



El futuro digital  
es de todos

MinTIC

**COMENCEMOS DESDE YA,  
A VIVIR LOS HECHOS  
QUE CONECTAN.**



[www.mintic.gov.co](http://www.mintic.gov.co)



El futuro digital  
es de todos

MinTIC

# CALIDAD EN DATOS ABIERTOS

HERRAMIENTAS Y RECURSOS

**Hechos**

QUE

CONECTAN

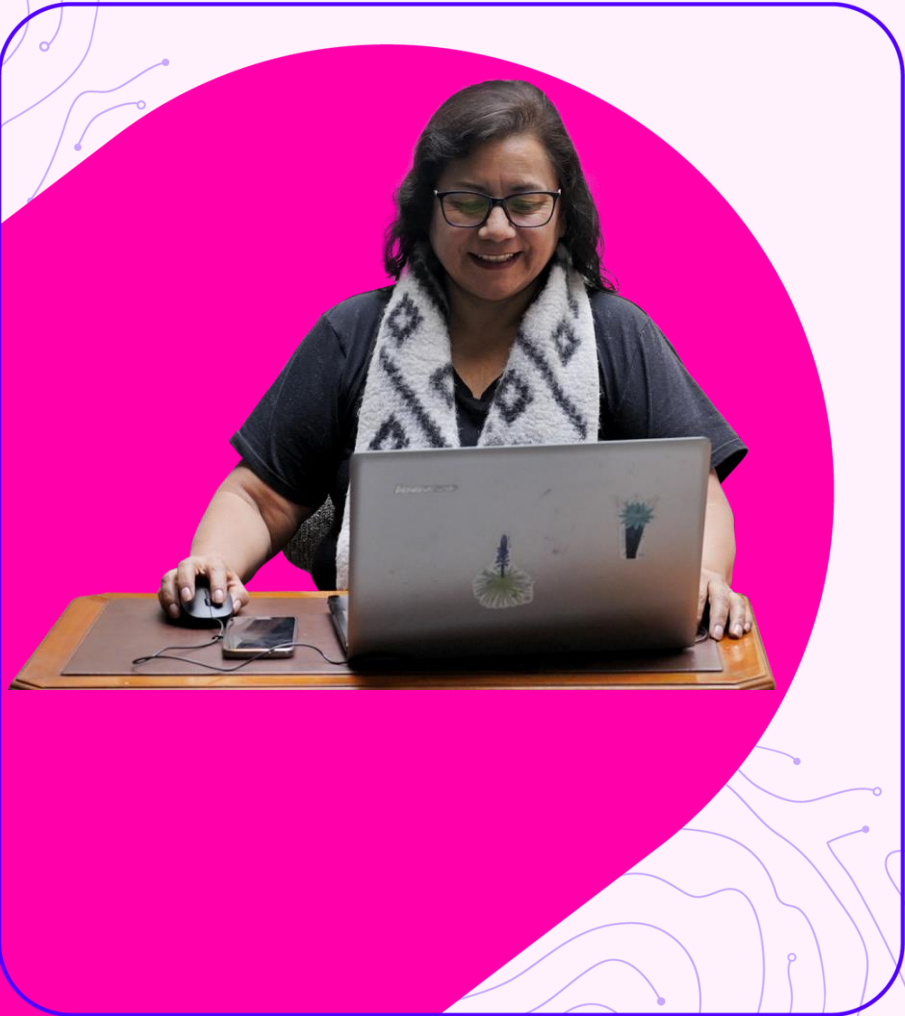


**Julio de 2022**



El futuro digital  
es de todos

MinTIC



# Agenda

1. Portal de datos abiertos
  - Publicación y *Metadata*
1. Cargues masivos
  - Socrata DataSync y sincronización de datos
  - API de Ingesta SodaPy
1. Herramientas de diagnóstico y calificación
  - Diagnóstico en lote
  - Herramienta de autodiagnóstico



El futuro digital  
es de todos

MinTIC

# Publicación y Metadata de Datos Abiertos

**Hechos**  
QUE CONECTAN



# Publicación de datos abiertos



El futuro digital  
es de todos

MinTIC

GOV.CO | DATOS ABIERTOS

Buscar

Inicio Descubre Publica Visualiza Participa Herramientas Novedades Usos Activos Calidad

Español ▼ Iniciar sesión

Debe iniciar sesión para acceder a esta página

Iniciar sesión en la plataforma de datos abiertos del gobierno colombiano

✉ datosabiertos@mintic.gov.co

🔒 .....

¿Ha olvidado su contraseña?

Iniciar sesión

¿Todavía no tiene una cuenta? Registrarse



El futuro digital  
es de todos

MinTIC

# Cargues e Ingesta Masivos

Hechos

QUE

CONECTAN



# Socrata y DataSync

## Republicando datos



El futuro digital  
es de todos

MinTIC

- Es un componente que permite la automatización de la publicación de datos en el portal de Datos Abiertos.
- Puede realizarse a través de una interfaz gráfica o por línea de comandos.
- Multiplataforma
- Manejo de archivos CSV

**Hechos**

QUE

CONECTAN



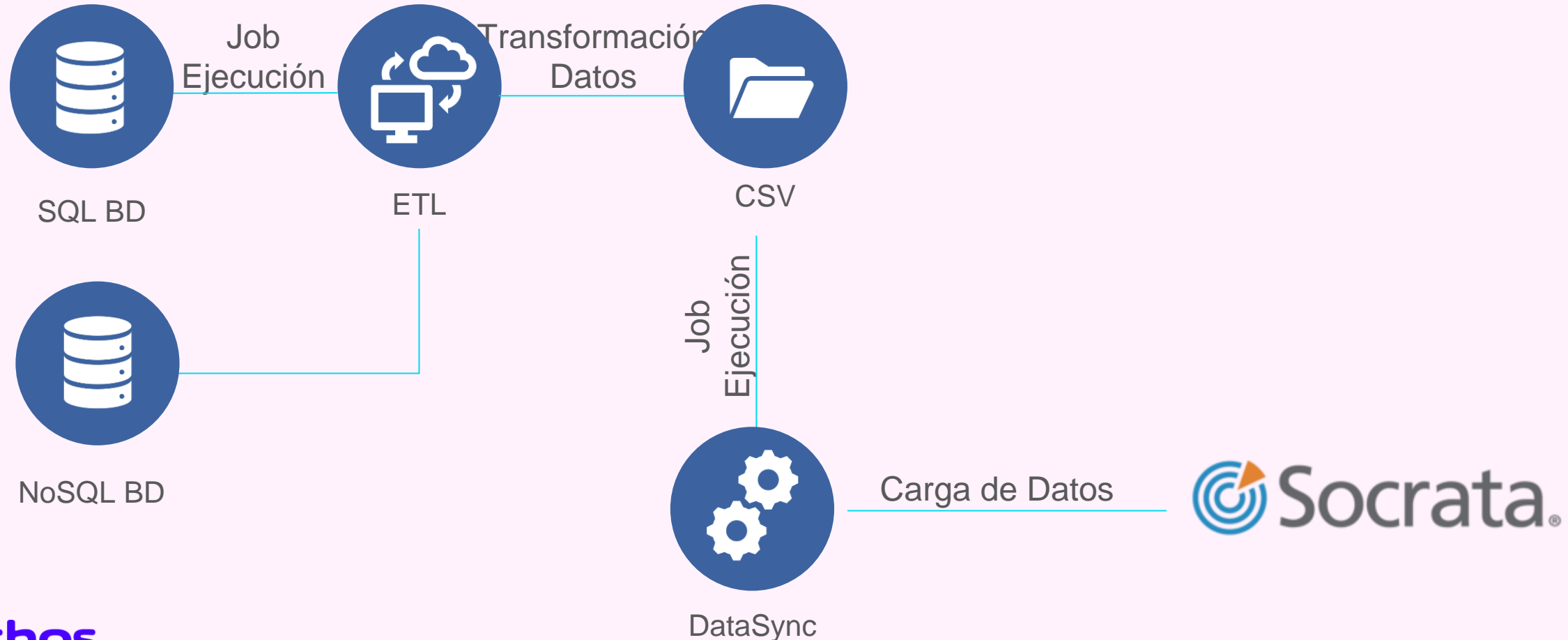
# Socrata y DataSync

## Republicando datos



El futuro digital  
es de todos

MinTIC



**Hechos**  
QUE CONECTAN ✓

<https://www.datos.gov.co/login>



# Socrata y SODAPY

## Ingesta de datos desde el portal



El futuro digital  
es de todos

MinTIC

```
{
  "cell_type": "code",
  "execution_count": null,
  "id": "34db2dbe-d6cf-4de3-94f6-dc0d6c34a1f4",
  "metadata": {},
  "outputs": [],
  "source": [
    "for i in range(1,iteraciones):\n",
    "    offset = i*50000\n",
    "    request_datset = requests.get(f'https://www.datos.gov.co/resource/{dataset}?$limit=50000&$offset={offset}&$order=nivel_entidad')\n",
    "    request_datset = eval(request_datset.text)\n",
    "    df_temp = pd.DataFrame(request_datset)\n",
    "    df_datASET = pd.concat([df_datASET,df_temp])\n",
    "    count += 1\n",
    "    print(count)\n",
    "print(df_datASET)"
  ]
},
```

- Documentación de Socrata. SodaPy y DataSync en <https://dev.socrata.com/>

Hechos

QUE CONECTAN 



El futuro digital  
es de todos

MinTIC

# Diagnósticos de Calidad y autodiagnóstico

Hechos

QUE

CONECTAN

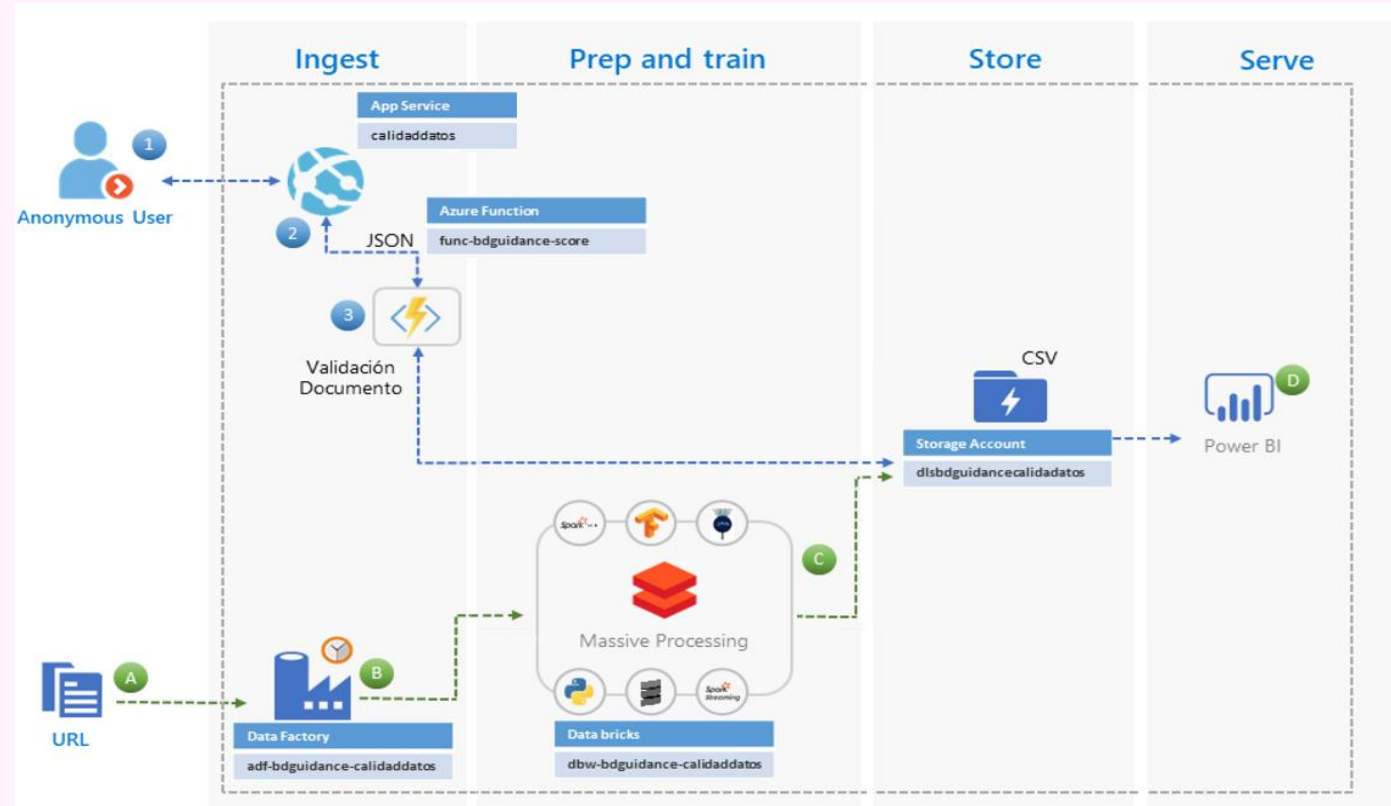


# Diagnóstico en lote



El futuro digital  
es de todos

MinTIC

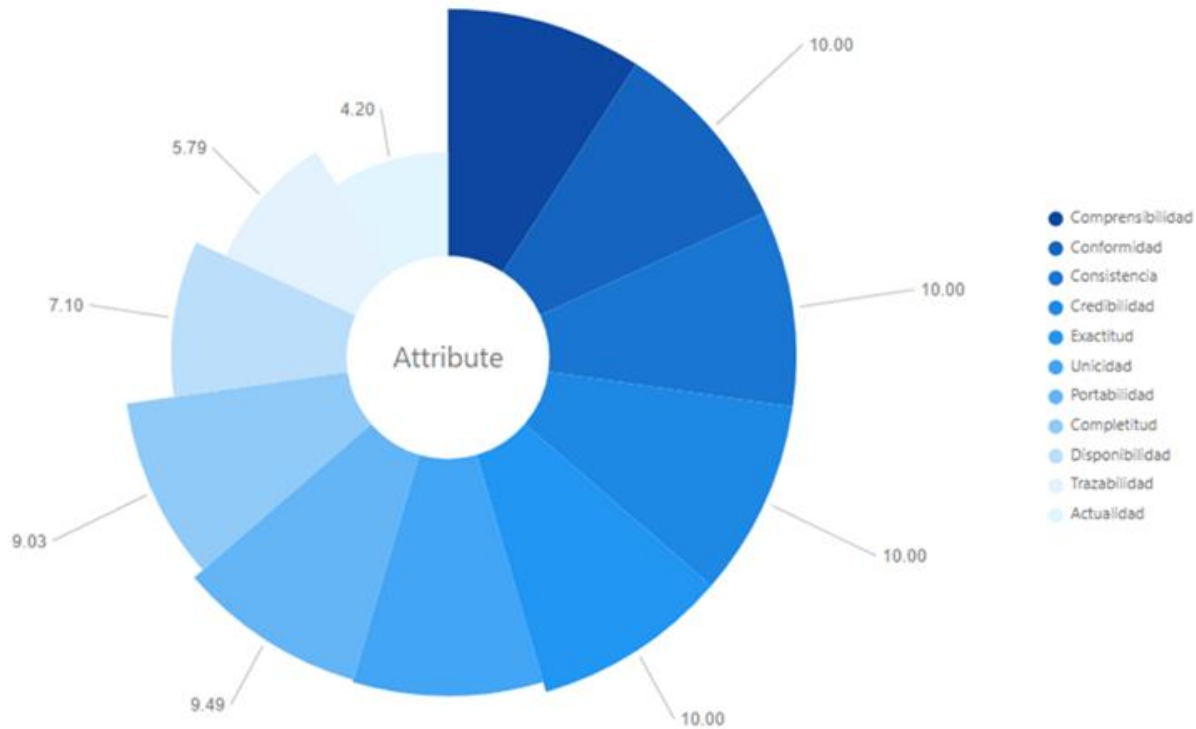


# Diagnóstico en lote



El futuro digital es de todos

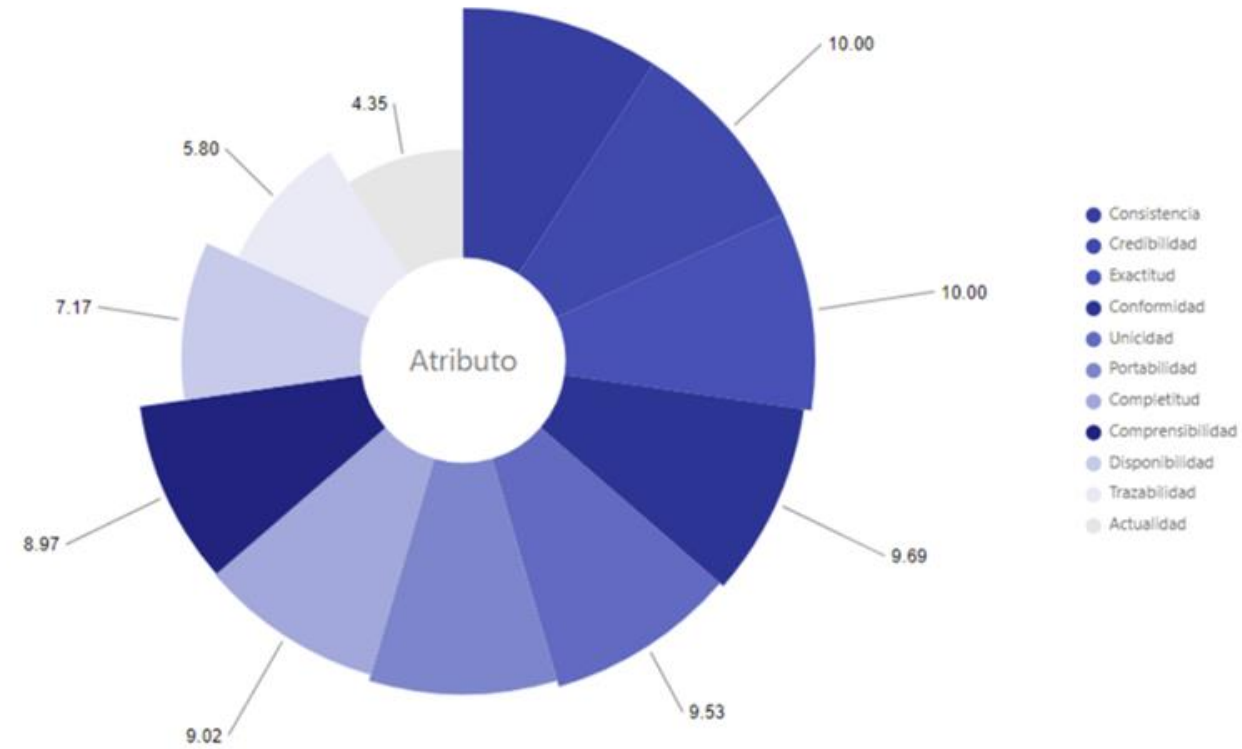
MinTIC



Mayo 31 / 2022

Hechos

QUE CONECTAN



Julio 8 / 2022

# Herramienta de Autodiagnóstico



El futuro digital  
es de todos

MinTIC

← → ↻ calidad-datos-front.azurewebsites.net

GOV.CO | DATOS ABIERTOS

### Diagnóstico de calidad en conjuntos de datos v2.1

A través de esta herramienta puede generar una **pre-evaluación** de referencia a un conjunto de datos en 8 criterios de calidad, de acuerdo con los estándares definidos por MinTIC para calificarlos en una escala de **0 a 10**. Esta calificación es una referencia y no garantiza la publicación de un conjunto de datos en el portal Datos.gov.co, debido a que la calificación para publicación incluye otros criterios adicionales como la Metadata, pero le servirá a su entidad para tener unan **referencia previa**.

**Tenga en cuenta lo siguiente antes de cargar el archivo:**

- Se aceptan únicamente archivos **.CSV**
- Tamaño máximo de archivo **100Mb** (Como referencia aproximadamente son: **200.000 registros**)
- El archivo debe encontrarse en codificación **UTF8**

**Paso 1**



**Seleccionar Archivo**

**Paso 2**

**Subir y procesar**

Hechos

QUE CONECTAN 

<https://calidad-datos-front.azurewebsites.net/>