Ink Document Metadata Format v2.0

Editor's note

The overview has been modified to align the terminology with the WILL Program DevDocs. In DevDocs the page is copied to: Documents SDK - Ink Document Metadata Format

https://developer-docs.wacom.com/display/DevDocs/Documents+SDK+-+Ink+Document+Metadata+Format

1. Overview

The Ink Document Metadata Format defines active areas within a PDF document by the use of Extensible Metadata Platform (XMP) packets with a custom schema. Each active area within a document has a matching AcroForm object within the PDF. An XMP packet embedded in the metadata stream of the matching AcroFormObject stores the data describing the behaviour and features of the active area. An XMP packet embedded in the metadata stream for the PDF page object describes the features such as barcode ID for each individual page. An XMP packet embedded in the metadata stream of the root PDF object stores the metadata related to the document as a whole.

Throughout the libraries and documentation there are references to 'Barbera' and 'Baxter' for the following reason: The Baxter SDKs were initially created to support CLB-Create and CLB-Paper applications developed for the Wacom Clipboard PHU-111, codenamed Barbera.

The source of the name Baxter is 'Barbera Extensible Translator'.

The SDKs can be used universally to extend the contents of PDF documents, as described in the Ink Document Metadata Format (aka Barbera File Format).

- 1. Overview
 - 1.1 Schema Overview
 - 1.2 Definitions
 - 1.2.1 PDF
 - 1.2.2 XMP
 - 1.2.3 AcroForm
 - 1.2.4 Smart Pad
 - 1.2.5 Pen data
 - 1.2.6 HWR (Hand writing recognition)
 - 1.2.7 Centripetal Catmull-Rom Spline
- 2. Scope
- 3. Common Data Type Definitions
 - 3.1 XMP Packet Wrapper
 - 3.2 Pen Data
 - 3.2.1 Stroke
 - 3.2.1.1 Point
 - 3.3 Compressed Pen Data
 - 3.3.1 Compressed Stroke
 - 3.3.2 Start Time
 - 3.3.4 Is Erased flag
 - 3.3.3 Point Time Offset
 - 3.3.5 Compressed Pen Data example
 - 3.4 WGSS Packet Type Marker
- 4. Document Level Definitions
 - 4.1 Authoring Tool Version
 - 4.2 Page ID List
 - 4.3 Page Completion Order
 - 4.4 Document Completion Time
 - 4.5 Smart Pad Device ID
 - 4.6 Smart Pad Device Characteristics
 - 4.7 Client App
 - 4.8 Client Device
 - 4.9 Full document level XMP sample
 - 4.10 InkDoc Meta
- 5. Page Level Definitions Section
 - 5.1 Page ID
 - 5.2 Field IDs for Page
 - 5.3 Pen Data

- 5.4 Full Page Level XMP Sample
- 6. Object Level Definitions Section
 - 6.1 Common attributes for object level fields
 - 6.1.1 Field UUID
 - 6.1.2 Field Location
 - 6.1.3 Field Type
 - 6.1.3.1 Text
 - 6.1.3.2 Number
 - 6.1.3.3 Boolean
 - 6.1.4 Tag
 - 6.1.5 Required
 - 6.1.6 Pen Data
 - 6.1.7 Completion Time
 - 6.1.8 Data
 - 6.2 Object level XMP sample packet
- 7. NHS Consent form worked sample
 - 7.1 XMP packets before data capture
 - 7.1.1 Document level XMP packet
 - 7.1.2 Page Level XMP packet
 - 7.1.3 Object Level XMP packets.
 - 7.1.3.1 Center number (Number)
 - 7.1.3.2 Study Number (Number)
 - 7.1.3.3 Patient ID number for trial (Number)
 - 7.1.3.4 Project Title (Text)
 - 7.1.3.5 Researcher Name (Text)
 - 7.1.3.6 Information Sheet Date (Text)
 - 7.1.3.7 Information Sheet Version (Text)
 - 7.1.3.8 Consent Checkbox 1 Consent (Boolean)
 - 7.1.3.9 Consent Checkbox 2 Provision for withdrawal (Boolean)
 - 7.1.3.10 Company Name (Text)
 - 7.1.3.11 Consent Checkbox 3 Consent for data to be passed to company (Boolean)
 - 7.1.3.12 Consent Checkbox 4 I Agree for GP access (Boolean)
 - 7.1.3.13 Consent Checkbox 5 I Agree to take part (Boolean)
 - 7.1.3.14 Participant Name (Text)
 - 7.1.3.15 Date 1 (Text)
 - 7.1.3.16 Participant signature (Signature)
 - 7.1.3.17 Name of person (Text)
 - 7.1.3.18 Date (Text)
 - 7.1.3.19 Person Signature (Text)

1.1 Schema Overview

The XML namespace for the Barbera file format is http://wacomgss.com/barbera/1.0/. XMP packets define the namespace in this way:

The file format has three types of XMP packet:

- Document Level Packet: Contains metadata covering the whole document
- Page Level Packet: Contains metadata for a particular page
- Object Level Packet: Contains metadata for a particular object

Before data capture, each XMP packet defines the characteristics and type of a particular area. After data capture, the XMP packet stores both raw and converted data alongside the area characteristics. The schemas outline below are placed between rdf:Description el ements of the XMP packet.

1.2 Definitions

This section gives a high level overview of the terms used and the underlying technology.

1.2.1 PDF

PDF or 'Portable Document Format' is a file format used to present documents in a program, software and hardware independent way, as defined by the ISO 32000–1 specification. The PDF format is a subset of COS (Carousel object structure). A COS tree mainly consists of objects, of where there are 8 types:

- Boolean (true and false)
- Number
- String (delimited in parentheses, may contain 8 bit chars)
- Name (starting with a '/')
- Array (ordered collection of objects delimited in square brackets)
- Dictionary (Collections of objects indexed by Names, stored in << >> pairs)
- Stream (Usually large amounts of data either compressed or binary)
- Null

Objects can be direct (embedded within another object) or indirect. Indirect objects have an object number and a generation number (defined between obj and endobj keywords). An index table (xref keyword) follows the main body of the data and gives the byte offset of each indirect object from the start of the file. At the end of the pdf the file trailer (trailer keyword) contains a dictionary, an offset to the start of the index table and the %%EOF (end of file marker). The dictionary contains a reference to the root object of the tree structure, and the count of indirect objects in the index table. This dictionary is also used to store metadata for the document.

1.2.2 XMP

XMP or 'Extensible Metadata Platform' is an ISO standard for the creation, processing and interchange of custom metadata for digital documents and data sets. PDFs support XMP packets embedded within metadata streams of COS objects, allowing the data definitions for the barbera file format to be part of the same file used to print the form itself. Using XMP in this way allows the flexibility to store the barbera metadata and captured pen data in the digitally signed document, readable on standard a PDF readers.

1.2.3 AcroForm

An AcroForm (sometimes called an Acrobat form) is a collection of active form objects such as a checkbox, text input field or radio button added to a PDF to create an interactive form. The Barbera file format uses AcroForms to bind active areas to underlying PDF objects.

1.2.4 Smart Pad

A smart pad refers to the Wacom device that captures pen data from an inking pen, with the intention of recording a digital copy of the real ink, drawn by the pen on paper. This includes the Barbera and Columbia devices.

1.2.5 Pen data

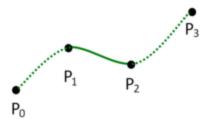
For the purposes of this document, pen data refers to the digital version of the data captured directly from the smart pad device. The pen data is a collection of strokes, with each stroke containing a collection of points. Each point contains x, y, pressure and time data. Each stroke also contains the ink colour for that particular stroke, and the ink thickness for the stroke.

1.2.6 HWR (Hand writing recognition)

For certain active areas, it is expected that the barbera client application will perform automatic hand writing recognition (HWR). This translated data will be stored in the the field level XMP data. The appearance stream for the active area will be set to show the stroke data so that the visual appearance of the digital and analogue documents is the same.

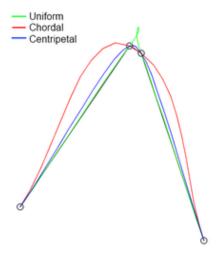
1.2.7 Centripetal Catmull-Rom Spline

A Centripetal Catmull-Rom spline is a type of interpolating spline (a curve that goes through its control points), defined by four control points (P0 to P3) where the curve is only drawn between P1 and P2:



(Catmull-Rom spline interpolation with four points)

Centripetal Catmull–Rom spline has several desirable mathematical properties compared to the original and the other types of Catmull-Rom formulation. First, it will not form loop or self-intersection within a curve segment. Second, cusp will never occur within a curve segment. Third, it follows the control points more tightly:



In this figure, there is a self-intersection/loop on the uniform Catmull-Rom spline (green), whereas for chordal Catmull-Rom spline (red), the curve does not follow tightly through the control points.

2. Scope

This document defines the Barbera file format. The format is a collection of XMP packets embedded within metadata streams of a PDF document. The three types of XMP packet (Document Level, Page Level and Object Level). Sections 3, 4 and 5 of this document define the schema for each packet, and section 7 contains a full document example.

3. Common Data Type Definitions

This section covers the XMP chunks that are common between sections, including:

- XMP Preamble
- Pen Data / Compressed Pen Data
- WGSS Packet Type Marker

3.1 XMP Packet Wrapper

Every XMP has a pair of XML processing instructions, that contain both the rdf:RDF element and the rest of the packet data. The header XML processing instruction contains a UUID and the unicode character U+FEFF used as a byte-order marker (shows the 0xFEFF character):

If the XMP block is writeable, the end attribute of the trailer should be 'w', for a read only block this should be 'r'.

3.2 Pen Data

The pen data sub-packet records input data received from the smart pad device. This is a collection of strokes, each of which is a collection of points that represent control points on a catmull-rom spline and a width at that point:

```
<wqss:PenData>
         <rdf:Seq>
            <rdf:li rdf:parseType="Resource">
               <wqss:Stroke>
                  <rdf:Seq>
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="1" wgss:y="2"
wgss:w="3" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="2" wgss:y="3"
wgss:w="3" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="3" wgss:y="4"
wgss:w="3" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="4" wgss:y="5"
wgss:w="4" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="5" wgss:y="6"
wqss:w="4"/>
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="6" wgss:y="7"
wgss:w="5" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="7" wgss:y="8"
wgss:w="6" />
                  </rdf:Seq>
               </wqss:Stroke>
            </rdf:li>
            <rdf:li rdf:parseType="Resource">
               <wqss:Stroke>
                  <rdf:Seq>
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="4" wgss:y="5"
wgss:w="4" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="5" wgss:y="6"
wgss:w="4" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="6" wgss:y="7"
wgss:w="5" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="7" wgss:y="8"
wgss:w="6" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="1" wgss:y="2"
wgss:w="3" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="2" wgss:y="3"
wgss:w="3" />
                     <rdf:li wgss:inkColor="#FFFFFF" wgss:x="3" wgss:y="4"
wqss:w="3" />
                  </rdf:Seq>
               </wgss:Stroke>
            </rdf:li>
         </rdf:Seq>
      </wgss:PenData>
```

3.2.1 Stroke

A stroke is an RDF sequence of points, the color of the stroke is defined in the first point in the sequence. If the inkColor attribute is present in any point other than the first of the sequence, it will be ignored.

3.2.1.1 Point

Each point is an RDF list item, representing a control point on a catmull-rom spline, along with the thickness of the stroke at that point. It will also contain the inkColor.

```
<rdf:li wgss:inkColor="#FFFFFF" wgss:x="1" wgss:y="2" wgss:w="3" />
```

The point element has the following attributes describing the control point:

| Attribute | Detail |
|-----------|---|
| X | The x coordinate of the point in device units |
| у | The y coordinate of the point in device units |
| W | The width of the stroke at this point |
| inkColor | The Color of the point (and or stroke) |

X,y,w are expressed in device units.

3.3 Compressed Pen Data

Sometimes it will be preferable to store the pen data in a compressed format in order to reduce the size of the meta data streams that have bee added to the document. In these instances, any 'Pen Data' sections within the file format can be substituted with a 'Compressed Pen Data' section. A compressed pen data section stores the stroke data as serialised, compressed base64. Serialisation and compression/decompression is handled by BaXter itself so data is presented uncompressed to developers.

For the purposes of encoding this point data into the meta data stream, the byte data is simply base64 encoded, and placed in the appropriate XMP tag.

3.3.1 Compressed Stroke

The compressedStroke list item contains the entire collection of Points within the stroke, compressed & encoded. Serialisation and compression/decompression is handled by BaXter itself.

3.3.2 Start Time

Strokes can also contain timing information. The startTime property contains the time the stroke began. In BaXter this is interfaced with a DateTime object, and serialised as a long.

3.3.3 Point Time Offset

To provide context to each Point's 't' value, the Point Time Offset can be set. It is an enum as follows:

| Nanosecond | 0 |
|-------------|---|
| Millisecond | 1 |
| Second | 2 |

3.3.4 Is Erased flag

To provide support for storing erased strokes (in case of document timing playback being required) you can mark a Stroke as erased. This is encoded as a boolean in the XMP.

3.3.5 Compressed Pen Data example

```
<wgss:CompressedPenData>
 <rdf:Seq>
  <rdf:li
wgss:compressedStroke="eNqdlTtuJDEMRA+zqQJSEn9w7vsfaat6gAWk5iTryJgHfkRWsVM
8RuTysacOGX9+n7/xk/+A+e5BSPYqb5BCUCNLT1BFsEb5CXQy1bShonESVxDdIHMdZMoCkdkQQ
0yg9oss2SARDWGd8PUmHFTEzoYks62mA9t4T2hTxzGriK5r5zg9mxgmCo+mtzCSrutIErM3SSH
Z1ZApJF3Meoj+J1G76mCivmLUtWxI6gMgkuY1K4drnQBiBaix7UwV+O2pvua+hjkJ5pjrVK1zm
773mHI6w41dbR+XaH0+q4xhZyKr+qz4/j1YANq8DGbGQRkKr/MNNpkID1vrinjeAMVuPye769E
DWlW/pPpICIVs9QD/9aD0Bos6pYm/ABU5q5sIRd8Y0tBncCVvsgxkNsa3x6qrkallfbGqU9oUy
5sYTZyNhZzZUpo6bAvnqjpDwt6wf2dIkuhsx22WfLNddQeG1zXKmpfmZky22RSnWjurwqApuOT
vrosx3mwunCQ7MpmtuuOX+FioNjG8OiDVHcwAmc1+LBgz2/NbzNYpBBNNlS4biuM9XQz2gxm0M
ZwB7kjW5Vx8sJJnz8MuS3M7td+WZmOCz5P6keovZsCiCg=="
wgss:startTime="636608694324516764" wgss:pointTimeOffset="1"
wqss:isErased="False"/>
  <rdf:li
wgss:compressedStroke="eNp1k01uxTAIhA/TrRdgsBnUfe9/pA5pVdV+vCwiJZ+AGX6QoSM
CORw2ZHx8Pc/4xB/YqieAEESMnbsHYbqA/8XmC94DzHwD9nwDcIHYBTCQ3oOUdYKVJTdH6iXX9
4/ztKsGvwMyR3qcQBkBpslbrrAlMB+14j8AFsHSUYYOEFWDBvPygV2psAlOuViMSMFQYbWDOEh
Mi2hDFnVlXsmcIKrKusrTYmbpulJxphCNFycpSWBMhavxLgSP92uBuDuQTZCHkZQ5C7x4T+GSE
jgdnlua+oSsIJ12k12600MXMrWymTTES51qQ6Iall1MGsluFJgU8dkRrkVqG1Pjb+vUSWBbQ56
TM3R+tBa2i1H6oYSmb0Ddi3UdfU6MczM/gdeJcQgh59zUKmJh2LqA/16rxrGY31Db/v8="
wgss:startTime="636608694324516764" wgss:pointTimeOffset="1"
wgss:isErased="False"/>
  <rdf:li
wgss:compressedStroke="eNqV1Tty5DAMBNDDOGUAECA+5dz3P5IbCrZqZlq7K6WvRIJAU+r
YujK613ksWV8/170++w94vIFdYOv04ZDir+ACqLOyhUOFcmjNN3BAxurTN1A3oKJ6J9v+V+yST
WSfESeiNlKfcjrnmOSdkz29VCKB1qQIEdcRts/UlkrOc3RDNqtAZjXTJ+I9tTnZx6sgIeydWY3
1+m+CKWQ7k1hZQgUVlN6dtHgPsE8ZFUyu3FhHZ7WzmWAKk3Qyn1kNI/8UGykm3nOfWA6wRbayC
hJJbIThUxDCbGMJmV5jfiTXquzqTSbodUc+EsQzu5TdrKm6hd3GOWn3I/G9aj4FjyQqW55JQuy
Z2FTAphDaIywHPUKTiGlDmn8pSg7N9ZyHzdRzaguSKrepINidk0vIeaymtiD7GD4SEFK1XVVT2
fUvaXn9ixkuIyBX1A28/xBtDySimy97/AKbKoo2"
wgss:startTime="636608694324516764" wgss:pointTimeOffset="1"
wgss:isErased="False"/>
  <rdf:li
wgss:compressedStroke="eNqVlDFyxTAIRA+TVgUIIWDS5/5HyiJX91/PJL/0G2BZVr9qyoj
YNeasIePr5/zGd5XNBjpmxh2sBeA2LPIOPAFAV+47iJ5hMbxeQIjeQRoAFEUJB7k3BzUfqqr3m
DoqXoCK1huZ9leSvbzOf5HAgiHFSBsjSYjvJkGIHbIJ0dPNP0leChg5V6ZE8kUB4tKbCiHeCnQ
RcmKmZNM4ycTHT6LRsXFWY1cEyZwmz1Sv9iBYt1CQIo5GoSaVXC57Ti5y7dxds8mcakcz2bUXN
\verb|i3m2817rc8kqgjch0piaOFCyBBDeFAp5BBAeDkpRhFkpDhtmAq0eUMHiqAyumGuF/FAdK/ZqKg| \\
bMnkVHtkZZcx2aX30VKuXWuy8bZKzw1vXWLCn1D7MOerxNxdWDRAY1N6j79eVzOuR/NMKbscN/
AKBbj9I" wgss:startTime="636608694324516764" wgss:pointTimeOffset="1"
wgss:isErased="False"/>
 </rdf:Seq>
</wgss:CompressedPenData>
```

3.4 WGSS Packet Type Marker

```
<wgss:PacketType wgss:level="document"/>
```

| level | Detail |
|----------|---|
| document | The contents of this XMP packet defines Document Level data |
| page | The contents of this XMP packet defines Page Level data |
| field | The contents of this XMP packet defines Field Level data |

4. Document Level Definitions

This section defines features specified in the document level XMP packet, including:

- Authoring tool version
- Page ID list
- Page ID order
- Document completion time
- Smart pad device ID
- Smart pad device characteristics
- Client application ID
- Client device ID

4.1 Authoring Tool Version

This element records the version number and name of the authoring tool used to create the file:

```
<wgss:AuthoringTool wgss:version="1.0" wgss:toolname="Barbera Authoring
Library"/>
```

The 'version' attribute denotes the version number of the client, and the tool name attribute is the human readable name for the authoring tool used.

4.2 Page ID List

This element lists the page IDs (barcodes) for each active page in the document, and which PDF page the page ID relates to:

| Attribute | Detail |
|-----------|---|
| pdfPage | The PDF page that the ID relates to |
| uuid | The barcode ID for the particular page in the PDF |

4.3 Page Completion Order

The page order element lists the order in which the PDF pages should be completed:

```
<wgss:PageCompletionOrder>
  <rdf:Seq>
    <rdf:li>1</rdf:li>
    <rdf:li>2</rdf:li>
    <rdf:li>3</rdf:li>
  </rdf:Seq>
</wgss:PageCompletionOrder>
```

In the example above the pages are completed 1, 2 then 3. This is an optional element.

4.4 Document Completion Time

This element is set to the time when a document is sealed, and marks the time the document was completed:

```
<wgss:DocumentCompletionTime>2016-06-17T13:32:45.5316112Z</wgss:DocumentCo
mpletionTime>
```

The time is specified in ISO-8601 format ("yyyy-MM-dd'T'HH:mm:ss.fffffffZ"") in UTC.

4.5 Smart Pad Device ID

The'SmartPad' element stores the ID of the last smart pad device used to complete the document:

```
<wgss:SmartPad wgss:id="1234567890" wgss:deviceName="Joss' Barbera
Device"/>
```

| Attribute | Detail |
|------------|--|
| id | The device ID of the smart pad used |
| deviceName | The human readable name of the smart pad device (as configured by the user). |

4.6 Smart Pad Device Characteristics

The SmartPadCharacteristics' element stores details about the smart pad used to capture data:

```
<wgss:SmartPadCharacteristics wgss:unit="inch" wgss:pointsPerUnit="200"
wgss:width="10400" wgss:height="10400"/>
```

| Attribute | Detail |
|-----------|--------|
| | |

| unit | The unit of measurement for the characteristics (e.g. meter, inch) |
|---------------|--|
| pointsPerUnit | The number of devices points per unit / the density of the sensor |
| width | The width of the device sensor in device points |
| height | The height of the device sensor in device points |

4.7 Client App

The 'ClientApp' element stores information about the application that was used to capture data from the smart pad device:

```
<wgss:ClientApp wgss:version="1.0" wgss:os="iOS"
wgss:applicationName="barbera for iOS"/>
```

| Attribute | Detail |
|-----------------|---------------------------------------|
| version | The client application version number |
| os | The operating system for the client |
| applicationName | The name of the client application |

4.8 Client Device

The 'ClientDevice' element stores information about the device that the client application is running on:

```
<wgss:ClientDevice wgss:id="1234567890" wgss:deviceClass="iPad"
wgss:deviceName:"Joss' iPad Pro"/>
```

| Attribute | Detail |
|-------------|--|
| id | The device ID as returned by the host OS |
| deviceClass | The type of device (e.g. iPhone) |
| deviceName | The human readable name of the device as returned by the host operating system |

4.9 Full document level XMP sample

This is the full document level XMP packet for a 3 page document that was completed on an iPad pro:

```
<?xml version="1.0"?>
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"</pre>
xmlns:xmp="http://ns.adobe.com/xap/1.0/"
xmlns:wqss="http://wacomqss.com/barbera/1.0/">
 <rdf:Description rdf:about="http://signature.wacom.eu">
  <wgss:PacketType wgss:level="document"/>
  <wqss:InkDoc wqss:version="2"</pre>
wgss:licence="eyJhbGci0iJSUzUxMiJ9.eyJpc3Mi0iJsbXMiLCJhdWQi0iJ3ZXM"/>
  <wgss:PageIDList<wgss:PageIDList>
         <rdf:Bag>
            <rdf:li wgss:pdfPage="1" wgss:uuid="1234567" />
   <rdf:li wgss:pdfPage="3" wgss:uuid="1234568" />
            <rdf:li wgss:pdfPage="2" wgss:uuid="1234569" />
         </rdf:Baq>
  </wqss:PageIDList>
  <wgss:PageCompletionOrder>
   <rdf:Seq>
    <rdf:li>1</rdf:li>
    <rdf:li>2</rdf:li>
    <rdf:li>3</rdf:li>
   </rdf:Seq>
  </wgss:PageCompletionOrder>
  <wgss:AuthoringTool wgss:version="1.0" wgss:toolname="Barbera Authoring</pre>
Library"/>
<wgss:DocumentCompletionTime>2016-06-17T13:32:45.5316112Z</wgss:DocumentCo</pre>
mpletionTime>
  <wqss:SmartPad wqss:id="1234567890" wqss:deviceName="Joss Barbera</pre>
Device"/>
  <wgss:SmartPadCharacteristics wgss:unit="inch" wgss:pointsPerUnit="200"</pre>
wgss:width="10400" wgss:height="10400"/>
  <wgss:ClientApp wgss:version="1.0" wgss:os="iOS"</pre>
wgss:applicationName="barbera for iOS"/>
  <wgss:ClientDevice wgss:id="1234567890" wgss:deviceClass="iPad"</pre>
wgss:deviceName="Joss' iPad Pro"/>
 </rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

4.10 InkDoc Meta

This field is used to store the license information of the library which generated the file and the version number of the format contained within. The full license string is stored for auditing purposes and the version information to defend against breaking changes in the interface

An exception will be thrown by the library if it does not support the opened IDMF version.

```
<wgss:InkDoc wgss:version="2"
wgss:licence="eyJhbGciOiJSUzUxMiJ9.eyJpc3MiOiJsbXMiLCJhdWQiOiJ3ZXM"/>
```

5. Page Level Definitions Section

This section defines the fields that can be set in a page level XMP packet, including:

- Page ID
- Field ID list
- Pen Data

These settings relate to a particular page within a document. The XMP packet is embedded into the metadata stream for the PDF page that it relates to.

5.1 Page ID

This element identifies the XMP packet within the document (ie this packet applies for PDF Page blah with barcode blah)

```
<wgss:PageID wgss:pdfPage="1" wgss:uuid="1234567"
wgss:uuid_type="Code128"/>
```

5.2 Field IDs for Page

This element lists the field IDs for the current page:

FieldIDs correspond to the underlying AcroForm objects PDFName (T:).

5.3 Pen Data

This element contains all pen data captured outside active areas on the page:

```
<wgss:PenData>
  <rdf:Seq>
   <rdf:li rdf:parseType="Resource">
    <wgss:Stroke>
     <rdf:Seq>
      <!-- sequence of points for stroke -->
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="10" wgss:y="13" wgss:w="53"</pre>
wgss:inkColor="#FFFFFF"/>
      </rdf:li>
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="11" wgss:y="14" wgss:w="52"/>
      </rdf:li>
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="12" wgss:y="16" wgss:w="48"/>
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="13" wgss:y="14" wgss:w="10"/>
      </rdf:li>
     </rdf:Seq>
    </wgss:Stroke>
   </rdf:li>
  </rdf:Seq>
 </wgss:PenData>
```

The format is defined in section 3.2 'Pen Data'.

5.4 Full Page Level XMP Sample

Here is a full sample of a page level XMP packet:

```
<?xml version="1.0"?>
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"</pre>
xmlns:wgss="http://wacomgss.com/barbera/1.0/">
 <rdf:Description rdf:about="http://www.w3.org/">
 <wgss:PacketType wgss:level="page"/>
 <wgss:PageID wgss:pdfPage="1" wgss:uuid="1234567"/>
  <wqss:FieldIDList>
   <rdf:Bag>
    <rdf:li>142532</rdf:li>
    <rdf:li>142533</rdf:li>
    <rdf:li>142534</rdf:li>
   </rdf:Baq>
  </wgss:FieldIDList>
  <wgss:PenData>
   <rdf:Seq>
    <rdf:li rdf:parseType="Resource">
     <wgss:Stroke>
      <rdf:Seq>
       <!-- sequence of points for stroke -->
       <rdf:li rdf:parseType="Resource">
        <wgss:Point wgss:x="10" wgss:y="13" wgss:w="53"</pre>
wgss:inkColor="#FFFFFF"/>
       </rdf:li>
       <rdf:li rdf:parseType="Resource">
        <wgss:Point wgss:x="11" wgss:y="14" wgss:w="52"/>
       </rdf:li>
       <rdf:li rdf:parseType="Resource">
        <wqss:Point wqss:x="12" wqss:y="16" wqss:w="48"/>
       </rdf:li>
       <rdf:li rdf:parseType="Resource">
        <wgss:Point wgss:x="13" wgss:y="14" wgss:w="10"/>
       </rdf:li>
      </rdf:Seq>
     </wgss:Stroke>
    </rdf:li>
   </rdf:Seq>
  </wgss:PenData>
 </rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

sample defines the data for the second page of a PDF document, with a barcode of 1234569. The page has 3 active fields, and one stroke has been captured outside any active areas.

6. Object Level Definitions Section

This section defines the fields that can be set in an object level XMP packet, including:

• Field ID

- Field Location
- Field type
- Tag
- Required
- Pen Data
- Completion Time

These settings relate to a particular object within a PDF document. The XMP packet is embedded into the metadata stream for the AcroForm object that it relates to.

6.1 Common attributes for object level fields

This section lists values that are common between all field types.

6.1.1 Field UUID

The UUID element specifies the UUID for the field:

```
<wgss:FieldUUID wgss:pdfID="5A234" wgss:fieldID="12345"/>
```

| Attribute | Detail |
|-----------|--------------------------------------|
| pdfID | The Nameof the underlying PDF object |
| fieldID | Barbera Specific Field ID |

6.1.2 Field Location

The field location element specifies the location of the active area on the page:

```
<wgss:FieldLocation wgss:x="123" wgss:y="345" wgss:w="100" wgss:h="300"/>
```

The location is a rectangle defined by an origin and a size element. The values are defined in PDF coordinates, with the following attributes:

| Attribute | Detail |
|-----------|--|
| X | The X coordinate of the origin for the active area |
| у | The Y coordinate of the origin for the active area |
| W | The width for the active area |
| h | The height for the active area |

6.1.3 Field Type

The field type attribute defines the field type of the active are:

```
<wgss:FieldType>Text</wgss:FieldType>
```

The field type can be one of the following values

| Field Type | Detail |
|------------|--------|
| | |

| Text | The area is hand written text for HWR conversion |
|-----------|---|
| Number | The area contains hand written numerics for HWR |
| Boolean | The area is a check box |
| Signature | The area is a hand written signature |
| Freehand | The area is used for freehand input (not converted) |

Specific details for each active area type are listed in the following sections.

6.1.3.1 Text

The field type "Text" is used for areas in the document where a user would write freehand text that should be converted using the HWR technology, such as the customer name. This type has no additional attributes.

6.1.3.2 Number

The field type "Number" is used for areas in the document where a user would write freehand number data that should be converted using the HWR technology, such as a total price field. This type has no additional attributes.

6.1.3.3 Boolean

The field type "Boolean" is used for checkbox areas in the document. Where a user writes in a checkbox this should be converted to a true or false value. This type has no additional attributes.

6.1.3.4 Signature

The field type "Signature" defines an area that is used for signature capture. This type supports the following additional elements:

For signature data, the element contains either base64 encoded FSS data in the case of non-encrypted signature data, or encrypted base64 encoded FSS data in the case of an encrypted signature. Additionally, a signature field can add the following elements:

| Element Name | Detail |
|--------------|---|
| Name | The signatory name |
| Reason | The signatory reason |
| Encrypted | If the encrypted attribute is true, then recorded signature data should be encrypted using the public key encryption. If the signature is not encrypted, then the signature data is stored in base64 encoded FSS data (see 6.1.9 'Data' for a more detailed description of how signature data is recorded). |
| KeyName | For encypted signatures, this is the name of the public key used to encrypt |

6.1.3.5 Freehand

The field type "Freehand" is used to mark areas that are to be used for freehand input that is not converted. This type has no additional attributes.

6.1.4 Tag

The tag element gives a user configurable name for the particular field:

<wgss:FieldTag>FirstName</wgss:FieldTag>

This is used to enable post processing of a document in workflows, where the user can define IDs that have contextual information for the field.

6.1.5 Required

The required element defines if a particular area on a page must be completed before the page can be 'finished':

```
<wgss:Required>True</wgss:Required>
```

If a page has any areas that are required, then that page implicitly becomes a required page for the document, and the page must be completed before the document can be finished.

6.1.6 Pen Data

The pen data element stores the pen data collected in the active area:

```
<wgss:PenData>
  <rdf:Seq>
   <rdf:li rdf:parseType="Resource">
    <wgss:Stroke>
     <rdf:Seq>
      <!-- sequence of points for stroke -->
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="10" wgss:y="13" wgss:w="53"</pre>
wgss:inkColor="#FFFFFF"/>
      </rdf:li>
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="11" wgss:y="14" wgss:w="52"/>
      </rdf:li>
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="12" wgss:y="16" wgss:w="48"/>
      </rdf:li>
      <rdf:li rdf:parseType="Resource">
       <wgss:Point wgss:x="13" wgss:y="14" wgss:w="10"/>
      </rdf:li>
     </rdf:Seq>
    </wgss:Stroke>
   </rdf:li>
  </rdf:Seq>
 </wgss:PenData>
```

enables post-processing of any area if required. The data format is detailed in section 3.2 'Pen Data'.

6.1.7 Completion Time

The FieldCompletionTime element defines when data was last captured in a field:

```
<wgss:FieldCompletionTime>2016-06-17T13:32:45.5316112Z</wgss:FieldCompleti
onTime>
```

The time is specified in ISO-8601 format ("yyyy-MM-dd'T'HH:mm:ss.ffffffZ"") in UTC.

6.1.8 Data

The data element is used to store the translated data (if any) from the active area:

```
<wgss:Data>Joss</wgss:Data>
```

In the case of text and number fields, this is simply the data as returned by the HWR system. For signature data, the element contains either base64 encoded FSS data in the case of non-encrypted signature data, or encrypted base64 encoded FSS data in the case of an encrypted signature. Additionally, a signature field can add the following elements:

| Element Name | Detail |
|--------------|---|
| Name | The signatory name |
| Reason | The signatory reason |
| KeyName | For encypted signatures, this is the name of the public key used to encrypt |

<wgss:Name>Joss Giffard-Burley</wgss:Name>

<wgss:Reason>Reason</wgss:Reason>

<wgss:KeyName>JossSuperKey</wgss:KeyName>

6.2 Object level XMP sample packet

Here is a full object level XMP sample packet:

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"</pre>
xmlns:xmp="http://ns.adobe.com/xap/1.0/"
xmlns:wgss="http://wacomgss.com/barbera/1.0/">
 <rdf:Description rdf:about="http://www.w3.org/">
     <wgss:PacketType wgss:level="field"/>
  <wgss:FieldUUID>127001</wgss:FieldUUID>
  <wgss:FieldLocation wgss:x="123" wgss:y="345" wgss:w="100" wgss:h="300"/>
  <wgss:FieldType>Text</wgss:FieldType>
  <wgss:FieldTag>FirstName</wgss:FieldTag>
  <wgss:Required>True</wgss:Required>
  <wgss:Encrypted>True</wgss:Encrypted>
  <wqss:PenData>
   <rdf:Seq>
    <rdf:li rdf:parseType="Resource">
     <wqss:Stroke>
      <rdf:Seq>
       <!-- sequence of points for stroke -->
       <rdf:li rdf:parseType="Resource">
        <wgss:Point wgss:x="10" wgss:y="13" wgss:w="53"</pre>
wgss:inkColor="#FFFFFF"/>
       </rdf:li>
       <rdf:li rdf:parseType="Resource">
        <wgss:Point wgss:x="11" wgss:y="14" wgss:w="52"/>
       </rdf:li>
       <rdf:li rdf:parseType="Resource">
        <wgss:Point wgss:x="12" wgss:y="16" wgss:w="48"/>
       </rdf:li>
       <rdf:li rdf:parseType="Resource">
        <wgss:Point wgss:x="13" wgss:y="14" wgss:w="10"/>
       </rdf:li>
      </rdf:Seq>
     </wgss:Stroke>
    </rdf:li>
   </rdf:Seq>
  </wgss:PenData>
<wgss:FieldCompletionTime>2016-06-17T13:32:45.5316112Z</wgss:FieldCompleti</pre>
  <wqss:Data>Joss</wqss:Data>
 </rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

7. NHS Consent form worked sample

Below is a sample consent form from the NHS that we will use as a worked example. This is a 1 page PDF document with 18 active areas. The example shows the structure and format of the XMP packets required to enable the PDF document to be used with the barbera system:



NHS Cumbria Clinical Commissioning Group Lonsdale Unit, Penrith Hospital, Bridge Lane Penrith, Cumbria, CA11 8HX

| Centre Number: | | Penrith, Cumbria, CA11 8HX | | | |
|--|------------|--|--|--|--|
| Study Number: | | tel. 01768 245 486 email. enquiries@cumbriaccg.nhs.uk | | | |
| Patient Identification Number for t | his trial: | web. www.cumbriaccg.nhs.uk | | | |
| CONSENT FORM | | | | | |
| Title of Project: [PROJECT TITLE] | | | | | |
| Name of Researcher: [NAME OF CHIEF INVESTIGATOR] | | | | | |
| | | Please initial all boxes | | | |
| 1. I confirm that I have read and understand the information sheet dated [DATE] (version [VERSION NUMBER]) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. | | | | | |
| 2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. | | | | | |
| 3. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from [COMPANY NAME], from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. | | | | | |
| 4. I agree to my GP being informed of my participation in the study. | | | | | |
| 5. I agree to take part in the above study. | | | | | |
| Name of Participant | Date | Signature | | | |
| _ | | | | | |
| Name of Person | Date | Signature | | | |

The active areas of the form are (proposed field type in brackets):

- Center number (Number)
- Study Number (Number)
- Patient ID number for trial (Number)
- Project Title (Text)
- Researcher Name (Text)
- Information Sheet Date (Text)
- Information Sheet Version (Text)
- Consent Checkbox 1 Consent (Boolean)
- Consent Checkbox 2 Provision for withdrawal (Boolean)
- Company Name (Text)
- Consent Checkbox 3 Consent for data to be passed to company (Boolean)
- Consent Checkbox 4 I Agree for GP access (Boolean)
- Consent Checkbox 5 I Agree to take part (Boolean)
- Participant Name (Text)
- Date 1 (Text)
- Participant signature (Signature)
- Name of person (Text)
- Date (Text)
- Person Signature (Text)

In this example, all fields are required for the form to be considered complete.

7.1 XMP packets before data capture

This section demonstrates what the XMP packetes for the document, page and object levels will look like before capture.

7.1.1 Document level XMP packet

Every document has one and only one document level XMP packet. As this document has only one page, the document level packet is fairly basic:

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"</pre>
xmlns:xmp="http://ns.adobe.com/xap/1.0/"
xmlns:wgss="http://wacomgss.com/barbera/1.0/">
 <rdf:Description rdf:about="http://signature.wacom.eu">
 <wgss:PacketType wgss:level="document"/>
  <wgss:PageList>
   <rdf:Bag>
    <rdf:li>
     <wgss:PageID wgss:pdfPage="1" wgss:uuid="127001"/>
    </rdf:li>
   </rdf:Bag>
  </wgss:PageList>
  <wgss:PageOrder>
   <rdf:Seq>
    <rdf:li>1</rdf:li>
   </rdf:Seq>
  </wgss:PageOrder>
   <wgss:AuthoringTool wgss:version="1.0" wgss:toolname="Barbera Authoring</pre>
Library"/>
 </rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

This assumes that the barcode on the document has the value of 127001.

7.1.2 Page Level XMP packet

As this document has only one page, there is only one page level XMP packet:

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"</pre>
xmlns:xmp="http://ns.adobe.com/xap/1.0/"
xmlns:wgss="http://wacomgss.com/barbera/1.0/">
 <rdf:Description rdf:about="http://signature.wacom.eu">
 <wgss:PacketType wgss:level="page"/>
  <wgss:FieldIDList>
   <rdf:Bag>
    <rdf:li>1</rdf:li>
    <rdf:li>2</rdf:li>
    <rdf:li>3</rdf:li>
    <rdf:li>4</rdf:li>
    <rdf:li>5</rdf:li>
    <rdf:li>6</rdf:li>
    <rdf:li>7</rdf:li>
    <rdf:li>8</rdf:li>
    <rdf:li>9</rdf:li>
    <rdf:li>10</rdf:li>
    <rdf:li>11</rdf:li>
    <rdf:li>12</rdf:li>
    <rdf:li>13</rdf:li>
    <rdf:li>14</rdf:li>
    <rdf:li>15</rdf:li>
    <rdf:li>16</rdf:li>
    <rdf:li>17</rdf:li>
    <rdf:li>18</rdf:li>
    <rdf:li>19</rdf:li>
   </rdf:Bag>
  </wqss:FieldIDList>
 </rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

you can see, this defines the correct barcode ID and lists the ID for the 19 active areas on the page.

7.1.3 Object Level XMP packets.

This document has 19 Object level XMP packets, each described in the sections below.

7.1.3.1 Center number (Number)

The field is for the 'Center Number' on the PDF document. This will be handwritten numerical data that we want the HWR to convert, so the field type is Number:

7.1.3.2 Study Number (Number)

The field is for the 'Study Number' on the PDF document. This will be handwritten numerical data that we want the HWR to convert, so the field type is Number:

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.3 Patient ID number for trial (Number)

The field is for the 'Patient ID' on the PDF document. This will be handwritten numerical data that we want the HWR to convert, so the field type is Number:

7.1.3.4 Project Title (Text)

The field is for the 'Project Title' on the PDF document. This will be handwritten tex we want the HWR to convert, so the field type is Text:

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.5 Researcher Name (Text)

The field is for the 'Researcher Name' on the PDF document. This will be handwritten tex we want the HWR to convert, so the field type is Text:

7.1.3.6 Information Sheet Date (Text)

The field is for the 'Sheet Date' on the PDF document. This will be handwritten tex we want the HWR to convert, so the field type is Text:

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.7 Information Sheet Version (Text)

The field is for the 'Sheet Version' on the PDF document. This will be handwritten tex we want the HWR to convert, so the field type is Text:

7.1.3.8 Consent Checkbox 1 - Consent (Boolean)

The field is for the first consent checkbox on the PDF document. This will be an area that the user ticks, so the field type is boolean:

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:xmp="http://ns.adobe.com/xap/1.0/"
xmlns:wgss="http://wacomgss.com/barbera/1.0/">
<rdf:Description rdf:about="http://signature.wacom.eu">
<wgss:PacketType wgss:level="field"/>
<wgss:FieldUUID wgss:pdfID="35242" wgss:fieldID="8"/>
<wgss:FieldLocation wgss:x="1038" wgss:y="660" wgss:w="60" wgss:h="60"/>
<wgss:FieldTag>Checkbox_1</wgss:FieldTag>
<wgss:FieldType>Boolean</wgss:FieldType>
<wgss:Required>True</wgss:Required>
</rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.9 Consent Checkbox 2 - Provision for withdrawal (Boolean)

The field is for the second consent checkbox on the PDF document. This will be an area that the user ticks, so the field type is boolean:

7.1.3.10 Company Name (Text)

The field is for the 'Company name' on the PDF document. This will be handwritten tex we want the HWR to convert, so the field type is Text:

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.11 Consent Checkbox 3 - Consent for data to be passed to company (Boolean)

The field is for the third consent checkbox on the PDF document. This will be an area that the user ticks, so the field type is boolean:

7.1.3.12 Consent Checkbox 4 - I Agree for GP access (Boolean)

The field is for the forth consent checkbox on the PDF document. This will be an area that the user ticks, so the field type is boolean:

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:xmp="http://ns.adobe.com/xap/1.0/"
xmlns:wgss="http://wacomgss.com/barbera/1.0/">
<rdf:Description rdf:about="http://signature.wacom.eu">
<wgss:PacketType wgss:level="field"/>
<wgss:FieldUUID wgss:pdfID="614EA" wgss:fieldID="12"/>
<wgss:FieldLocation wgss:x="1038" wgss:y="660" wgss:w="60" wgss:h="60"/>
<wgss:FieldTag>Checkbox_4</wgss:FieldTag>
<wgss:FieldType>Boolean</wgss:FieldType>
<wgss:Required>True</wgss:Required>
</rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.13 Consent Checkbox 5 - I Agree to take part (Boolean)

The field is for the forth consent checkbox on the PDF document. This will be an area that the user ticks, so the field type is boolean:

7.1.3.14 Participant Name (Text)

The field is for the 'Participant Name' on the PDF document. This will be handwritten tex we want the HWR to convert, so the field type is Text:

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.15 Date 1 (Text)

The field is for the 'Date' on the PDF document, next to participant name. This will be handwritten tex we want the HWR to convert, so the field type is Text:

7.1.3.16 Participant signature (Signature)

The field is for the 'Participant Signature' on the PDF document. This will be a digital signature field, so the type will be signature:

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:xmp="http://ns.adobe.com/xap/1.0/"
xmlns:wgss="http://wacomgss.com/barbera/1.0/">
<rdf:Description rdf:about="http://signature.wacom.eu">
<rdf:Description rdf:about="field"/>
<wgss:PacketType wgss:level="field"/>
<wgss:FieldUUID wgss:pdfID="AB33612" wgss:fieldID="16"/>
<wgss:FieldLocation wgss:x="745" wgss:y="1340" wgss:w="230" wgss:h="90"/>
<wgss:FieldTag>Participant_Signature</wgss:FieldTag>
<wgss:FieldType>Signature</wgss:FieldType>
<wgss:Required>True</wgss:Required>
</rdf:Description>
</rdf:RDF>
<?xpacket end="w"?>
```

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.17 Name of person (Text)

The field is for the 'Name of Person' on the PDF document. This will be handwritten text we want the HWR to convert, so the field type is Text:

7.1.3.18 Date (Text)

The field is for the 'Date' on the PDF document, next to the 'Name of Person' field. This will be handwritten text we want the HWR to convert, so the field type is Text:

As all fields on the PDF are mandatory, this field has the required attribute set to true.

7.1.3.19 Person Signature (Text)

The field is for the 'Person Signature' on the PDF document. This will be a digital signature field, so the type will be signature: