Permanent Pixels: Building Blocks for the Longevity of Digital Surrogates of Historical Photographs

R. van Horik*, H. Koppelaar#, K. van der Meer#, and P. Doorn*
***Netherlands Institute for Scientific Information Services, Amsterdam**
#Delft University of Technology, The Netherlands

Abstract

The conversion of historical photographs, into digital surrogates facilitates the easy access to the intellectual content. Durable high quality digital master images available today enable the creation of specific derivatives in the future. As historical photographs often are very vulnerable and the conversion process is expensive, the creation of high quality digital master images is important.

This paper presents three building blocks that can be used to realize a situation in which high quality digital surrogates of historical photographs are accessible and usable in the long run. The first building block for digital preservation is the application of a standardized image file format. This building block is based on the assumption that standards can be considered as durable. The second building block concerns the storage of the bit stream that makes up the digital surrogate in a durable data format. It is assumed that the XML data format as a non-proprietary data-encoding standard facilitates digital longevity. The third building block takes the creation of preservation metadata into consideration and is based on the assumption that metadata of digital objects is essential to understand and process the objects in the future.

Introduction

For archives, libraries, and museums the digitization of analogue source material is an exciting way to open up and exploit their holdings. Depending on its characteristics the digital surrogates of the analogue sources can be used for a wide range of purposes, such as a truthful representation of the original or a global reference to the original. As the creation of digital copies of analogue source material is expensive and the condition of the vulnerable originals deteriorates by the scanning process, it is important to create digital surrogates that can be used in the long run also for a number of purposes. Thus, the longevity of the digital objects is important. Digital preservation refers to all the actions required to maintain access to digital materials beyond the limits of media failure or technological change.

The variety and complexity of digital objects resulted in a number of preservation strategies and a wide range of implementation alternatives.[1] This paper presents building blocks that can be used to create durable digital surrogates of a specific digital object, namely digitized historical photographs. Building blocks are procedures, tools, specifications and guidelines available to realize the creation of durable digital surrogates. The most appropriate method for digital preservation is difficult to give, because it is not certain how information technology will evolve and what users expect from digital objects in the future. An evolutionary approach is required in this respect.

This paper consists of three parts. First the digitization of historical photographs is discussed. Than the research approach is covered upon which the formulation of the building blocks for digital longevity is based. The third part of this paper presents building blocks that can be used to improve the longevity of the digital surrogates of historical photographs. The pixels of the digital image that represents an analogue original should meet all benchmarked conversion criteria and be stored in a durable manner. For this situation the metaphor “permanent pixels” is used.

Digital Surrogates of Historical Photographs

For the digitization of photographic material a number of guidelines and references are available (e.g.[2],[3]). A main advice in these publications is to create “use-neutral” digital master images. With this approach, a photograph is digitized once, at the highest level of quality affordable, and settings such as color matching and contrast levels are set so that the image can be used for multiple applications, also in the long run.

What are Historical Photographs?

The features of the analogue documents to be converted into digital master images determine the requirements of the digitization process. In the course of time a number of photographic techniques were introduced.[4] A thorough assessment of the photograph is required prior to the creation of a digital surrogate in the form of a “use-neutral” digital master file. As the documentation of resources is a significant aspect for its longevity it is important to strive to an unambiguous determination of photographic items. Photographs can be considered as documents containing an image that has a continuous tone scale. The following pragmatic classification makes it possible to almost unambiguously identify a photograph, also by non-experts. The classification of photographic material consists of four categories:

- Type: reflective or transparent. A photograph is either a medium that reflects light or a medium that allows light to pass through.
- Polarity: positive or negative. Negatives contain tonal values opposite from reality.
- Color: monochromatic or polychromatic. A photographic with only one base color is monochromatic, e.g. black and white photographs, blue prints and also sepia toned photographs.
- Carrier: paper, glass, plastic, metal or other.

The identification of a photographic technique is even more explicit in case another two classification categories are applied.

- Dimension: the length and width of the photograph in a suitable format.
- Date: the date the photograph is created.

The classification can be used to determine the photographic technique of the original. Some techniques are considered as cultural more valuable than others. The distinction between fine-art photographs and historical, documentary photographs is not always clear as historical collections can contain important, high quality items. In general it can be stated that a fine-art photograph collection consists of a relative small number of valuable items, whereas an historical photograph collection contains a relative large number of items whose main function is to document events and objects. To a large extend the requirements of the digital capture device can be deduced from the classification.

Assessment of Historical Photographs Relevant for Digitization

Once the physical features and intellectual value of the photographic items are assessed the next steps in the digitization chain are the determination of the rendering intent and the determination to what extend the digital surrogate should resemble the original. Rather than the technical attributes of a digital capture device it is better to consider the features of the photographs and the rendering intents as the starting point for a digitization process. According to Frey[2] four rendering intents can be distinguished: rendering of the photographic image "as is", rendering the intent of the photographer, rendering of the original appearance of the photograph and rendering of the "real life object" visible on the photograph.

Also on the function of the digital surrogate decisions have to be made. The two extremities here are on the one hand a visual reference to the original and on the other hand a digital surrogate that can act as a replacement of the original in terms of spatial and tonal information content. This second principle is the closest to the notion of high quality, use-neutral "permanent pixels" that facilitate the passive preservation of the vulnerable original. In order to determine to what extent the digital surrogate resembles the original three levels of preservation can be distinguished as introduced in the "Illustrated book study".[5] The three levels of preservation are: Preservation of structure, representing the photographic technique, preservation of detail, representing the smallest significant detail of the image and preservation of essence, representing what the unaided eye can detect at a normal reading distance.

There are no guidelines or accepted standards for the creation of digital images according to one of the levels of preservation or one of the rendering intents. But the classification of the physical features of the photographic material, the rendering intent and the intended preservation level provides a good framework to fix the requirements of the digital surrogate. The relation between the analogue original and the digital surrogate is made explicit.

Benchmarked Digitization of Historical Photographs

The image capture device must be able to create a digital surrogate that meets the assessed requirements in terms of tonal range and image resolution. The quality of a capture device is determined by the characteristics of the signal registered by the device.

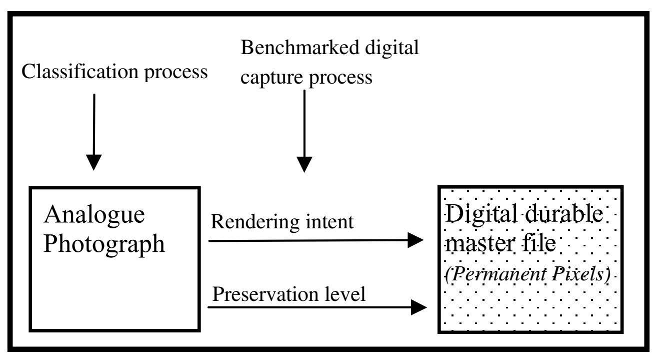

Figure 1. Durable digitization of historical photographs

As a capture device consists of a number of hardware components, such as a light source and optics, a benchmark phase is required to assess the performance of the capture device. With the help of calibration tools the features of the signal can be measured, such as the color and tone reproduction, the signal-noise ratio and optical frequency. It can be expected that in the near future no longer a trained operator and special calibrating software is needed to determine to what extend an image capture device is able to convert an original according to the assessed specifications. Increasingly consumer image capture systems will contain functions to create digital surrogates whose features are based on objective quality standards.

Figure 1 illustrates the relation between an analogue historical photograph and its benchmarked digital surrogate. A high quality digital surrogate containing all significant details of the original can be considered as a digital durable master file.

Despite the fact that digital surrogates can very close match its analogue counterpart regarding the structure, detail and essence both the digital and analogue images should be considered as separate entities. Both entities are close related to each other when the quality and longevity of the conversion is considered as important. Thus for "permanent pixels" the features of the analogue original and the digital surrogate it is based on will be considered as essential.

The value and importance of the analogue original will be one of the most important factors in the determination of the requirements of its digital surrogate. In case high quality digital master files are created based on a benchmarked digitization process the assurance of long-term access is eminent.

Research Approach

For this paper empirical research is carried out to determine what kinds of tools, procedures, specifications and guidelines are available that enable the long-term access and usage of a specific digital object. The most relevant building blocks to enable the durability of digital surrogates of historical photographs are collected and analyzed. The building blocks are presented in the next section.

A number of assumptions are formulated upon which building blocks for digital longevity can be based. These assumptions are part of the research framework. The attention is focused on the digital surrogate as an object that must be accessible and usable in the long term. The characteristics of the digital surrogate are based on a benchmarked conversion process as described in the previous section.

The first assumption is that the longevity of the digital surrogates is not obstructed by the actual state of art on storage media of digital data. The migration of digital objects is a matter of copying the bit stream to a new storage medium.

Secondly, XML is considered as a durable data format. It is a standardized, non-proprietary, and both human and machine-readable data format.

The third assumption is related to the importance of the strict formulation of the data elements that facilitate the digital preservation. A data element is a unit of data for which the definition, identification, representation and permissible values are specified by means of a set of attributes.[6] A number of important data element sets do exist that are used in a number of projects, such as the prospective NISO standard for technical metadata for digital still images, the Dublin Core metadata element set aimed at "resource discovery" or VRA core set developed by the Visual Resources Association. In practice very often data elements of different sets are "mixed and matched" or specific data elements are created on the spot. The ultimate preservation metadata element set will never be established resulting in the necessity to support the formulation and usage of so-called profiles that contain the data elements applied in a specific situation. The strict and unambiguous formulation of the data elements is important. For this the ISO/IEC 11179 family of standards on the standardization and specification of data elements can be used. This standard helps to "discriminate exactly what we know vaguely".

Building Blocks for Permanent Pixels

In this section three building blocks for durable digital surrogates of historical photographs are presented. These building blocks are: standardized image file formats, coding of the bit stream representing the pixels that make up the bitmap in XML format and the creation of preservation metadata.

It is not the ambition of this paper to present all possible building blocks that enable the longevity of digital surrogates of historical photographs, but the data structure of the digital surrogate and metadata relevant for durability can be considered as the most important variables for the long term access.

Standardized Image File Format

An obvious way to create durable digital objects is to use standardized file formats. A standard has the connotation of a well designed, widely used, and broad supported object. Empirically the requirements for standard data formats are given, followed by an assessment of existing data formats for digital raster files. A standard data format for digital objects must meet three conditions:

1. A large community must use the data format during a considerable period of time.
2. The specifications of the data format must be in the public domain or be published by a standards developing organization (SDO) such as ISO.
3. A wide range of relevant systems has to support the format, such as image capture devices as well as image processing systems.

For digital raster images acting as digital surrogates of historical photographs another three requirements for a standardized data format can be formulated:

4. Data compression should not be used due to a number of drawbacks. Most data compression leads to loss of image quality. Next to that a compressed image has a bigger risk to become unreadable in the future than an uncompressed digital image as a corrupt bit in a compressed image results in a "dead image", chances are that a corrupt bit in an uncompressed image is just a "dead pixel".
5. A durable data format for digital surrogates of historical photographs should contain facilities to store

preservation metadata. The quality and granularity of the metadata is an important factor for the future usage of the data format and thus its longevity.

6. The image file format must enable the coding of all significant characteristics of the analogue original that it is based on. This means that all colors, details and the dynamic range of the original can be represented in the digital image.

Table 1. Durability requirements and raster file formats **("+" = file format does meet requirement, "-" = file format does not meet requirement)**

	Raster file requirements	TI FF	JPEG	GI F	PNG
1	Used by a large community over a long time	+	+	+	-
2	File format specification is published	+	+	+	+
3	Supported by a wide range of applications	+	+	+	+
4	Supports un-compressed / single page images	+	-	-	-
5	Facilities for preservation metadata	+	-	-	+
6	Enables "full informational capture"	+	-	-	+

The first requirements of an image file format for the creation of high quality digital surrogates of historical photographs is that the format must be used for a considerable amount of time. The list of image file formats stated in the *Encyclopedia of Graphics File Formats*[7] published about 10 years ago, is used as a reference to raster image file format standards that are potential relevant. Four raster image file formats are still used today: TIFF, JPG, GIF and PNG. To what extent the four file formats do meet the durability requirements is stated in table 1. The PNG file format did not become the widely used standard as was foreseen by the developers. The JPEG, GIF and PNG data formats apply data compression. Thus, the TIFF file format seems the most durable raster image file format to be used for the coding of digital images.

TIFF as Durable Raster Image File Format

The TIFF ("tagged image file format") is largely hardware and software independent and has been around for more than 10 years. TIFF version 6 was made public in 1992 and supports the coding of color, compression methods and metadata. The TIFF version 6.0 specification is divided into two parts: baseline TIFF and TIFF extensions. Baseline TIFF is the core of TIFF, the essentials that all mainstream TIFF developers should support in their products.

TIFF extensions are TIFF features that may not be supported by all TIFF readers and this can hamper successful interchange and thus lowering the durability of the digital image. An essential feature of the TIFF standard is that it consists of fields that contain information on the bitmap data. This information is required by image processing systems in order to render the image. Other fields are used to store textual documentation on the image.

TIFF extensions are features that may not be supported by all TIFF readers. The official TIFF 6.0 standard contains a number of TIFF extensions and there are a number of TIFF extensions that are published independently from the officially published TIFF 6.0 standard, such as the TIFF/ "Electronic Photography" (ISO 12234) extension. The only extension relevant for the durability of a digital surrogate of an historical color photograph is the TIFF 6.0 extension for better color management. The CIELAB color space, supported by the extension on the TIFF 6.0 standard, has excellent applicability for device independent manipulation of continuous tone images.

The TIFF 6.0 specification can be considered as durable, provided that it is applied in a specific way. No compression method should be applied and no multiple page images must be created. TIFF 6.0 contains a number of information tags that enable the storage of preservation metadata. But support for storage of preservation metadata in baseline TIFF is rather limited. Only a limited amount of information fields is available for the storage of preservation metadata. The coding of all significant characteristics of color photographs is also problematic, because baseline TIFF does not support accurate coding of color information. The TIFF 6.0 extension contains a high quality color space (CIE LAB) that supports the accurate coding of colors.

Durable Encoding of the Bit Stream

The fundamental unit of digital data is the binary digit or bit. A bit stream is the sequence of bits that make up a digital object such as a digital raster image. The specification of the data format prescribes the organization or the bit stream. By using non-proprietary data coding methods the data object will remain accessible in case new software or hardware is introduced. The correct interpretation of the bit stream is the most challenging problem to be solved. For this a durable data structure of the bit stream is required as well as a method to interpret the data in a sensible way.

To a large extent the dependence on specific software and hardware for the processing of digital objects can be avoided by using the XML data format. This standardized data format is considered as self-descriptive and it does not require proprietary software to get access to the data. An essential requirement to access an XML formatted bit stream is to recognize the internal character encoding of the XML file. The auto detection or bootstrap mechanism of the XML encoding declaration is based on the ISO/IEC 10646 standard. This standard contains the legal binary codes that can be used for the markup codes or character data in the XML file. The XML markup codes are restricted in position and content to enable auto detecting the applied character encoding in use. According to the standard the first sequence of characters of an XML file must be <?xml>. The XML processor must detect this sequence of characters. After reading two or four octets the processor will be able to understand the byte ordering method and thus solve the bootstrapping problem.

The following three functions are important regarding the creation and usage of a bit stream representing a digital image expressed in XML:

- Expression of the content model in XML. The determination of the elements and attributes that are part of the image bit stream. The quality of the image file formatted in XML improves if the content model on which the XML instantiation is based meets the features of a standardized rich image file format.
- Binary to XML conversion. A generator is required that converts the binary image, e.g. in TIFF standard into an XML formatted file. This conversion process must of course take a content model into consideration. The content model must cover all features of the binary image file format.
- XML to binary conversion. The conversion of the XML formatted bit stream representing the digital image into a binary format that can be processed by a (future) computer platform. The archived, original binary format needs to be re-generated since it is the only understandable and usable format for the final user.

Currently there are three methods available that use the XML data format to create application-independent multimedia objects. These are the “Bitstream Syntax Description Language” (BSDL), the “Universal Virtual Computer” (UVC) and the “Formal Language for Audio-Visual Object Representation Language”(FLAVOR).

Bit Stream Syntax Description Language (BSDL) [8]

BSDL is developed to enhance the interoperability between different multimedia formats. A main benefit of BSDL is that objects expressed in XML according to BSDL are easily accessible by text retrieval engines and that multi-layered multimedia objects can be manipulated on an individual layer basis. The XML representation of the bit stream enables the definition of editing operations on the XML document e.g. by modifying some element or attribute values. For each binary format a specific content model is required and adjusted for the specific features of the binary format. The application of BSDL requires thorough knowledge on the multimedia format with respect to the way the bits are organized. BSDL uses the W3C style sheet language XSLT to transform the XML descriptions of multimedia formats. This transformation does not result in a binary object. The XSLT transformation results in a “transformed XML description” and must be converted into a binary bit stream. An XSLT style sheet contains one or several templates defining the modifications to be applied to the elements or to the attributes matching a set of conditions. The result of the XSLT transformation is also an XML file “from which it is possible to re-generate an adapted and compliant bit stream from the resulting description”.

The value of the BSDL approach for creating durable digital surrogates of historical photographs can be determined when the file format used for digital master images (e.g. TIFF 6.0) is expressed in BSDL and when the system as described in the system architecture is implemented. A strong point of BSDL is that it provides a framework for the expression of a binary format resulting in a durable expression of this format. The absence of “XML to binary” and “binary to XML” functionality is a drawback of this method in its current state.

Universal Virtual Computer (UVC) [9]

The goal of the UVC method is to enable a solution for the problem that digital data files and computer programs cannot be used anymore in the future. As a digital image is a data file and not a computer program, only the preservation of bit streams representing data files with the help of the UVC method is taken into consideration.

The basic idea of the UVC method is that the bit stream representing the data object is stored together with a logical view of the data. A logical view of the data follows the way the user normally thinks about the data, rather than the internal representation often designed for efficiency. Not only the logical view, but also the specification to process the data on a future platform is archived. The processing specification is based on a Universal Virtual Computer (UVC). The UVC program is independent of the architecture of the computer on which it runs. An UVC interpreter that has to be written for any future target machine. This means that every machine manufacturer has to produce an UVC interpreter. The instruction set of the UVC is kept to a minimum. The UVC examples given in publications on the UVC can be considered as simple data structures. It is not clear whether for complex data structures logical views can be designed that contain all relevant features.

The logical view of a data format is stored in a schema that contains the structure of the data format. As there are many data formats a lot of schemas will exist. A generic “schema to read schemas” approach is introduced that should be simple and intuitive and should endure for a long time to come, and be published in many places so that it remains known. It is assumed that decoding rules for the data are simple and will therefore easily survive for a very long time. The question is whether it is possible to create simple coding rules e.g. to code the colors of pixels according to a benchmarked color space.

The responsibility of extracting the logical data elements from the data stream lies with the methods. The UVC approach is based on the principle that the only standards needed now are the UVC and the data model for data and metadata. They are simple enough to endure.

The most challenging issue in using the UVC method to create a durable bit stream of a digital surrogate of an historical photograph is the formulation of a logical view in XML format of a digital surrogate that contains all relevant details of the benchmarked conversion process.

Formal Language for Audio-Visual Object Representation (Flavor/XFlavor) [10]

The formal language Flavor is a media representation language. It is developed for the description of binary multimedia objects, in order to simplify and speed up the development of software that processes multimedia data. The purpose of converting multimedia data in binary format into an equivalent XML document is for easier and more flexible manipulation of the data. The bit stream layer is abstracted from applications and the semantic values of the data (e.g. width and height of an image) are directly available, whereas with the proprietary bit stream format, such values must be extracted via bit string manipulations. Durability of digital objects is not mentioned as a design goal for Flavor.

In order to simplify interoperability among different applications XFlavor is developed. XFlavor has three major functions: (1) given a Flavor description of the bit stream syntax, a corresponding XML Schema can be created. (2) Given a Flavor description of a bit stream syntax, XML document generating code can be created. (3) Conversely, for an XML representation, an equivalent bit stream can be generated.

XFlavor offers an alternate format for multimedia data so that different applications can be created for the same data even if the expected syntax of the data is different among the applications. The XSLT language can be used for transforming XML documents. "Universal interoperability" is achieved by transforming XML documents into different structures usable for different applications. The style sheet along with the source document can be fed into an XSLT processor to generate a new document with desired structure and information. XFlavor has the ability to transform a given XML representation of multimedia content back into the bit stream representation. The current version of the XFlavor application contains a demonstrator procedure to convert images formatted in GIF file format into an XML Schema. XML Schemas containing the structure of other image file formats are not available yet

XFlavor is a translator that generates XML schemas from Flavor descriptions. Additionally, a method is available for producing an XML document from a Flavor described bit stream. As a complement, a software tool for converting the XML document back into an original bit stream format is also provided as part of the Flavor package.

Evaluation of Bit Stream Preservation Methods

Table 2 gives an overview of the three methods concerning their ability to perform the three essential tasks in creating a durable bit stream in XML format representing a digital image: (1) the conversion of the binary data format into an XML data format (2) the expression of the content model of the XML file and (3) the conversion of the XML data format into a binary data format.

The UVC method has the best approach concerning the conversion between XML and binary data representations of digital objects, because the virtual computer will be able to process the data, provided that on all future computer platforms an UVC emulator is available. As the UVC is intentionally kept very simple it is assumed that it will not be difficult to create this emulator. For each data format a content model has to be created. Flavor/XFlavor has implemented all required functionalities, but is based on the Java and C++ language. The stability and durability of these languages is the most important factor in the assessment of the method for the conversion between binary and XML format. The durability of the C++ and Java platform is the central issue in comparing the Flavor method with the UVC approach. The UVC approach is based on the principle that a simple virtual machine instruction set is the core of the method.

Table 2. Comparison of three methods with respect to the ability to express the bit stream of a digital surrogate in XML **("-" = task is only available in the system design, "+" = task can be performed with the method, but adjustments are required to enable the processing of digital master images, "++" = task can be performed by the method)**

	Binary to XML conversion	Content model in XML	XML to Binary conversion
BSDL	-	++	-
UVC	++	+	++
XFlavor	+	+	+

Application of Preservation Metadata

Preservation metadata is documentation that is required to assess the features of a digital object, also in the future. Preservation metadata serves a wide range of functions: resource discovery, object management, control of intellectual property rights, version identification, certification of authenticity, rendering instructions, etc. For the application of preservation metadata that enables the longevity of digital objects a number of data elements are connected to the digital object. Both technical specifications as well as administrative information are part of the preservation data element set.

Issues

For several issues related to the application of preservation metadata more than one approach can be applied. The first issue concerns the actual storage of the metadata. In practice three general methods can be distinguished. Metadata is stored as part of the digital object, e.g. in the information tags of a digital file, in a database or as part of the object name or file system. The durability of the metadata as electronic data will improve in case the XML data format is used as syntax for the data elements instead of a proprietary data format.

The second issue concerning the application of preservation metadata is the lack of a general accepted list of preservation data elements for digital surrogates. The diversity of metadata element sets makes clear that a common standard that is general accepted is far from reality. The long time and organizational effort it took for the library community to adhere to standardized principles for bibliographic metadata learns that it will be merely impossible to come to a general accepted and correctly applied set of metadata elements that enable the durability of digital surrogates. Some communities however are able to adhere to a common metadata standard

The third issue is related to the second one and concerns the difficulties to assess and apply relevant existing domain knowledge that results in high quality metadata and thus improves the longevity of the object that is documented. In case a thesaurus term or classification scheme can be linked to a data element the quality of the metadata will improve.

Metadata Registries

Two methods are available to formulate preservation metadata for digital surrogates of historical photographs. The first one is to bring together a group of stakeholders that discuss and assess the requirements for preservation metadata and compile a set of data elements. To a large extend the prestige and communication skills of the stakeholders determine the acceptance and usage of the data elements.

The second method is to use existing sets of data elements. A metadata registry can help to determine which data elements are relevant for a specific object. In case data elements are lacking they may be added to the registry. One can benefit from both methods by submitting to the registry the data elements created by a group of stakeholders. The data elements can be reviewed, assessed and applied by others.

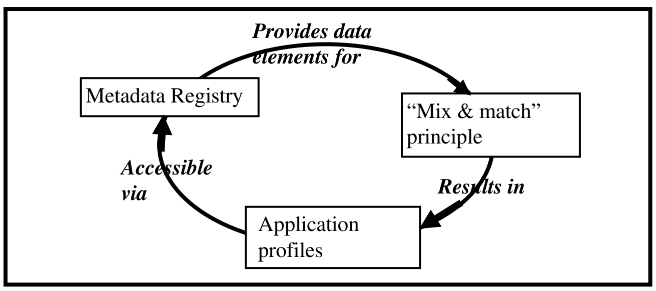

Figure 2. Registries and the application of preservation metadata.

The first approach was carried out within the framework of a EU funded project and resulted in the SepiaDES[11] set of data elements.

For the second approach a system has to be used that contains detailed information on data elements. The ISO/IEC 11179 family of standards contains an overview of the basic attributes of data elements. The standard lists 23 attributes providing that the value and usability of a data element cannot be misunderstood. The process of the creation and usage of data elements with the help of a metadata registry is illustrated in the figure 2. Application profiles are element sets optimized for a particular local application. It may include data elements drawn from one or more other element sets. The "mix and match" principle concerns the compilation of a set of data elements intended for specific use.

A metadata registry system relevant for preservation metadata of digital surrogates of cultural materials that enables the storage according to the ISO/IEC 11179 standard does not exist at the moment. But the CORES registry can be used to assess the relevance and efficiency of metadata registries to create and apply data elements[6]. The two main factors that obstruct the wide spread use of metadata registries are the labor-intensive procedure to register data elements into a registry and the lack of a metadata registry as a structural, permanent service. Also the CORES registry does not have a structural character because the project funding has ended.

Conclusion

A number of quality requirements for durable digital surrogates of historical photographs are distinguished. The conversion process must be based on benchmarked settings. In case high quality digital master images are created the long-term access and usage requires special attention. Despite the fact that it is impossible to predict which method in the end will be the most efficient, some methods currently available can be assessed as building blocks for durable digital surrogates of historical photographs, or permanent pixels. These methods are: the application of image file format standards, the expression of the digital surrogate in the XML data format and the creation and employment of preservation metadata. This paper reports on research to what extent assessed durability building blocks can be realized in practical situations.

References

1. M. Jones and N. Beagrie, *Preservation management of digital materials*, British Library, London, UK, 2001.
2. F. Frey and J. Reilly, *Digital imaging for photographic collections. Foundations for technical standards*. Image Permanence Institute, Rochester, 1999.
3. *Guides to quality in visual resource imaging*, DLF, RLG, CLIR, 2000. Online available at: <http://www.rlg.org/visguides>.
4. M. Frizot (ed.), *A new history of photography,* Könemann, Köln, 1998.
5. Kenney and L. H. Sharpe, *Illustrated book study: digital conversion requirements of printed illustrations*, Report to the Library of Congress directorate, 1999. Online available at: <http://www.library.cornell.edu/preservation/ill_bk_cover.htm>.
6. R. Heery and others, Metadata schema registries in the partially Semantic Web: the CORES experience, *Proceedings of the 2003 Dublin Core Conference* (2003).
7. J.D. Murray and W. vanRyper, *Encyclopedia of graphics file formats*, O'Reilly, Sebastopol, CA, 1994.
8. M. Amielh and S. Devillers, Bitstream syntax description language: Application of XML-Schema to multimedia content adaption, *Proceedings of the 11th International World Web Conference*, Honolulu, 2002.
9. R. Lorie, *The UVC: a method for preserving digital documents. Proof of concept,* Den Haag Koninklijke Bibliotheek, 2002.

10. Eleftheriadis, and D. Hong, 'Flavor: a language for media representation'. In: B. Fuhrt and O. Marques *Handbook of video databases.* Boca Raton, CRC Press, pp.69-96 (2003).
11. K. Aasbø and others (ed.), *SEPIADES, Recommendations for cataloguing photographic collections: advisory report by the SEPIA working group on descriptive models for photographic collections*, European Commission on Preservation and Access, Amsterdam, 2003.A.

Biography

René van Horik received his M.A. degree in History at the University of Nijmegen. He works as a researcher and project leader at the Netherlands Institute for Scientific Information Services. He is involved in several projects in which ICT is applied in the Humanities. On a part-time basis he is working on his dissertation on the longevity of digital images at Delft University of Technology.