

PDF y datos variables (los estándares PDF/VT y PDF/VCR)

Gustavo Sánchez Muñoz

(Octubre de 2024)

Los datos variables son aquellos que varían de un ejemplar impreso a otro mientras que lo que los rodea permanece igual. Su uso típico, por ejemplo, es la preparación de miles de cartas comerciales cada una de ellas con el nombre propio y la dirección de un usuario diferente.

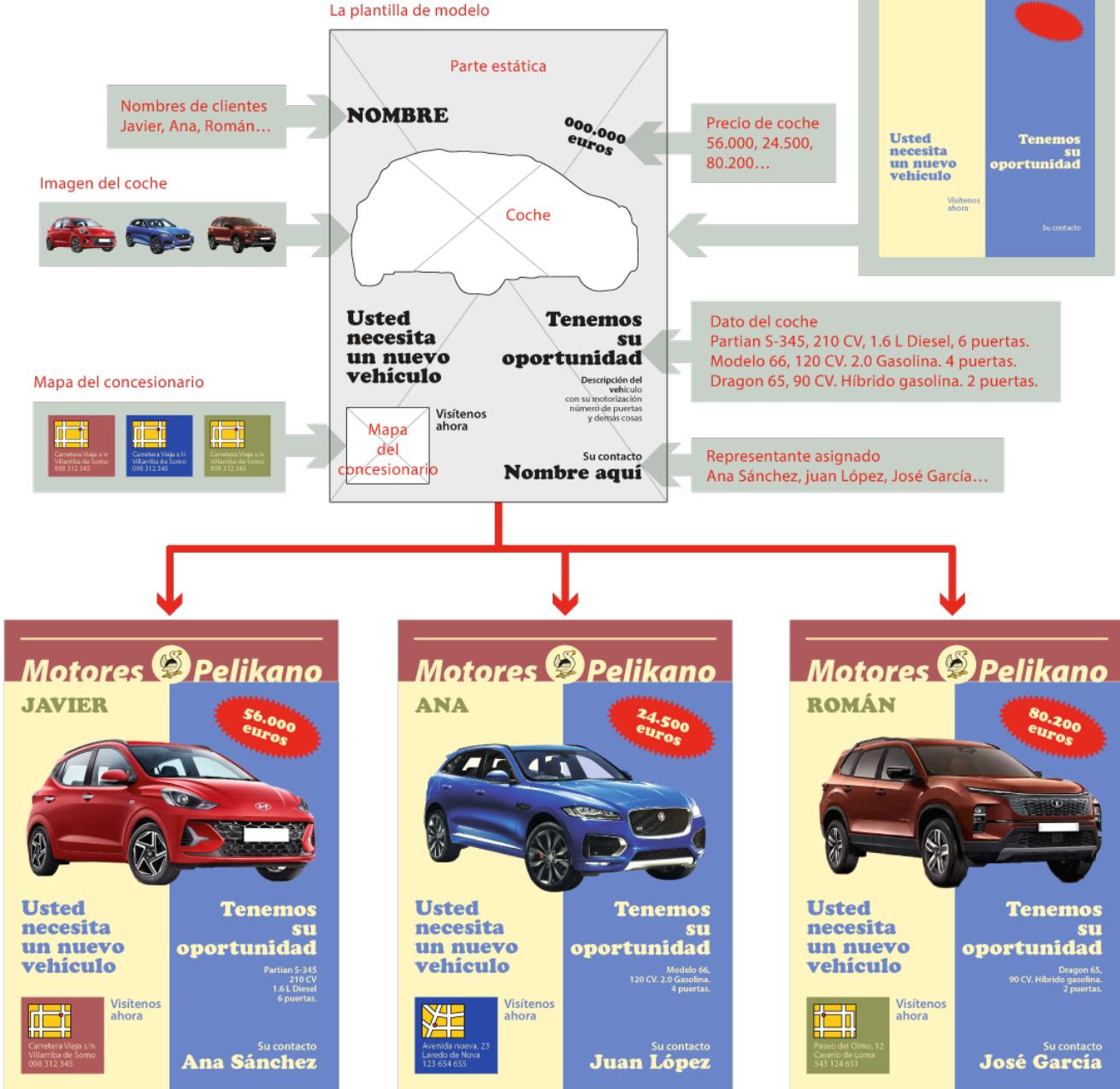
El empleo del formato PDF mejora notablemente la impresión de datos variables ya que permite crear documentos a color de gran calidad de presentación fácilmente accesibles o imprimibles.

Además, la estandarización del uso de datos variables con PDF, un formato públicamente conocido y consolidado, facilita la interacción de maquinaria y programas de distintos fabricantes que hasta ahora usaban sus formatos privados.

PDF con datos variables

PDF/VT, PDF/VCR ...

Elementos estáticos,
iguales en todos los ejemplares



Arriba vemos un ejemplo de las posibilidades del uso de datos variables con formato PDF: Un distribuidor de vehículos combina sus bases de datos de posibles clientes y vehículos disponibles para crear anuncios de oferta personalizados en los que se ofrece un vehículo concreto a un cliente y se le propone una cita en un centro concreto con un representante asignado.

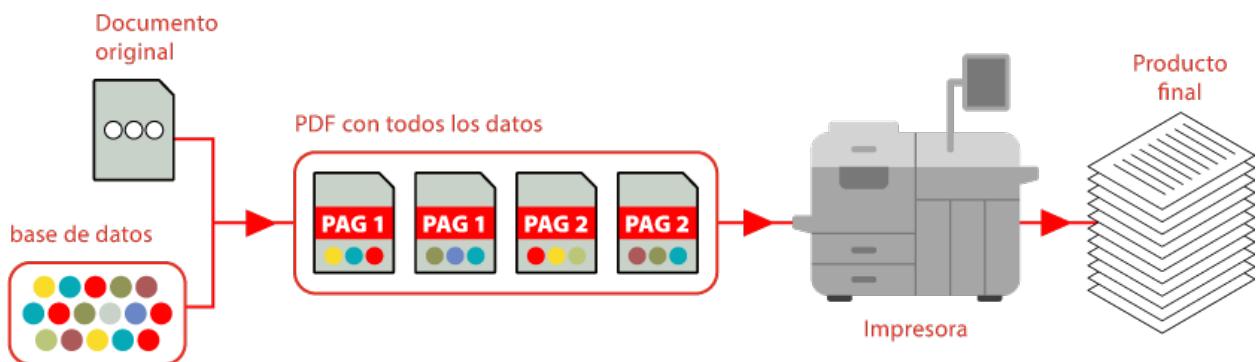
Para crear estos documentos se necesita un programa capaz de combinar todos los elementos. Una vez creado el PDF, sólo hace falta una impresora digital de calidad suficiente para imprimir y ensobrar cada documento. La ubicuidad y popularidad del formato PDF facilita esto último enormemente.

Modos de combinar PDF y datos variables

Hay dos aproximaciones principales en el uso de datos variables con estándares de PDF. Cada una de ellas tiene sus ventajas e inconvenientes. En ambas, los datos variables se tratan o identifican de modo especial. No son meros datos estáticos, como las partes del documento que no cambian de ejemplar a ejemplar.

- **Incluir los datos variables en el documento**

En este caso, el contenido final va incluido íntegramente en el PDF. La fusión de los datos variables y las partes estáticas se produce al crear el PDF. Las mismas páginas con distintos datos van incluidas las veces necesarias para la tirada y cada parte se identifica ante la maquinaria con los metadatos adecuados para que ésta pueda manipularla correctamente.



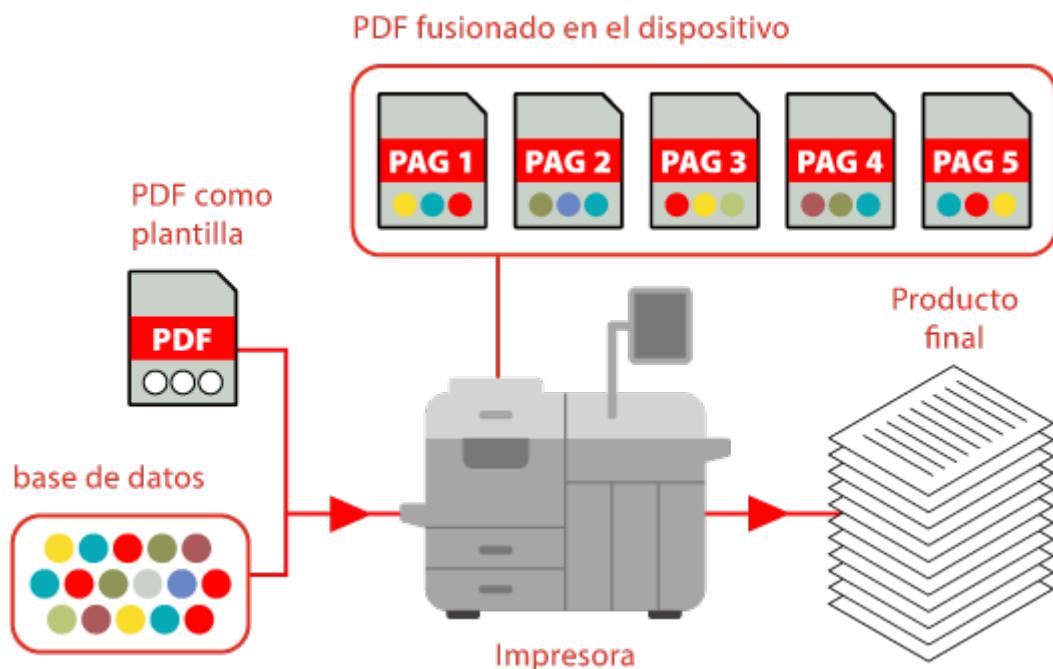
Este tipo de PDF, que incluye todo el contenido, no es un PDF estático, ya que los datos variables, identificados con metadatos, se pueden extraer y cambiar.

Una ventaja de este planteamiento es que el PDF resultante se puede imprimir en cualquier dispositivo, aunque los preparados para este tipo de metadatos agilizan el rasterizado y facilitan mucho la manipulación automatizada de muchos documentos.

En los estándares, esta aproximación se identifica con PDF/VT (donde todo el documento PDF debe estar generado por completo antes de comenzar su rasterizado, incluso en PDF/VT-2).

- **Llamar a los datos externos al crear el documento**

Los datos variables se fusionan con el resto del diseño usando el formato PDF en el momento de su impresión. El resultado se rasteriza en el dispositivo pero no se guarda.



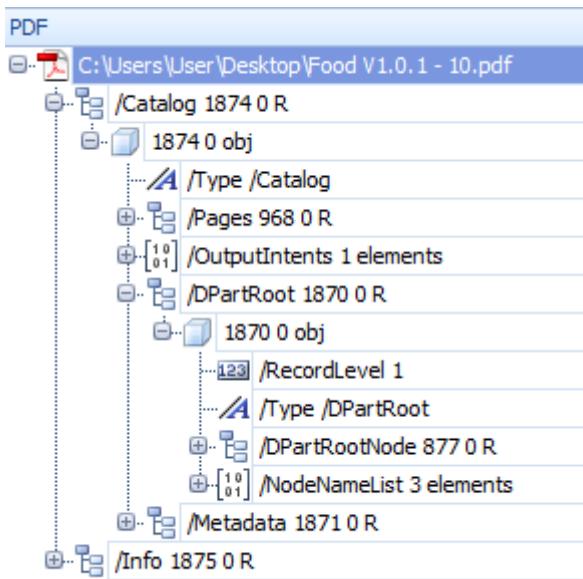
Cada documento es una especie de plantilla. Sólo incluye las páginas necesarias para un ejemplar y los datos se van inyectando en su ubicación aplicando el formato especificado al rasterizarse en el dispositivo. Los datos variables y todos los parámetros necesarios para la operación se identifican mediante metadatos.

En los estándares, esta aproximación se identifica con PDF/VCR.

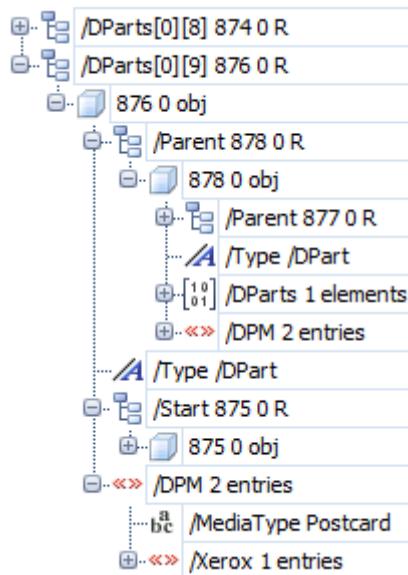
Los DPart metadata

Las bases del uso de los datos variables son unas estructuras de metadatos llamadas "metadatos de partes de documentos" (*Document Part Metadata*), es decir: datos referentes a páginas y grupos de páginas.

Estos metadatos se organizan jerárquicamente en un árbol de páginas o conjuntos de páginas, que forman partes (*DParts: Data Parts*) pueden llevar sus metadatos (*DPM: Document Part Metadata*).



Estos metadatos no se refieren sólo a contenidos (como diferentes nombres propios o direcciones postales, como en las fusiones de correo tradicionales), también se usan para indicar los tipos de reproducción de esos grupos de páginas (tipo de papel, formato, etc.).



En estas estructuras hay tres tipos de elementos:

- “DPartRoot”: Un diccionario raíz llamado “DPartRoot”, de la que dependen los demás nodos y que debe ir referenciado directamente en el diccionario “Catalog”.
- “dparts”: Un conjunto de una o más páginas del documento. Lleva identificadores de comienzo y final. Debe haber al menos uno (puede haber más).

- “`dpm`”: Un diccionario que contiene los metadatos en sí (los valores). Debe haber al menos uno, asociado con un “`dpart`”.

PDF/VT

PDF/VT es el primer estándar (creado por la ISO en 2010) para la reproducción de datos variables con PDF para imprenta. Su uso original está pensado para la impresión de datos variables transaccionales (de ahí las siglas VT: *Variable and Transactional Printing*); es decir: Transacciones comerciales como facturas, albaranes, ofertas y correo personalizado, etc.

Es un subestándar de PDF/X. Su uso no excluye que un documento sea PDF/X (de hecho, *debe* ser al menos PDF/X-4).

PDF/VT está pensado especialmente para su uso con documentos de definición de trabajo JDF con los sistemas y métodos desarrollados por la organización CIP4 para la automatización de las artes gráficas impresas.

El uso de las versiones más avanzadas (es decir, superiores a la versión 2.5) del motor de impresión de PDF de Adobe ([Adobe PDF Print Engine](#)) es lo que permite imprimir grandes volúmenes de este tipo de documentos PDF/VT de forma muy rápida y fiable, ya que el rasterizado final de los datos (en documentos adecuadamente preparados) es muy eficiente.

PDF/VT Tiene de momento tres variantes:

- **PDF/VT-1** y **PDF/VT-2**: Ambas están definidas con el mismo estándar (ISO 16612-2:2010). La versión 1 se refiere al uso con PDF/X-4 y la 2 con PDF/4p, PDF/X-5g y PDF/X-5pg. Ambas se basan en el nivel 1.6 del formato PDF.

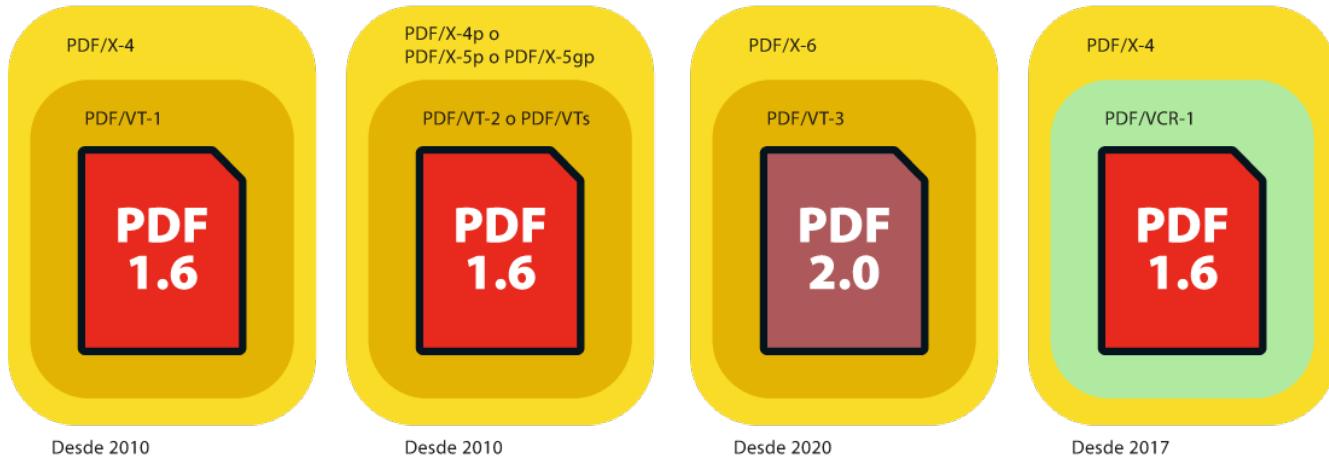
Con PDF/VT-1 los datos estáticos (diseño) y variables van juntos dentro del mismo documento. Todo es un único archivo que contiene todo lo necesario para su reproducción (no hay llamada a datos externos de ningún tipo).

En PDF/VT-2, se admite la referencia a objetos externos (perfiles de color y datos). De hecho, tiene una subvariante PDF/VT-2s que admite la llamada por paquetes a flujos de datos (*streams*).

Todo ello está pensado para facilitar el uso de conjuntos de datos grandes, ya sea en conjunto o por llamadas sucesivas.

- **PDF/VT-3:** Está definida para el uso de datos variables con PDF/X-6 (ISO 16612-3:2020) y se basa en el uso del nivel 2.0 del formato PDF (ISO 32000-2:2020), que ya incluye como parte integral de ese formato los DPart Metadata (cosa que en el nivel 1.6 no ocurría).

Datos variables y estándares ISO de PDF



PDF/VCR

En 2017, la ISO creó el estándar PDF/VCR. para la reproducción con el Reemplazo de Contenido Variable (*Variable Content Replacement: VCR*).

En PDF/VCR-1 se crean plantillas de una o varias páginas con contenido fijo y marcadores de posición de contenido variable. Los datos variables se proporcionan al sistema mediante un formato de valores separados por comas (CSV) en el momento de la impresión que ensambla el PDF en ese momento, sin guardar en principio copia del mismo.

Números de serie de documentos de pago



Esto es muy útil en sistemas donde hace falta calidad de impresión (proporcionada por el formato PDF) pero no se quiere almacenar lo impreso; por ejemplo: Tarjetas de crédito con distintos números y códigos PIN, envases farmacéuticos con números de lote y serie generados a partir de sensores, números de series de billetes, etc.

Estándares que los definen

PDF/VT-1, PDF/VT-2 y PDF/VT-2s se especifican en el estándar [ISO 16612-2:2010](#).

PDF/VCR-1 se especifica en el estándar [ISO 16613-1:2017](#).

PDF/VT-3 se especifica en el estándar [ISO 16612-3:2020](#).