



# DATOS E INFORMACIÓN PÚBLICA

Guía Práctica

---

Para la publicación de datos en formatos abiertos

# ÍNDICE

- P3. [Introducción](#)  
[Objetivo del documento](#)
- P4. [Formatos abiertos de archivos](#)  
[CSV](#)  
[JSON](#)
- P7. [Fragmentación de archivos](#)  
[Nomenclatura de archivos](#)
- P8. [Codificación](#)  
[Estructura y características de los datos tabulares](#)  
**Recomendaciones Generales**  
[Nomenclatura de los campos \(nombres de las columnas\)](#)  
[Nivel de granularidad de los datos](#)  
[Utilizar orientación vertical en lugar de horizontal](#)  
[Incluir sólo un atributo por campo](#)  
[Valores nulos, desconocidos o en blanco en campos numéricos](#)  
**Recomendaciones para estructurar planillas de cálculo**  
[Utilizar celdas simples](#)  
[Fila de encabezado](#)  
[Celdas vacías en filas para agrupar conceptos](#)  
[Formato de celdas](#)  
**Exportación a CSV**
- P16. [Estándares según el tipo de Datos](#)  
**Texto**  
[Entidades](#)  
[Nombres Propios](#)  
[Siglas](#)  
**Número**  
[Moneda](#)  
[Números telefónicos](#)  
[Coordenadas](#)  
**Tiempo**  
[Rangos horarios](#)  
**Booleano**

# Introducción

---

Este documento tiene su fundamento en la necesidad de los organismos de instrumentar la Política de Datos Abiertos impulsada desde el Gobierno Nacional, cuyo hito fundacional es el Decreto N° 117/2016 del 12 de enero de 2016.

## Objetivo del documento

---

El presente documento tiene por objetivo ofrecer al lector una guía de buenas prácticas para la publicación de datos en formatos abiertos.

Las recomendaciones que aquí pueden encontrarse están basadas en estándares utilizados a nivel nacional e internacional y en la experiencia de trabajo del equipo de la Dirección Nacional de Datos e Información Pública del Ministerio de Modernización de la Nación.

Es nuestra intención que este documento sea considerado como un trabajo colaborativo y en progreso. Por lo dicho, alentamos a organizaciones y ciudadanos a considerarlo de esa manera, y plantear ideas, sugerencias, y comentarios que contribuyan a un mejor documento. De ningún modo entendemos a este trabajo como una versión final y completa.

Consideramos que el documento está orientado a lectores que pueden no ser especialistas en tecnología. Aquellos que sí lo sean seguramente comprendan de manera más integral las implicancias que algunas de las recomendaciones aquí establecidas puedan tener.

En rasgos generales el documento está estructurado de la siguiente manera:

Formatos abiertos de archivos: cuáles son los formatos más usuales en los que se publican datos y cuáles son los más recomendables.

- **Fragmentación de archivos:** cuáles son los criterios para decidir que un archivo es demasiado grande (y hay que fragmentarlo) o demasiado chico (y hay que juntarlo con otros).
- **Nomenclatura de archivos:** cómo nombrar adecuadamente un archivo.
- **Codificación:** cuál es la codificación en que se debe guardar un archivo.
- **Estructura y características de los datos tabulares**
  - o Recomendaciones generales: aplican a todos los recursos.

- Recomendaciones para estructurar planillas de cálculo: aplican exclusivamente al trabajo en planillas de cálculo.
- Exportación a CSV: cómo exportar adecuadamente planillas de cálculo a un archivo de formato abierto.
- **Estándares según el tipo de datos**: cuáles son los estándares recomendados para manejar distintos tipos de datos.

La presente guía aborda los primeros aspectos relevantes para la estandarización de datos. Para una discusión sobre los estándares recomendados para el manejo de datos básicos y fundamentales que son transversales a diversas áreas temáticas, se remite al lector al Guía para la identificación y uso de entidades interoperables.

## Formatos abiertos de archivos

A la hora de producir y almacenar datos existe una gran variedad de tecnologías disponibles (planillas de cálculo, bases de datos, software estadístico más específico, etc). Esto deriva en una profusión, a veces caótica, de datos en una amplia diversidad de formatos.

Considerando que algunos de esos formatos, y en particular algunos ampliamente utilizados no siempre se adecuan a los niveles de apertura deseados. A lo largo de esta sección, ofreceremos algunas pautas y recomendaciones que facilitarán la adaptación y/o transformación de estos formatos hacia otros más abiertos y fácilmente reutilizables.

En el siguiente cuadro consideramos algunos de los formatos más ampliamente utilizados y evaluamos su grado/nivel de apertura:

Formato	Descripción breve	Tipo de datos	Nivel de apertura
XLS	Los XLS son archivos de planilla de cálculo. Es un formato propietario de MS Office.	Tabulares	Baja
XLSX	Los XLSX son archivos con la estructura de un XML. Es un formato abierto basado en Office Open XML (ISO/IEC 29500:2008). Popularizado por ser el formato por defecto del procesador de planillas de cálculo desde MS Office 2007.	Tabulares	Media
ODS	Los ODS son archivos con la estructura de un XML. Es un formato abierto basado en	Tabulares	Media

	OASIS OpenDocument Format (ISO/IEC 26300) . Es el formato por defecto del procesador de planillas de cálculo Open Office.		
CSV	Los archivos CSV son archivos de texto plano donde las columnas se separan por comas y las filas por saltos de línea.  Es un formato abierto.	Tabulares	Alta
JSON	Es un formato para el intercambio de datos. En mucha mayor medida que los formatos anteriormente descritos, JSON es especialmente útil para datos entre máquinas. Es un formato abierto basado en la especificación RFC 7159.	Estructurados	Alta

Un recurso en formato tabular es un archivo plano que se ajusta a un esquema predefinido de columnas, incluyendo el nombre de la columna y el tipo de datos. En la mayoría de los casos se corresponden con datos provenientes de bases de datos, reportes y planillas de cálculo en general. A diferencia de los formatos tabulares, los archivos JSON siguen una estructura diferente donde se definen listas de objetos con pares "clave" - "valor".

En función de lo mencionado recomendamos enfáticamente la publicación de los datos en formato CSV y/o JSON. En caso de utilizar formatos propietarios o aún no estandarizados, resulta de utilidad indicar el software, versión y aplicación que permite procesar esos formatos.

## CSV

El CSV es un formato de archivo de texto plano donde las campos (columnas) se separan por comas y los registros (filas) por saltos de línea. Algunas versiones alternativas de esta forma de publicar datos utilizan otros separadores como punto y coma (";") o pipe ("|").

Los recomendaciones, siguiendo lo indicado por la especificación [RFC4180](#) y las pautas de la [W3C](#), para la publicación de archivos en formato CSV son las siguientes:

- Las filas deben finalizar con los caracteres de "retorno de carro" (\r) y "salto de línea" (\n) unidos (\r\n). Esta forma de separar las líneas se denomina usualmente CRLF ("Carriage Return Line Feed").
- La primera fila siempre contiene los nombres de los campos.

- No se deben repetir nombres entre los campos.
- No se debe colocar espacios al principio ni al final del nombre de un campo, o de un valor.
- Tanto los campos como los valores deben estar separados por comas (",").
- En el caso de que un valor contenga el caracter separador (",") o cualquiera de los caracteres que separan las líneas ("r", "n" o "r\n"), el valor completo debe ser encerrado entre comillas dobles "". Esto indica que el caracter no cumple el rol de separar columnas o filas, sino que es parte de un valor.

Ejemplo:

```
col1,col2\r\n
"La tasa de Juan, está vacía", La tasa de Pablo está llena\r\n
"La tasa de Juan\n está vacía", La tasa de Pablo está llena\r\n
"La tasa de Juan\r\n está vacía", La tasa de Pablo está llena\r\n
```

- En el caso de que un valor contenga el caracter comilla doble ("), el valor debe ser encerrado entre comillas dobles como en el caso anterior (") y, además, los caracteres comilla doble que se encuentren dentro del valor deben escribirse dos veces (").

Ejemplo:

```
col1,col2\r\n
"La tasa de ""Juan"" , está vacía", La tasa de Pablo está llena\r\n
```

- Para todos los tipos de datos se considera válido el valor indefinido. Este se expresará con la ausencia de todo caracter y no con un caracter o string especial como podrían ser ".", "null", "none", "nan", etc.

Ejemplo:

```
col1,col2\r\n
a,,b\r\n
a,c,\r\n
```

## JSON

JSON es un formato de texto popular para el intercambio de datos, es un acrónimo de JavaScript Object Notation. Por su característica de ser un formato de tipo estructurado es especialmente útil para el intercambio de datos entre máquina (machine - readable format).

El formato JSON ha sido definido por la especificación [RFC 7159](#), y tal como CSV también es un estándar abierto.

Las recomendaciones que adoptamos para el formato JSON están fuertemente vinculadas con nuestras propuestas para metadatos, por lo que remitimos al lector a visitar ese apartado específico.

En la próxima sección daremos recomendaciones orientadas a mejorar el proceso de trabajo con datos en los casos mencionados, tendientes a facilitar su posterior exportación a CSV.

## Fragmentación de archivos

---

A los fines de garantizar una adecuada accesibilidad a los datos, y teniendo en cuenta las limitaciones de las herramientas más ampliamente difundidas, es necesario considerar la fragmentación de archivos excesivamente grandes, entendiendo por esto aquellos archivos que superen el millón de filas.

Para realizar la fragmentación de archivos se recomienda la utilización de conceptos simples, como ser: (a) por períodos en caso de tratarse de información temporal (Ej. Años, semestres, trimestres, meses, semanas, días), (b) por zonas en caso de tratarse de

información geográfica (Ej. provincias, municipios, barrios, secciones, o manzanas) o (c) por dimensiones temáticas propias del dominio particular de la información.

En todos los casos, siempre que se decida fragmentar un archivo por alguna de sus cualidades se recomienda incorporar una versión completa de los datos a los fines de evitar la tarea de consolidación.

Para estos casos se recomienda utilizar protocolos de compresión, en especial en caso de tratarse de archivos muy grandes, y altamente compresibles. De utilizarlos se aconseja protocolos sin pérdida, y abiertos.

## Nomenclatura de archivos

---

Para nombrar los archivos se recomienda seguir las siguientes convenciones:

- Utilizar palabras siempre en minúsculas
- No usar artículos
- Usar únicamente letras y números ASCII, siempre en minúsculas, comprendidos en el rango "a-z" y "0-9"
- Separar las palabras con guión medio "-"
- En caso de corresponder, ubicar la referencia temporal o del atributo de fragmentación siempre al final

### Ejemplos:

- *"acceso-informacion-publica.csv"*: Versión completa del recurso.
- *"acceso-informacion-publica-2013.csv"*: Versión del recurso fragmentada por año.
- *"acceso-informacion-publica-201302.csv"*: Versión del recurso fragmentada por mes.
- *"acceso-informacion-publica-caba.csv"*: Versión del recurso fragmentada por división político-territorial (provincia o caba).
- *"acceso-informacion-publica-caba-2013.csv"*: Versión del recurso fragmentada por división político-territorial (provincia o caba) y año.
- *"acceso-informacion-publica-jujuy-20130208.csv"*: Versión del recurso fragmentada por división político-territorial (provincia o caba) y fecha.

En el caso de la fragmentación temporal, se recomienda adoptar el estándar que se vislumbra en los ejemplos ya que es compacto y ordena los recursos adecuadamente por tiempo: YYYYMMDD. Debe recordarse mantener siempre dos dígitos para el mes y el día, a pesar de que el número sea menor a 10.



En el caso de la fragmentación por zonas, se remite al Guía para la identificación y uso de entidades interoperables para una discusión más detallada sobre cómo nombrarlas adecuadamente. En el caso de utilizar dimensiones temáticas propias del dominio particular de la información, se remite a la misma guía o se recomienda utilizar el mejor estándar identificado para esa temática particular.

## Codificación

---

Todos los recursos de datos, incluyendo los geográficos, deben publicarse utilizando la codificación UTF-8 siguiendo las recomendaciones de la W3C.

Una de las principales razones para esta selección, entre otras, es que UTF-8 soporta una gran variedad de lenguajes, según la W3C es un: “estándar en el que se definen todos los caracteres necesarios para la escritura de la mayoría de los idiomas hablados en la actualidad. Su objetivo es ser, y, en gran medida, ya lo ha logrado, un superconjunto de todos los sets de caracteres que se hayan codificado”.

## Estructura y características de los datos tabulares

---

En el apartado anterior, referente a los formatos de archivo, hemos comentado la conveniencia de publicar datos en formatos abiertos. En esta sección introduciremos (a) recomendaciones generales para el trabajo con datos y (b) recomendaciones para el trabajo con planillas de cálculo, orientadas tanto a facilitar su exportación a formatos abiertos como su propia usabilidad en el contexto de cualquier aplicación de planillas de cálculo.

### Recomendaciones Generales

A continuación realizaremos algunas recomendaciones para el trabajo con datos tabulares en general, que consideramos conveniente adoptar más allá de las tecnologías específicamente utilizadas.

#### Nomenclatura de los campos (nombres de las columnas)

Por “nomenclatura de los campos” nos referimos al nombre de las columnas en los datos de estructura tabular. Las siguientes recomendaciones aplican a la generalidad de los

casos, pero cuando existan convenciones particulares según la temática o rubro de datos de que se trate y estas entren en conflicto, puede ser conveniente privilegiar primero la convención de la temática específica y luego la convención general.

Los nombres de los campos deben:

- Estar en español.
- Ser lo más explícitos, descriptivos y declarativos como sea posible.
  - o Es preferible que el nombre de un campo sea claro antes que corto, pero se recomienda no superar los 50 caracteres.
  - o No usar abreviaturas si no es estrictamente necesario o recomendado por una convención ampliamente difundida. En caso de usarlas, incluirlas en el diccionario de datos.
- Estar en minúsculas, no incluir caracteres especiales, ni estar subrayados.
  - o Utilizar palabras compuestas únicamente de caracteres en minúsculas comprendidos en el rango a-z (letras sin tildes) y en el rango 0-9 (dígitos).
  - o Las palabras deben unirse con guión bajo " \_ ".
  - o No contener espacios.
  - o Las palabras deben separarse siempre con " \_ ", en lugar de no tener separación alguna: fecha\_audiencia\_solicitada en lugar de fechaaudienciasolicitada
- Referirse a un sólo atributo de los datos, indivisible en más de un campo.
  - o Los campos deben separar los atributos de los datos en la forma más desagregada que sea posible.
  - o Se debe evitar definir campos que contengan más de un tipo de información (por ejemplo: e-mail y sitio web, número de teléfono, etc bajo "datos\_de\_contacto").
- Si existe una entidad que engloba varias características separadas en campos diferentes, comenzar nombrando los campos con esa entidad y luego con los atributos más específicos (de lo más general a lo más específico).
  - o Ej.: solicitante y solicitante\_documento son entidades más generales que se repiten en varios campos, que corresponden a atributos más específicos.
    - solicitante\_nombre
    - solicitante\_cargo
    - solicitante\_documento\_tipo
    - solicitante\_documento\_numero
  - o Resulta más fácil agrupar mentalmente campos relacionados entre sí, queda más claro que son atributos de una misma entidad y el orden alfabético de los campos los dejaría automáticamente agrupados.
  - o Incluso cuando la entidad de un atributo parezca evidente (Ej.: un dataset llamado "audiencias" donde todos los campos son atributos de la entidad "audiencia") se recomienda nombrar el campo incluyendo la entidad a la que hace referencia el atributo.
    - **No recomendado:** "id", "fecha\_hora"

- **Recomendado:** “audiencia\_id”, “audiencia\_fecha\_hora”.

Los campos que sean identificadores o códigos deberán incluir el sufijo “\_id” en el nombre del campo, salvo casos excepcionales donde un nombre alternativo sea más conveniente porque ofrece información sobre el sistema de identificación utilizado.

- En cuanto a los campos que contengan la descripción de ese identificador, se recomienda que incluyan el sufijo “\_desc” (por “descripción”) en caso de que no exista una forma más conveniente de nombrar el campo.

sucursal_id	sucursal_desc
1	Nueva Pompeya
2	Barracas
3	La Quiaca

## Nivel de granularidad de los datos

No incluir totales, subtotales ni agrupamientos de datos. Un dataset debe ser consistente en el nivel de granularidad de los datos que contiene. Está bien tener un dataset con la cantidad de convenios firmados por provincia y está bien tener un dataset con la cantidad de convenios firmados por municipio. No está bien tener un dataset que mezcle ambos.

Dicho esto, el dato agregado “convenios firmados por provincia” siempre se puede calcular a partir de un proceso del dataset más desagregado, pero esto no es así a la inversa (es imposible recuperar los datos a nivel de municipio desde el dataset provincial).

**No recomendado** - datos con subtotales y/o totales incluidos

(diferentes niveles de granularidad)

provincia_nombre	municipio_nombre	convenios_firmados_anio	convenios_firmados_numero
Provincia X	Municipio W	2011	10
Provincia X	Municipio X	2011	15
Provincia X	Subtotal	2011	25
Provincia Y	Municipio Z	2011	5

Provincia Y	Subtotal	2011	5
	TOTAL	2011	30

**Recomendado** - datos de un mismo nivel de granularidad

provincia_nombre	municipio_nombre	convenios_firmados_anio	convenios_firmados_numero
Provincia X	Municipio W	2011	10
Provincia X	Municipio X	2011	15
Provincia Y	Municipio Z	2011	5

### Utilizar orientación vertical en lugar de horizontal

Es preferible que la orientación de los datos sea “vertical” en lugar de “horizontal” en los casos en que esto sea posible. La principal razón es que los datos orientados de manera vertical facilitan el tratamiento y análisis de los datos.

**No recomendado** - Orientación horizontal

municipio_nombre	solicitudes_anio	solicitudes_poda_y_arbolado_numero	solicitudes_recoleccion_residuos_numero
Municipalidad X	2015	340	198

**Recomendado** - Orientación vertical

municipio_nombre	solicitudes_anio	solicitudes_categoria	solicitudes_numero
Municipalidad X	2015	Poda y arbolado	340
Municipalidad X	2015	Recolección de residuos	198

## Incluir sólo un atributo por campo

Se recomienda definir los campos de forma atómica de modo de incluir un sólo atributo por elemento, en lugar de datos múltiples, generando campos adicionales de ser necesario.

**No recomendado** - múltiples datos en una celda

municipio_nombre	solicitudes_anio	categoria_solicitudes_numero_y_tipo
Municipalidad X	2015	Poda y arbolado - 340 Recolección de residuos - 198

**Recomendado** - un dato por celda

municipio_nombre	solicitudes_anio	solicitudes_categoria	solicitudes_numero
Municipalidad X	2015	Poda y arbolado	340
Municipalidad X	2015	Recolección de residuos	198

## Valores nulos, desconocidos o en blanco en campos numéricos

Los valores de los datos deben ser siempre explícitos y respetando el tipo de datos del campo de que se trate. Los elementos o celdas en blanco se interpretarán siempre como “valor ausente”.

Si existen distintas interpretaciones posibles de un “valor ausente”, estas deben ser explicitadas en un campo aparte. Si sólo hay “valores ausentes” (no hay distintas interpretaciones, es siempre un “valor ausente”) no es necesario agregar una columna adicional.

Es importante destacar, por ejemplo, que cuando un valor numérico sea “0” siempre debe ponerse un “0” como dato (y no un valor nulo, en blanco o vacío).

**No recomendado** - texto presente en campos numéricos

solicitudes_numero
34
NA
NS/NC

...
567

**Recomendado** - texto excluido de campos numéricos

solicitudes_numero	solicitudes_numero_missing_desc
34	
	No disponible
	No sabe / No contesta
	Valor ausente
	Valor ausente
0	
567	

## Recomendaciones para estructurar planillas de cálculo

Las recomendaciones de esta sección aplican exclusivamente al trabajo sobre planillas de cálculo.

### Utilizar celdas simples

En el trabajo con planillas de cálculo recomendamos utilizar celdas simples, es decir, en ningún caso combinar celdas.

**No recomendado** - celdas combinadas

proveedor_nombre	contacto_datos	
	correo_electronico	telefono
Ejemplo Sociedad Anónima	ejemplo@ejemplo.com.ar	1143XXXXXX

**Recomendado** - celdas simples, sin combinar

proveedor_nombre	contacto_correo_electronico	contacto_telefono
Ejemplo Sociedad Anónima	ejemplo@ejemplo.com.ar	1143XXXXXX

### Fila de encabezado

Los datos deben contener solo una fila de encabezado. Desde la segunda fila en adelante, sólo debe haber datos, pero nunca un encabezado.

**No recomendado** - múltiples filas de encabezado

Nombre del	Dirección de correo	Teléfono de
proveedor	electrónico de contacto	contacto
Ejemplo Sociedad Anónima	ejemplo@ejemplo.com.ar	1143XXXXXX

**Recomendado** - encabezado de una única fila

proveedor_nombre	contacto_correo_electronico	contacto_telefono
Ejemplo Sociedad Anónima	ejemplo@ejemplo.com.ar	1143XXXXXX

### Celdas vacías en filas para agrupar conceptos

Se recomienda no dejar celdas vacías en filas bajo la presunción de que valores en blanco posteriores a un valor positivo contienen implícitamente a ese mismo valor en una suerte de “agrupamiento conceptual”.

Este error es muy común en la construcción de planillas de cálculo y suele generar problemas graves cuando cambia el orden original de las filas, además de que impide el uso de tablas dinámicas y otras formas de analizar los datos.

**No recomendado** - primera celda de la segunda fila vacía

municipio_nombre	solicitudes_anio	solicitudes_categoria	solicitudes_numero
Municipalidad X	2015	Poda y arbolado	340

	2015	Recolección de residuos	198
--	------	-------------------------	-----

**Recomendado** - primera celda de la segunda fila completa

municipio_nombre	solicitudes_anio	solicitudes_categoria	solicitudes_numero
Municipalidad X	2015	Poda y arbolado	340
Municipalidad X	2015	Recolección de residuos	198

## Formato de celdas

Las celdas de una planilla de cálculo deben estar formateadas en forma acorde al tipo de datos de que se trate. Específicamente los números siempre deben estar en celdas de formato/tipo "número", los campos de tipo textual deben estar en celdas de formato/tipo "texto" y los campos de tipo fecha deben estar en celdas de formato/tipo "fecha".

audiencia_fecha_hora	audiencia_sujeto_obligado_nombre	audiencia_numero
1/3/16	Juan	3456

**Fecha**

**Texto**

**Número**

Mantener el formato correcto de las celdas según el tipo de datos que contengan (a) mejora las las probabilidades de que una exportación a otro formato salga correctamente y (b) hace que los datos sean más operables en la propia planilla de cálculo, aprovechando mejor sus funcionalidades.

## Exportación a CSV

Como se dijo anteriormente, el CSV es el formato más recomendado para la publicación de archivos tabulares. A la hora de exportar una planilla de cálculo a CSV hay 3 parámetros que deben ser especificados durante la exportación, independientemente del software de que se trate:

- Codificación (encoding, en inglés): Siempre debe ser UTF-8.
- Caracter separador (separator character, en inglés): Siempre debe ser "," (coma).
- Caracter calificador (quote character o enclosing character, en inglés): Siempre debe ser " " " (comillas dobles).



# Estándares según el tipo de Datos

---

El formato recomendado para los distintos tipos de datos está mayormente basado en las especificaciones de la W3C. En los otros casos, las recomendaciones surgen de la experiencia de trabajo del equipo de la Dirección Nacional de Datos e Información Pública y del esfuerzo realizado en la búsqueda de estándares más adecuados.

## Texto

- Los campos de texto no deben contener espacios en blanco innecesarios al principio ni al final.

## Entidades

Las entidades que aparezcan entre los datos de un campo textual deben tener una descripción única. Es decir, toda mención que se realice a una entidad dada debe hacerse utilizando exactamente la misma cadena de caracteres cada vez:

- Las descripciones de entidades deberían elegirse siempre de forma tal que cumplan con el estándar específico que las describe, en caso de que este exista.
- Cuando este estándar no existe y hay dudas respecto del criterio a adoptar para elegir la descripción única de una entidad, debe privilegiarse siempre aquella que sea lo más explícita, descriptiva y declarativa posible.

## Nombres Propios

Se capitalizan (primera letra de cada palabra es mayúscula, el resto de las letras son minúsculas) todas las palabras significativas, salvo las siglas. Las palabras significativas son aquellas que no cumplen la función de artículos o preposiciones.

## Siglas

Todas las siglas se escriben en mayúsculas, sin usar puntos ni espacios intermedios.

Ejemplo:

**No recomendado**

localidad_nombre
Ciudad Autónoma de Buenos Aires
CABA
Capital Federal
Ciudad de Buenos Aires

Recomendado - primera celda de la segunda fila completa

localidad_nombre
Ciudad Autónoma de Buenos Aires
Ciudad Autónoma de Buenos Aires
Ciudad Autónoma de Buenos Aires
Ciudad Autónoma de Buenos Aires

En el ejemplo anterior, los 4 valores de texto refieren a la misma entidad. Debe elegirse una única forma de referirse a la misma y utilizarla en todos los casos.

Siempre que sea posible, la elección deberá fundamentarse en el estándar establecido para ese tipo de entidad (para más información ver el Guía para la identificación y uso de entidades interoperables) o en el caso de no existir un estándar deberá adecuarse a las pautas generales contexto del dataset de que se trate.

## Número

- El separador decimal debe ser el caracter ".".
- No se utilizará separador de miles.
- No se deberán utilizar espacios en blanco.
- Para los números negativos debe incluirse el símbolo menos "-" inmediatamente antes del número, sin espacio en blanco intermedio.

## Moneda

A todos los efectos, los valores numéricos que sean además valores monetarios se consideran números y, por lo tanto, valen las mismas recomendaciones que para ellos. Adicionalmente, se agregan las siguientes recomendaciones:

- La cantidad de decimales debe limitarse a dos, salvo que el uso de una mayor cantidad de decimales sea significativo para el caso particular.
- En ningún caso se incluirán símbolos o letras en el campo numérico -ya sea "\$", "DOL", "USD", etc.

De indicarse en el recurso diferentes tipos de moneda se recomienda hacerlo en un campo aparte utilizando los códigos alfabéticos definidos en la [ISO 4127](#).

Ejemplo:

- ARS, para el peso argentino.
- USD, para el dólar estadounidense.

## Números telefónicos

En el presente apartado nos abocaremos a proponer una solución para la inclusión de números telefónicos nacionales en los recursos de datos. A nivel internacional el estándar para los números telefónicos ha sido desarrollado por el "Sector de Normalización de las Telecomunicaciones de la Unión Internacional de Telecomunicaciones" (ITU Telecommunication Standardization Sector) bajo la recomendación [E.164](#), estándar que no trataremos en el presente apartado.

Para el caso de los números nacionales, el ENACOM tiene la competencia sobre el sistema de numeración telefónica. El organismo mencionado indica que el número nacional de abonado debe estar compuesto por 10 dígitos. Estos 10 dígitos están conformados por un indicativo interurbano más un número de abonado. Pudiendo el indicativo interurbano tener entre 2 y 4 dígitos, y el número de abonado entre 6 y 8 dígitos.

Indicativo interurbano (ámbito geográfico al que corresponde)	Número de abonado	Número Nacional interurbano = Indicativo interurbano + Número de abonado
11 (AMBA)	4343XXXX	114343XXXX
351 (Ciudad de Córdoba)	434XXXX	351434XXXX
3837 (Tinogasta)	43XXXX	383743XXXX

La numeración citada es válida para los teléfonos móviles, pero dado que para llamar a un móvil desde un teléfono fijo es necesario anteponer "15" al número de abonado, será necesario que el registro del número telefónico especifique de alguna manera si se trata de un móvil o de un teléfono fijo.

Al no existir estándares nacionales para la inclusión de números telefónicos en recursos de datos, los números telefónicos suelen indicarse de múltiples maneras como podemos ver en la tabla a continuación:

**No recomendado** - Múltiples formas de indicar un número telefónico

proveedor_nombre	contacto_correo_electronico	contacto_telefono
Ejemplo Sociedad Anónima	ejemplo@ejemplo.com.ar	01143XXXXXX
Ejemplo2 Sociedad Anónima	ejemplo2@ejemplo2.com.ar	011-45XXXXXX
Ejemplo3 Sociedad Anónima	ejemplo3@ejemplo3.com.ar	351 434-XXXX
Ejemplo4 Sociedad Anónima	ejemplo4@ejemplo4.com.ar	011 15 6344-XXXX

Para los recursos que contengan números telefónicos nacionales se recomienda como mínimo:

- Respetar el estándar definido por el Número Nacional Interurbano utilizando la conformación de números mediante 10 dígitos.
- Asegurarse de indicar si el teléfono es móvil o fijo.
- Omitir el agregado de dígitos adicionales en el indicativo interurbano, se recomienda no indicar cero inicial antes del código de área.

Con las salvedades que comentaremos al final de este apartado, un posible abordaje sería el de la tabla a continuación:

**Recomendado** - adecuado al estándar del Número Nacional Interurbano

proveedor_nombre	contacto_correo_electronico	tipo_numero_telefono	contacto_telefono_indicativo_interurbano	contacto_telefono_numero_abonado
Ejemplo Sociedad Anónima	ejemplo@ejemplo.com.ar	Fijo	11	43XXXXXX
Ejemplo2	ejemplo2@ejemplo2.com.ar	Fijo	11	45XXXXXX

Sociedad Anónima				
Ejemplo3 Sociedad Anónima	ejemplo3@ejemplo3.com.ar	Fijo	351	434XXXX
Ejemplo4 Sociedad Anónima	ejemplo4@ejemplo4.com.ar	Móvil	11	6344XXXX

Esta recomendación no contempla los siguientes casos específicos:

- No será aplicable a números de uso público, ejemplo: 100, Bomberos; 911, Policía Federal; etc.
- Para casos que requieran la inclusión de más de un número telefónico. Deberán agregarse campos o modificar la estructura de la base de datos.
- Para teléfonos que requieran la inclusión de un número de interno, y al estar este definido por la persona u organización específica, deberá considerarse la inclusión de otro campo de tipo texto. Ya que los números de interno pueden incluir texto, ejemplo: “\*86”, “#36”, etc.
- Para el casos de números internacionales se recomienda contemplar el estándar internacional E.164.

## Coordenadas

Para registrar datos de coordenadas geográficas de puntos se utilizarán números decimales. Los campos deberán llamarse “latitud” y “longitud”. Cuando sea conveniente especificar el nombre de la entidad de la cual se consignan las coordenadas, se utilizarán los sufijos “\_latitud” y “\_longitud”.

### No recomendado

latitud	longitud
34° 11' 30"	58° 43' 20"

### Recomendado

latitud	longitud
---------	----------

-34.6043222	-58.4134862
-------------	-------------

## Tiempo

Se usará el estándar ISO 8601 (YYYY-MM-DDTHH:MM:SS[.mmmmmm][+HH:MM]). A menos que se indique lo contrario, se asumirá que la zona horaria es UTC-03:00 (Argentina).

- Fecha: YYYY-MM-DD
- Hora: HH:MM:SS[.mmmmmm][+HH:MM]
- Fecha y Hora: YYYY-MM-DDTHH:MM:SS[.mmmmmm][+HH:MM]
- Duración: YYYY-MM-DDTHH:MM:SS[.mmmmmm]

### Rangos horarios

- Los rangos estarán divididos en dos partes separadas por un doble guión bajo "\_\_", la primera indica el día y la segunda, la hora.
- Se puede omitir la parte del día o bien de la hora pero nunca ambas
- Si se omite la parte que indica el día se asumirá que el rango abarca todo el horario indicado
- Si se omite la parte que indica el horario se asumirá que el rango abarca todo el día indicado
- El día se puede indicar tanto mediante rangos separando los días con guiones medios "-" o bien como particulares con el guión bajo "\_".

Ejemplos de formatos válidos para días:

<b>DAY:</b> Un solo día
<b>DAY1-DAY2:</b> Entre entre DAY1 y DAY2
<b>DAY1_DAY2:</b> DAY1 y DAY2
<b>DAY1-DAY2_DAY3:</b> DAY1 a DAY2 y DAY3

- La hora se indica mediante rangos separando los horarios con guiones medios ("-"). También se pueden indicar varios horarios con el guión bajo "\_".

Ejemplos de formatos válidos para horas:

**HH:MM-HH:MM:** Rango simple  
**HH:MM-HH:MM\_HH:MM-HH:MM:** Dos rangos

Más ejemplos de formatos válidos completos:

**HH:MM-HH:MM** para indicar un rango que ocurre todos los días.  
**DAY** para indicar que el rango ocupa todo el día *DAY*.  
**DAY\_\_HH:MM-HH:MM** para indicar un rango que ocurre los días *DAY* entre *HH:MM* y *HH:MM*.  
**DAY\_\_HH:MM-HH:MM\_HH:MM-HH:MM** para indicar mas un rango horario en el mismo día  
**DAY1-DAY2\_\_HH:MM-HH:MM** para indicar un rango que ocurre los días *DAY1* a *DAY2* entre *HH:MM* y *HH:MM*  
**DAY1-DAY2\_\_HH:MM-HH:MM\_HH:MM-HH:MM** para indicar mas un rango horario en el mismo rango de días

- En caso de que se necesite cubrir más de una franja horaria y esta sintaxis sea insuficiente, se pueden incluir varias separadas por espacios.
- Los días se indicarán con sus iniciales en castellano: LUN, MAR, MIE, JUE, VIE, SAB y DOM

Ejemplos:

**24hs** -> "00:00-23:59"  
**Jueves 24hs** -> "JUE"  
**Jueves de 14:30 a 17 hs** -> "JUE\_\_14:30-17:00"  
**Jueves de 8 a 12 hs y de 16 a 20 hs** -> "JUE\_\_08:12-17:00\_16:00-20:00"  
**Jueves de 8 a 15 hs y Viernes de 8 a 15 hs** -> "JUE\_\_08:00-15:00 VIE\_08:00-15:00"  
**Lunes a Viernes 7:30 a 17 hs y Sábados 8 a 12 hs** -> "LUN-VIE\_\_07:30-17:00 SAB\_\_08:00-12:00"  
**Lunes a Viernes 8 a 11 y 14 a 18 hs** -> "LUN-VIE\_\_08:00-11:00\_14:00-18:00"  
**Lunes y Miercoles 8 a 11 y 14 a 18 hs** -> "LUN\_MIE\_\_08:00-11:00\_14:00-18:00"  
**Lunes a Miercoles y Viernes 8 a 11 y 14 a 18 hs** -> "LUN-MIE\_VIE\_\_08:00-11:00\_14:00-18:00"  
**Lunes a Miercoles 8 a 11 y de Viernes a Domingo 9 a 10** -> "LUN-MIE\_\_08:00-11:00 VIE-DOM\_\_09:00-10:00"

## Booleano

- A menos que se indique lo contrario, se identificarán con los valores true o false.

- Esta convención puede variar en algunos rubros específicos de datos, pero en caso de no existir una convención clara y definida aplicable al rubro o contexto del dataset, se recomienda utilizar true o false.
- Este campo puede contener “valores ausentes”, en cuyo caso el campo deberá estar totalmente vacío, no conteniendo ningún carácter.
- Si existe la posibilidad de que haya otro valor que no sea true, false o “valor ausente” significa que se eligió un tipo de datos incorrecto: este no es booleano, el tipo de dato booleano es binario y sólo admite 2 valores de verdad (aparte del caso del “valor ausente”).