db2, MySQL, PostgreSQL, ...
- Generación de vistas de negocio virtualizadas. Buscar y combinar tablas automáticamente entre sistemas para poder consultarlas como una sola tabla virtual.
- Capacidad de conexión desde aplicaciones de terceros.
- Seguridad. Estrictos controles de acceso y comunicaciones totalmente encriptadas.

¿Qué ofrece? En el ámbito del acceso a datos

## ■ Catálogo y Marketplace de Datos

- Marketplace de datos, para compartir los datos con otros usuarios y departamentos, accediendo a rates y comentarios de otros usuarios.
- Permite publicar todos los datos y activos, curados, con calidad, securizados y encriptados.



¿Qué ofrece? En el ámbito del acceso a datos

## ■ Calidad de Datos

- Descubrimiento automático de la calidad de los datos, ofreciendo una primera visión de la confiabilidad de los mismos.
- Histórico de evolución de la calidad.
- Permite definir reglas personalizadas de calidad. Ejemplo, convertir formatos de fechas automáticamente.

Dashboard

Data assets

Data rules

Relationships

Settings

Data assets

6 data assets

Discover

Create SQL virtual table

Find a data asset

Sort by:

Default

Filter by:

All data assets

BANK\_CLIENTS

Analyzed

Data quality:

97%

Delta:

—

Location

jdbc:db2://toulon1.fyre.ib...

Last analysis

Oct 30, 2020, 5:53 AM

Last published

Never

Last imported

Oct 29, 2020, 9:33 AM

Threshold

80%

BANK\_ACCOUNTS

Analyzed

Data quality:

97%

Delta:

—

Location

jdbc:db2://toulon1.fyre.ib...

Last analysis

Oct 30, 2020, 5:50 AM

Last published

Never

Last imported

Oct 29, 2020, 9:33 AM

Threshold

80%

CHECKING\_ACCOUNTS

Analyzed

Data quality:

99%

Delta:

—

Location

jdbc:db2://toulon1.fyre.ib...

Last analysis

Oct 30, 2020, 5:50 AM

Last published

Never

Last imported

Oct 29, 2020, 9:33 AM

Threshold

80%

BANK\_CUSTOMERS

Analyzed

Data quality:

98%

Delta:

—

ACCOUNT\_HOLDERS

Analyzed

Data quality:

96%

Delta:

—

SAVINGS\_ACCOUNTS

Analyzed

Data quality:

98%

Delta:

—

¿Qué ofrece? En el ámbito del acceso a datos

## ■ Glosario de negocio. Se definen

- Términos como: conceptos de negocio, propiedades del concepto de negocio, relaciones entre activos, términos que definen las medidas y KPIs de la empresa
- Otras estructuras como: relaciones personalizadas entre activos, Cardinalidad, Tipos de datos o datos de referencia.

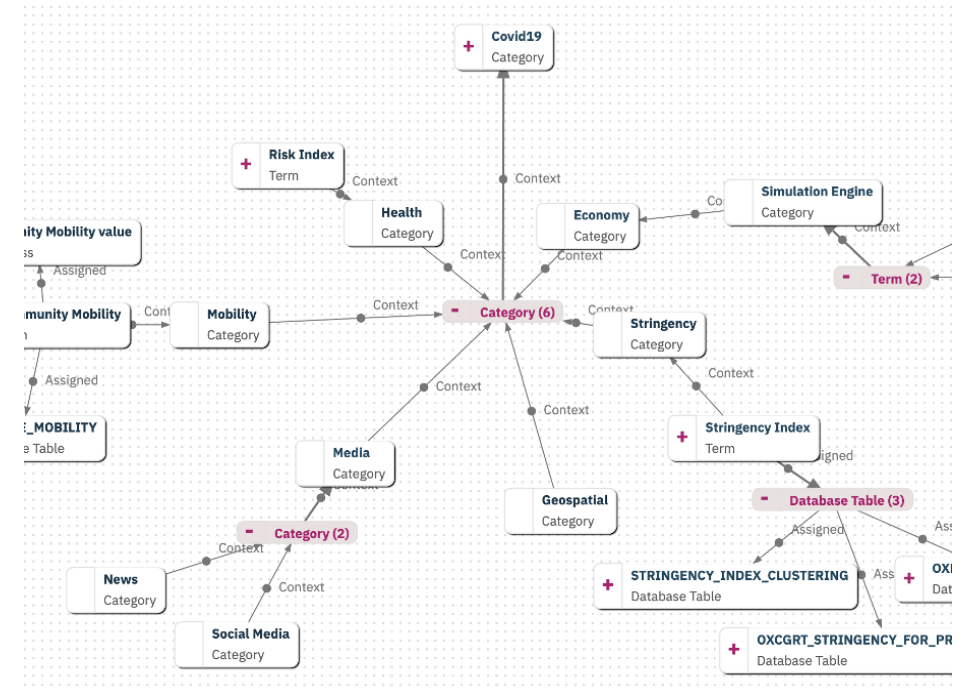
IBM Cloud Pak for Data

Search results for Social Security number

Any type Any tag Steward/Owner Modified by Clear all

Showing 25 of 183 items

Name	Type	Tags	Modified by	Modified on
<b>Social Security Number</b> Categories > Sensitive Information The social security number entered o...	Business te...	Mortgage Ap... Sensitive Inf...	SD Scott D'Angelo	May 26, 2020
<b>US Social Security Number</b> Categories > [uncategorized] In the United States, a Social Secur...	Data class		System	Apr 15, 2020
<b>US Social Security Number Last 4</b> Categories > [uncategorized] The last four digits of a United States ...	Data class		System	Apr 15, 2020
<b>Social security number</b> Categories > CCPA-v2019 Personal information includes, but is ...	Business te...		DW David Wilde	May 13, 2020

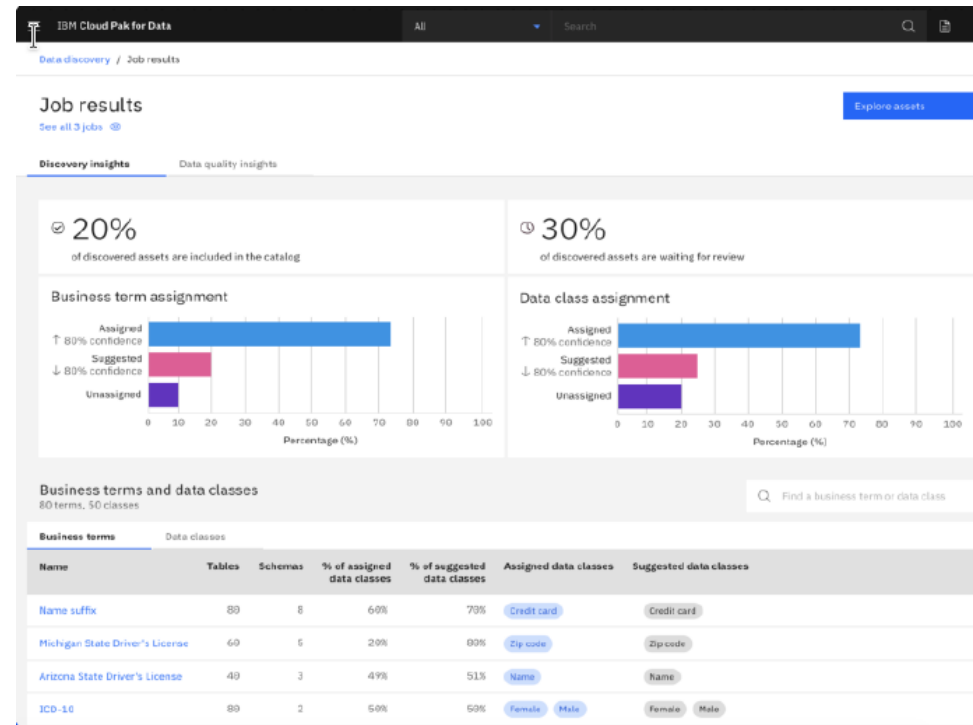




¿Qué ofrece? En el ámbito del acceso a datos

## ■ Auto descubrimiento de activos

- Autodescubrimiento de los datos de la empresa a través de procesos de AI: Detectan los metadatos, las clases de datos, los tipos de datos, etc...
- Asignación automática de activos de Gobierno a los datos descubiertos.



¿Qué ofrece? En el ámbito del acceso a datos

## ■ Linaje de datos y de gobierno

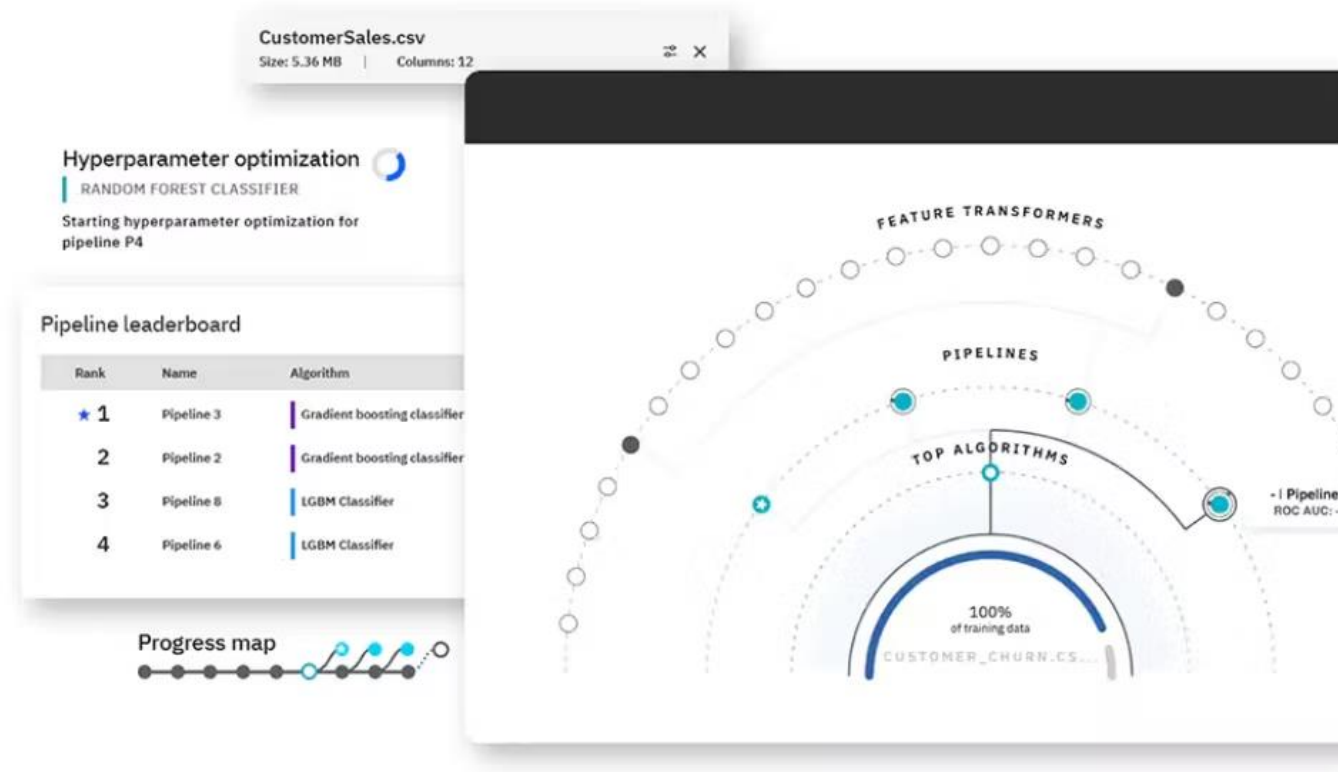
- Linaje del dato, nos muestra gráficamente las fuentes origen de los datos y qué procesos los han generado y transformado.
- Linaje de gobierno, nos permite ver la relación que existe entre los diferentes activos de gobierno, creando knowledge graphs y enriqueciendo la información de negocio del activo.



¿Qué ofrece? En el ámbito de ML

## ■ Auto-AI

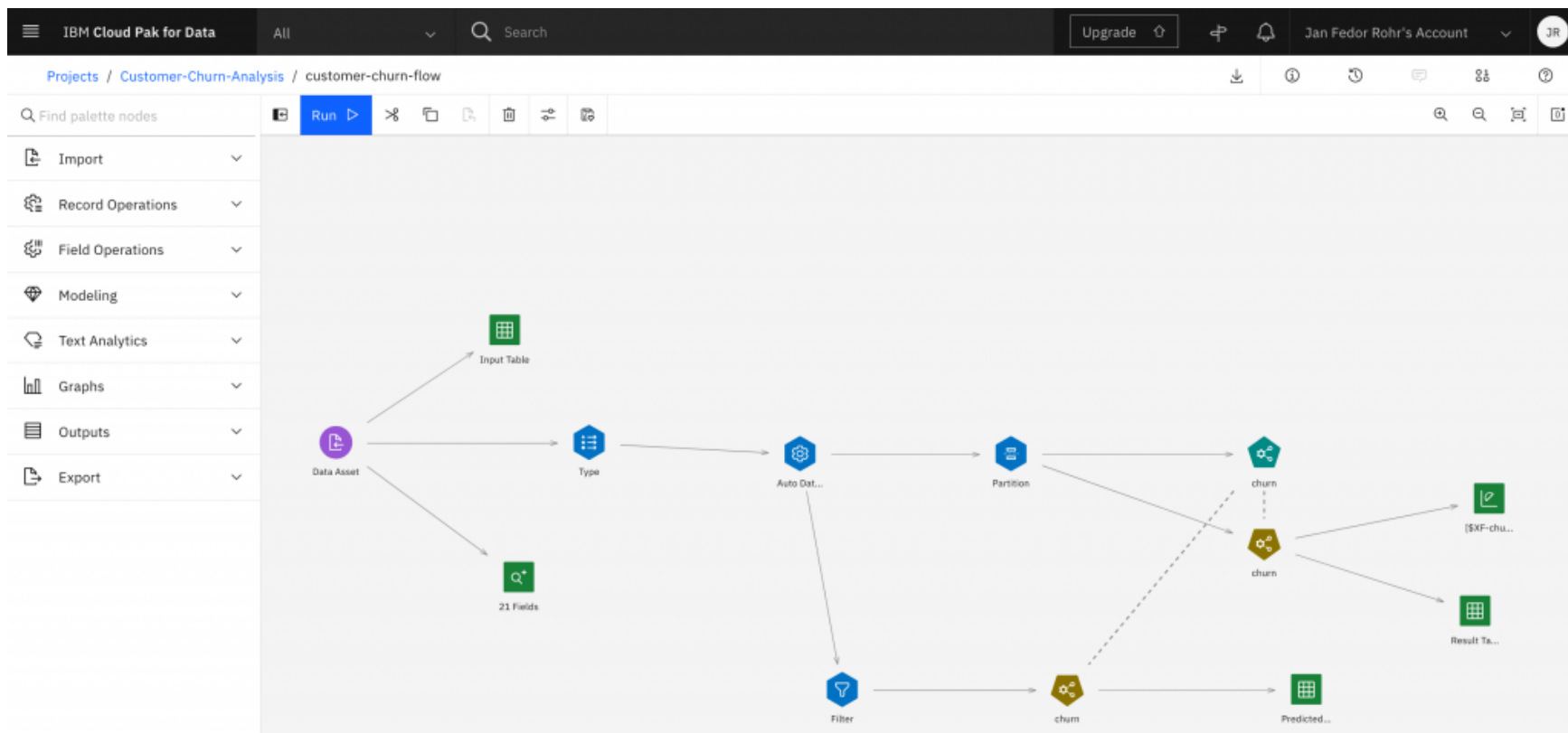
- Generación automática de modelos de ML y comparativa entre ellos.



¿Qué ofrece? En el ámbito de ML

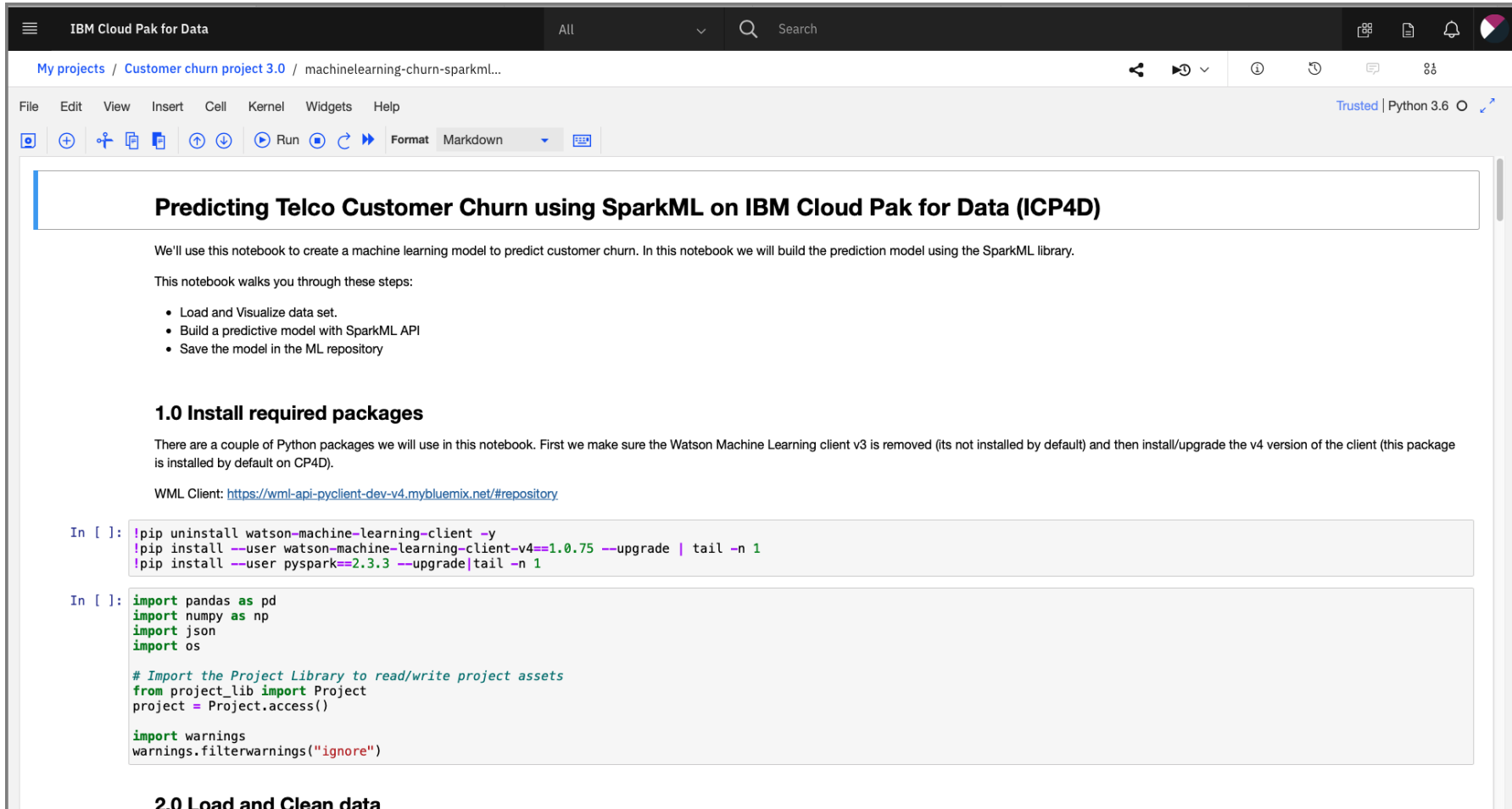
## ■ Generación gráfica de flujos de ML

➤ Sin código



¿Qué ofrece? En el ámbito de ML

## ■ Posibilidad de generar modelos mediante código Python o R



The screenshot shows the IBM Cloud Pak for Data (ICP4D) Jupyter Notebook interface. The notebook is titled "Predicting Telco Customer Churn using SparkML on IBM Cloud Pak for Data (ICP4D)". The interface includes a top navigation bar with "IBM Cloud Pak for Data", a search bar, and various icons. Below the navigation bar is a menu bar with "File", "Edit", "View", "Insert", "Cell", "Kernel", "Widgets", and "Help". The notebook content is displayed in a Markdown format, showing the title, an introduction, a list of steps, and code snippets for installing packages and importing libraries.

**Predicting Telco Customer Churn using SparkML on IBM Cloud Pak for Data (ICP4D)**

We'll use this notebook to create a machine learning model to predict customer churn. In this notebook we will build the prediction model using the SparkML library.

This notebook walks you through these steps:

- Load and Visualize data set.
- Build a predictive model with SparkML API
- Save the model in the ML repository

**1.0 Install required packages**

There are a couple of Python packages we will use in this notebook. First we make sure the Watson Machine Learning client v3 is removed (its not installed by default) and then install/upgrade the v4 version of the client (this package is installed by default on CP4D).

WML Client: <https://wml-api-pyclient-dev-v4.mybluemix.net/#repository>

```
In [ ]: !pip uninstall watson-machine-learning-client -y
!pip install --user watson-machine-learning-client-v4==1.0.75 --upgrade | tail -n 1
!pip install --user pyspark==2.3.3 --upgrade | tail -n 1
```

```
In [ ]: import pandas as pd
import numpy as np
import json
import os

# Import the Project Library to read/write project assets
from project_lib import Project
project = Project.access()

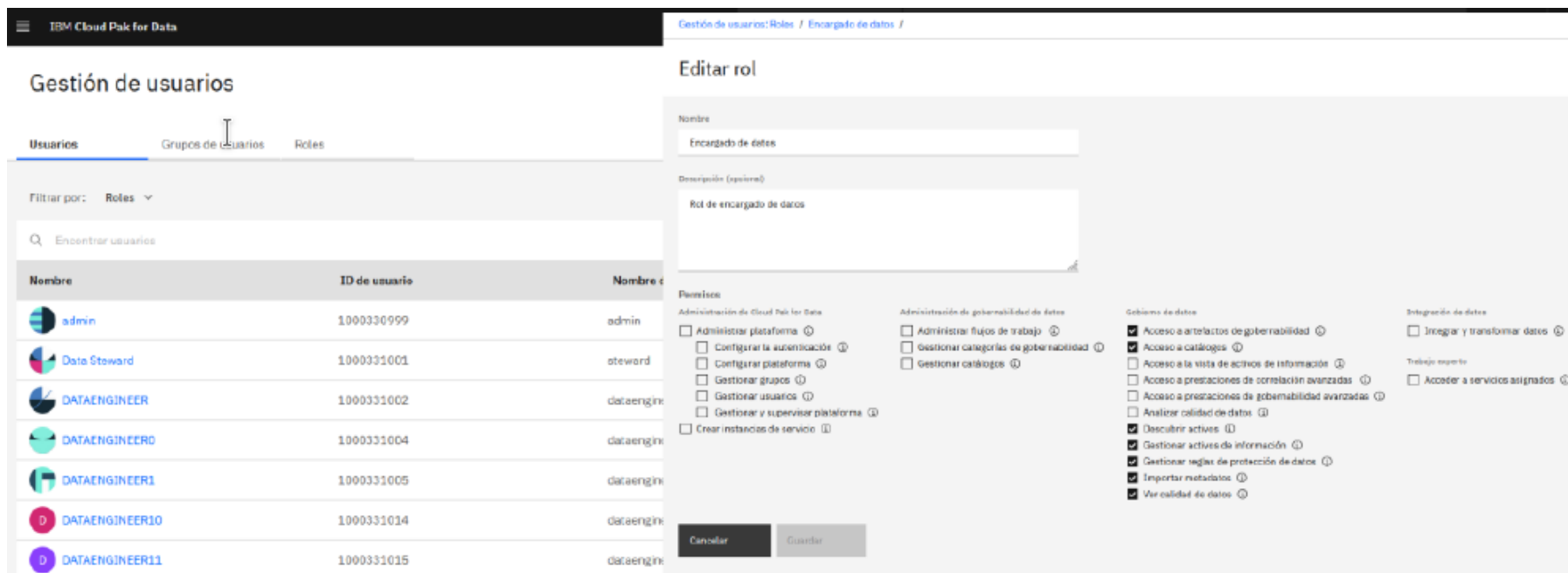
import warnings
warnings.filterwarnings("ignore")
```

**2.0 Load and Clean data**

¿Qué ofrece? En el ámbito de la seguridad

## ■ Gestión de usuarios y roles

- Generación y asignación de Roles y permisos a los usuarios de la plataforma según sus perfiles
- Acceso a datos de grano fino. Permiso de acceso a los datos de grano fino, a nivel usuario, o bien a través de reglas de negocio (RuBAC).



**Gestión de usuarios**

Usuarios | Grupos de usuarios | Roles

Filtrar por: Roles

Encuentra usuarios

Nombre	ID de usuario	Nombre de rol
admin	1000330999	admin
Data Steward	1000331001	oteward
DATAENGINEER	1000331002	dataengin
DATAENGINEER0	1000331004	dataengin
DATAENGINEER1	1000331005	dataengin
DATAENGINEER10	1000331014	dataengin
DATAENGINEER11	1000331015	dataengin

**Editar rol**

Nombre: Encargado de datos

Descripción (opcional): Rol de encargado de datos

**Permisos**

- Administración de Cloud Pak for Data
  - ☐ Administrar plataforma
  - ☐ Configurar la autenticación
  - ☐ Configurar plataforma
  - ☐ Gestionar grupos
  - ☐ Gestionar usuarios
  - ☐ Gestionar y supervisar plataforma
  - ☐ Crear instancias de servicio
- Administración de gobernanza de datos
  - ☐ Administrar flujos de trabajo
  - ☐ Gestionar categorías de gobernanza
  - ☐ Gestionar catálogos
- Gobierno de datos
  - ☒ Acceso a artefactos de gobernanza
  - ☒ Acceso a catálogos
  - ☐ Acceso a la vista de activos de información
  - ☐ Acceso a prestaciones de correlación avanzadas
  - ☐ Acceso a prestaciones de gobernanza avanzadas
  - ☐ Analizar calidad de datos
  - ☒ Descubrir activos
  - ☒ Gestionar activos de información
  - ☒ Gestionar reglas de protección de datos
  - ☒ Importar metadatos
  - ☒ Ver calidad de datos
- Integración de datos
  - ☐ Integrar y transformar datos
  - ☐ Trabajar experto
  - ☐ Acceder a servicios asignados

Cancelar | Guardar

¿Qué ofrece? En el ámbito de la seguridad

## ■ Anonimización y acceso a datos automático

- Aplicación de reglas de acceso o protección de datos automáticamente en el descubrimiento de los datos, sin necesidad de tener que configurar manualmente la seguridad y el acceso a los datos de cada activo.

### IBM Cloud Pak for Data

#### New data protection rule

Rule details

Name

Redact Social Security Number

Type

Access

Business definition ⓘ

Rule to redact Social Security Number

Rule builder

Criteria

Condition 1

Remove

If

Business term

Social Security Number X

contains any

+

Add new condition +

Action

Then

mask data

Social Security Number X

in columns

Business term

Select how to mask data:

Redact

Before

XXXXXXXXXX

Replace data with Xs

Substitute

Before

452-821-1120

Replace data with values that don't match the original format

Obfuscate

Before

452-821-1120

Replace data with similarly formatted values

Cancel

Create

¿Qué ofrece? En el ámbito de la productivización

## ■ Productivización

- Posibilidad de publicar el modelo de AI mediante.
  - API REST
  - En modo Batch

### Create a deployment

Associated asset  
autoai - P1 DecisionTreeClassifierEstimator

Deployment type

**Online** ✔

Run the model on data in real-time, as data is received by a web service.

**Batch**

Run the model against data as a batch process.

Name

flight-delays-model

flight-delays-model ✔ Deployed Online

API reference

Test

Direct link

Endpoint

`https://zen-cpd-zen.apps.cpd2.power.10.3.75.50.nip.io/ml/v4/deployments/048f8f92-1e5b-4a57-bfb4-a01cb1ea3987/predictions?version=2021-04-12`

Code snippets

cURL

Java

JavaScript

Python

Scala

# TODO: manually define and pass values to be scored below:  
`curl -X POST --header 'Content-Type: application/json' --header 'Accept: application/json' --header 'Authorization: Bearer $IAM_AUTH_TOKEN' -d '{"input_data": [{"fields": [$ARRAY_OF_INPUT_FIELDS], "values": [$ARRAY_OF_VALUES_TO_BE_`



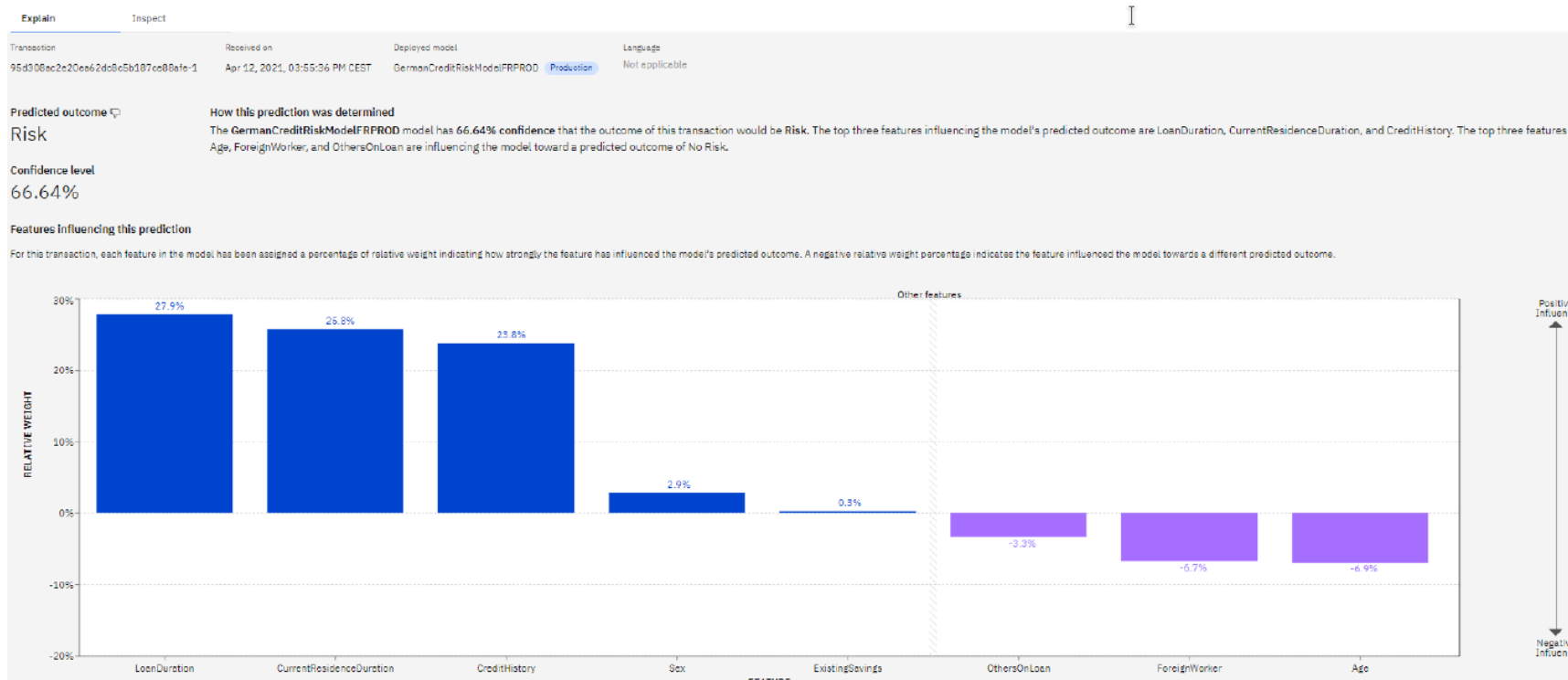
¿Qué ofrece? En el ámbito de la productivización

## ■ Explicabilidad de las decisiones

- Para una transacción concreta explica cómo se ha determinado la predicción y cuáles son los factores de mayor impacto.

Explanations / 95d308ac2820ea62dc8e5b187ce88afe-1

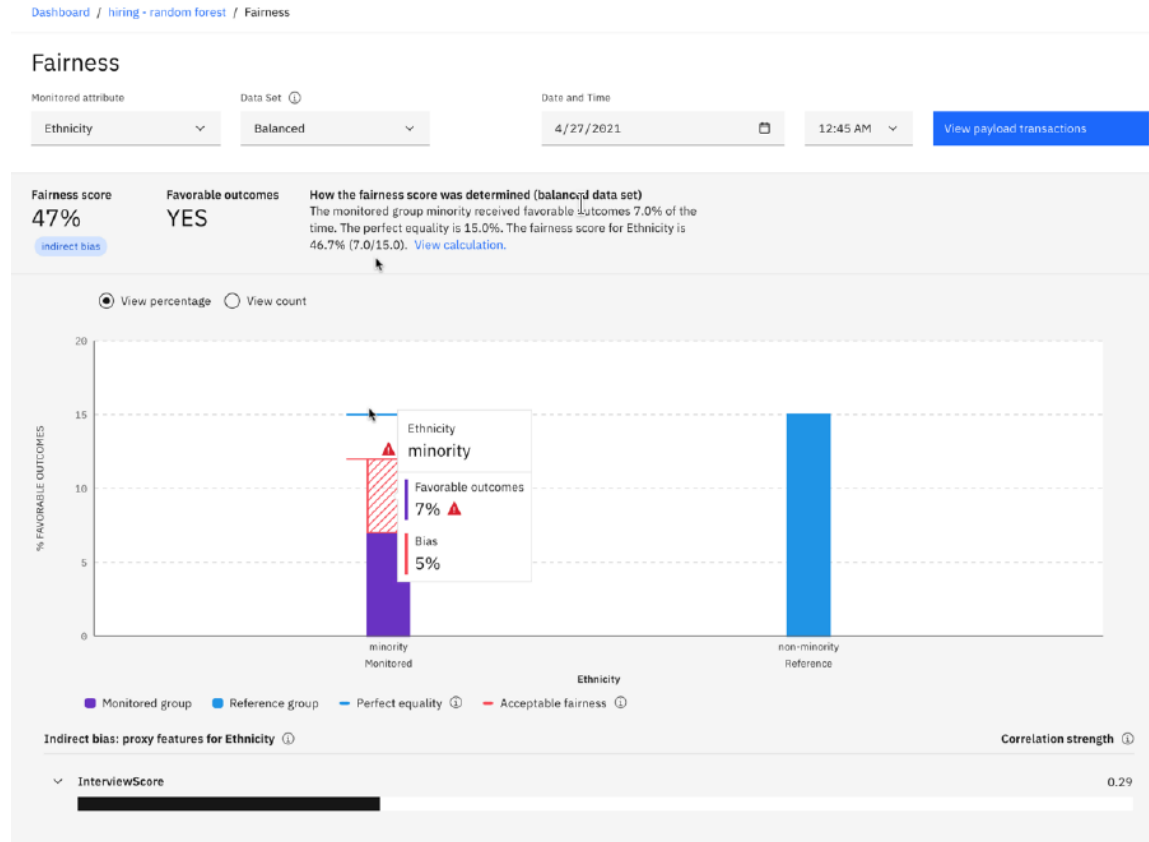
Transaction details



¿Qué ofrece? En el ámbito de la productivización

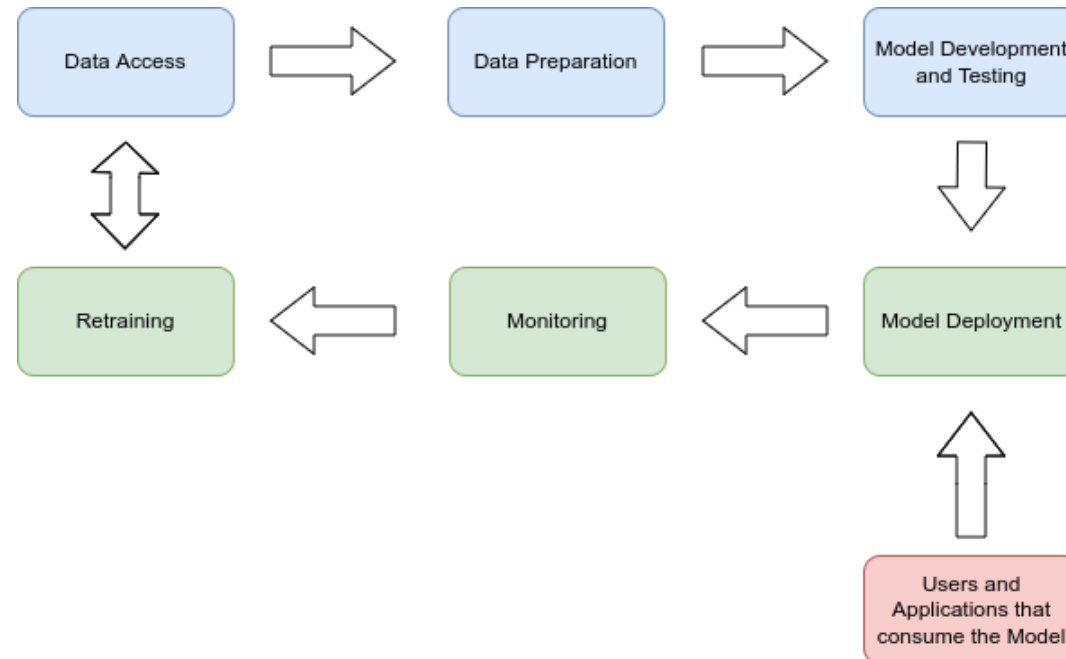
## ■ Detección de sesgos

- Detección de sesgos. Analiza las predicciones de los modelos desplegados en busca de sesgos.



## ■ En resumen...

- Listado no exhaustivo de características.
- Sistema capaz de dar una respuesta integral (extremo a extremo).
- Facilita el control de un ecosistema de elevada complejidad.
- Permite la aceleración en la creación de modelos de IA.



# Servicio Inteligencia Artificial Aplicada

¿Quiénes somos?

■ Servicio integrado actualmente por 3 personas. Todos ellos de la región:

➤ 1 en Talavera (Jorge)

➤ 2 en Toledo (Claudia y Nuria)

■ Muy jóvenes pero... con muy alto nivel de especialización

➤ Jorge:

- Grado en Ingeniería Electrónica Industrial y Automática - UCLM
- Máster en Project Management - Universidad Politécnica de Cataluña
- Máster en Sistemas Inteligentes (IA y Ciencia de Datos) - Universidad de las Islas Baleares

➤ Nuria:

- Grado Ingeniería Informática - Universidad Alcalá
- Máster en Ciencia de Datos y Aprendizaje Automático - Universidad de la Rioja

➤ Claudia:

- Doble Grado Ingeniería Informática y Matemáticas - Universidad Complutense Madrid
- Máster en Formación del profesorado – UNED
- Experiencia en Centro Vasco de Matemática aplicada (BCAM)

¿Qué hacemos?

- Potenciar el uso de sistemas de Inteligencia Artificial en el ámbito de la JCCM.
  - Procedimentar y estandarizar las distintas fases implicadas en los sistemas de Inteligencia Artificial.
  - Definir la arquitectura corporativa y el ecosistema de herramientas.
  - Datos: Elaboración de un catálogo de datos "curados" para su posterior explotación: AI, BI, Open Data,... Generación de datasets.
  - Colaborar/supervisar la redacción de contratos que incluyan sistemas de inteligencia artificial.
  - Velar por el correcto funcionamiento de los sistemas de Inteligencia Artificial desplegados en la casa.

¿Qué hacemos?



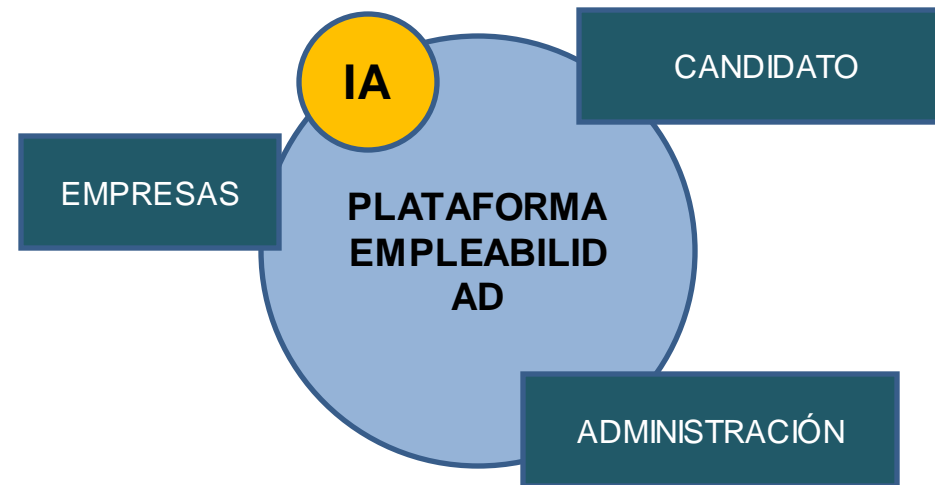
¿Qué estamos haciendo?

Aplicación de modelos de IA para la mejora de la empleabilidad en Castilla-La Mancha

- Se pone el foco en el servicio al ciudadano. Punto en común con Espacio Ciudadano.
- Proyecto que surge a iniciativa de la Viceconsejería de empleo.
- Papel primordial RTIC de Empleo.
- Aprovecha otras inversiones:  
Plataforma IA (Espacio Ciudadano), HW específico IA.

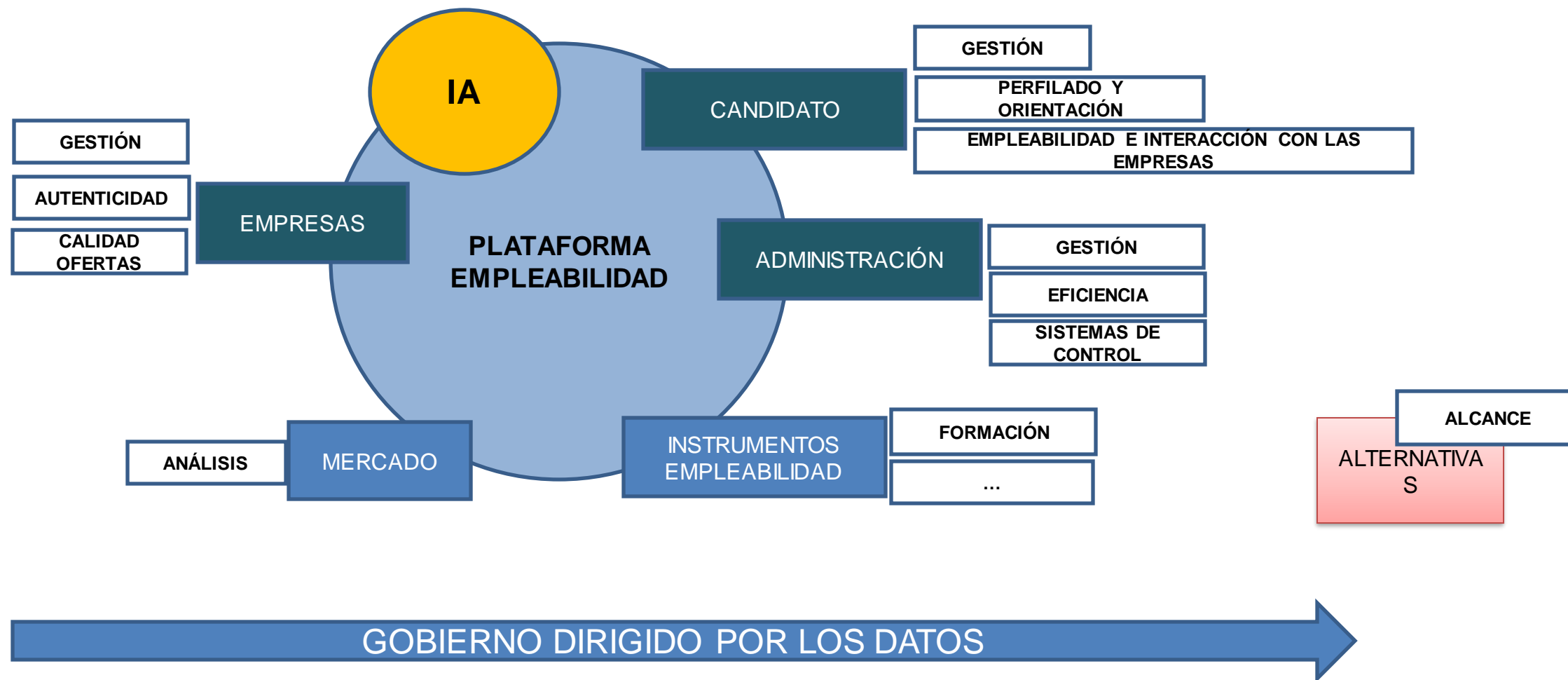
¿Qué estamos haciendo?

- Creación de un espacio virtual de intermediación, que facilite la comunicación entre empresas, administración pública y trabajadores.
- Mejora de la empleabilidad mediante el perfilado de los candidatos y el análisis del mercado.
- Conseguir un gobierno dirigido por los datos .





¿Qué estamos haciendo?



¿Qué estamos haciendo?

- Trabajo conjunto con el Servicio de Modernización
- Se desarrollarán una serie de sistemas de Inteligencia Artificial que serán puestos a disposición mediante servicios REST para su consumo desde diferentes aplicativos.
- Estos sistemas tienen su propio ciclo de vida. Independiente de los desarrollos del aplicativo.
- Apuesta muy ambiciosa.

¿Qué estamos haciendo?

## ■ Sistema 1: Evaluación individual de demandantes de empleo.

- Destinatarios: Personal AAPP y demandantes de empleo.
- Objetivo: Proporciona estimación/predicción de las opciones de integración de un demandante de empleo en el mercado de trabajo en un momento dado, en relación con su situación geográfica y ocupacional.
- Calcula las probabilidades del individuo de cambiar su situación laboral en un período.
  - Inferior a 3 meses.
  - Inferior a 6 meses.
  - Inferior a 12 meses.

¿Qué estamos haciendo?

## ■ Sistema 2: Detección de necesidades formativas

- Destinatarios: Personal AAPP y demandantes de empleo.
- Objetivo: Identificar las necesidades de acciones formativas recomendables para todo el territorio, de manera georreferenciada, y con capacidad de establecer el potencial beneficio laboral obtenido tanto en la oferta como en la demanda.
- Se tiene en cuenta: Conocimiento de las necesidades del tejido productivo y conocimientos ofrecidos por los demandantes de empleo, así como de la evaluación de la ejecución de las políticas desarrolladas en este ámbito en ejercicios anteriores

¿Qué estamos haciendo?

## ■ Sistema 3: Prescripción

- Destinatarios: Personal AAPP y demandantes de empleo.
- Objetivo: Obtener una serie de recomendaciones de posibles acciones que pueden ofrecerse a los demandantes de empleo para mejorar su empleabilidad. Estas acciones tienen que ser valoradas por expertos.
- Las prescripciones se engloban en 3 grupos:
  - Acciones que puede tomar el demandante modificando su perfil. Por ejemplo, ampliar el área de búsqueda de empleo.
  - Inclusión en política activa de empleo
  - Recomendaciones para autoempleo

¿Qué estamos haciendo?

## ■ Sistema 4: Anticipación

- Destinatarios: Personal AAPP.
- Objetivo: Capacidad de anticipación sobre las políticas activas de empleo que deben emprenderse para paliar necesidades potenciales detectables a 12 y 24 meses en el tejido productivo.
- Principalmente sobre la oferta formativa como mecanismo principal de acción en políticas activas de empleo.
- Centrado en predecir las tendencias del mercado laboral a futuro.
- También detección de nuevos yacimientos de empleo.

¿Qué estamos haciendo?

## ■ Sistema 5: Scoring ofertas de empleo y generador de propuestas de mejora de ofertas.

- Destinatarios: Empresas que ofertan puestos de trabajo.
- Objetivo: Predicción de perspectivas de éxito de una oferta mediante el análisis de la misma en un momento dado, en relación a su situación geográfica y respecto a la información histórica existente.
- Obtención de recomendaciones de mejora para mejorar el scoring de una oferta.

¿Qué estamos haciendo?

## ■ Sistema 6: Búsqueda inteligente para demandantes y empresas.

- Destinatarios: Demandantes de empleo y persona de AAPP.
- Objetivo: Sistema de "matching" entre ofertas y demandas de modo que permita:
  - A los demandantes: Obtener un listado ordenado de las ofertas que más se adecúan a su perfil.
  - A los orientadores. Obtener un listado de los demandantes más adecuados para cada oferta disponible.



¿Qué estamos haciendo?

■ Sistema 7: Alertas de condiciones de una oferta.

- Destinatarios: Personal de AAPP.
- Objetivo: Sistema que evaluará las ofertas y alertará cuando sus condiciones económicas y laborales no se ajusten a derecho.

¿Qué estamos haciendo? Fuentes de información

## FUENTES DE INFORMACIÓN INTERNAS



## FUENTES DE INFORMACIÓN EXTERNAS




A screenshot of the 'Informe de Vida Laboral' form from the Spanish government. The form is titled 'INFORME DE VIDA LABORAL' and 'Situaciones'. It includes sections for 'DATOS IDENTIFICATIVOS' (Name and Surname, Social Security Number, DNI) and 'SITUACIONES' (Regimen, Company, Date of Hire, Date of Effect of Hire, Date of Termination).

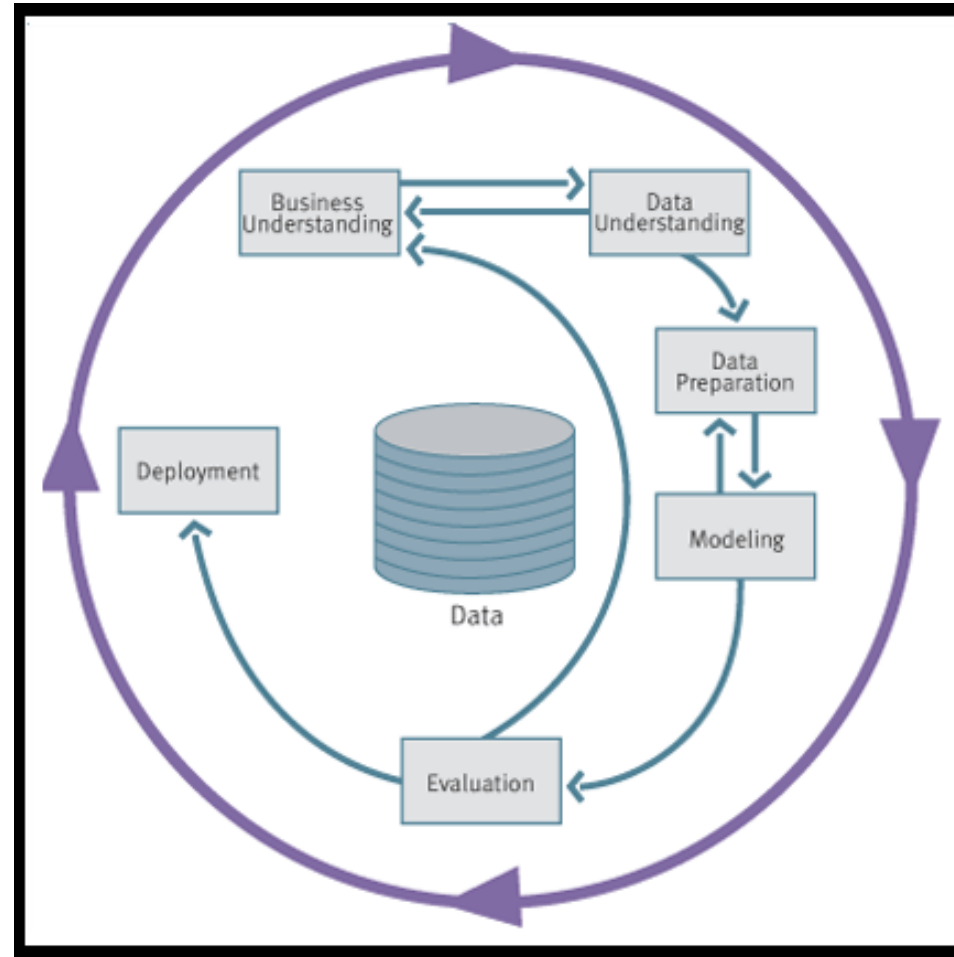
DATOS IDENTIFICATIVOS	
NOMBRE Y APELLIDOS	Nº SEGURIDAD SOCIAL
	DOC D.N.

SITUACIONES				
RÉGIMEN	EMPRESA SITUACIÓN ASIMILADA A LA DE ALTA	FECHA DE ALTA	FECHA DE EFECTO DE ALTA	FECHA DE BAJA

¿Somos autosuficientes?

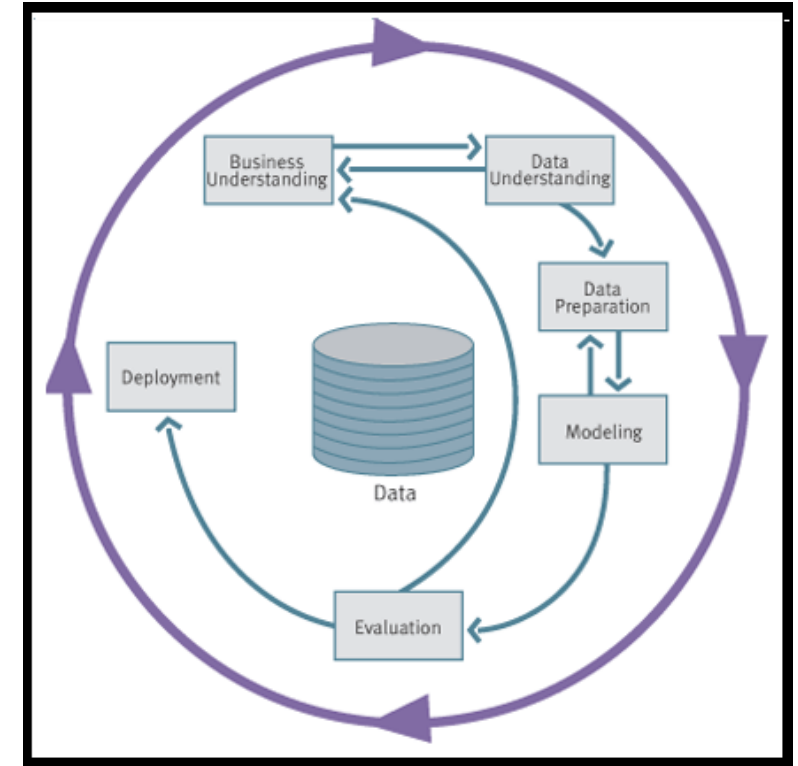
## ■ Metodología CRISP-DM



¿Somos autosuficientes?

■ El proceso comienza con:

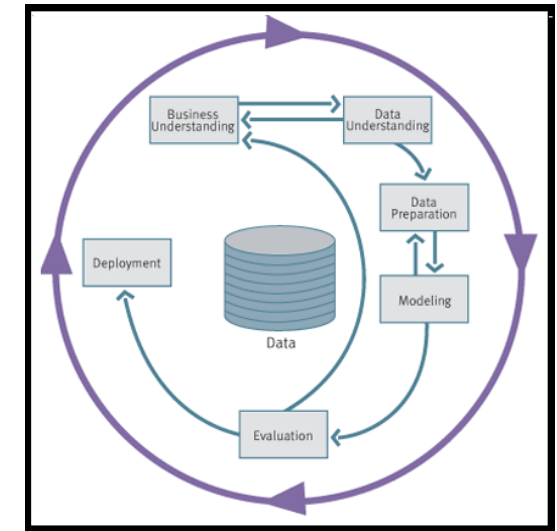
- Conocimiento del negocio. Aplicar IA con sentido.  
Conocimiento indispensable para fijar los objetivos del negocio y estudiar cómo la IA puede ayudar a estos objetivos
- Conocimiento de los datos. Se requiere conocer las fuentes de información disponibles para abordar el desarrollo de este tipo de sistemas.



¿Somos autosuficientes?

## ■ Enfoque multidisciplinar:

- Aspectos jurídicos, protección de datos, ...
- Aspectos técnicos (datos y algoritmos)
- Y sobre todo, conocimiento del negocio:
  - Principales competencias.
  - Focos de interés de las diferentes Consejerías.
  - Oportunidades...



Papel fundamental de los RTICs!!

¡Gracias por vuestra  
atención!