# Vatican Library Long term digital conservation for manuscript collection









### VATICAN LIBRARY MISSION

The Vatican Library is institutionally a library of preservation which for more than 500 years has fulfilled its fundamental role, preserving, protecting and restoring the patrimony of books kept there. This daunting task was made even more difficult because the Vatican Library has never forgotten that free consultation of its immense patrimony for scholars from around the world was a mandate from which the Library could not be dispensed.

### VATICAN LIBRARY CHALLENGES

Quite in keeping with this role and knowing the needs to also preserve its manuscripts for the future generations, the Library started a project of "long-term digital preservation" that on one hand gives all possible guarantees of longevity of the realized technological product and on the other would allow the disclosure of the manuscripts possessed by the Vatican Library to a wide representation of users worldwide.

### GOSPEL OF LUCA WITH THE TEXT OF THE "OUR FATHER"



# SECURITY SYSTEMS

# RFID tags store all data:

- Silicon solid memory
- Reprogrammable
  - Archives historical book data and history
  - Inventory
  - Transactions
  - Removals
  - Restoration
  - Who managed and red the book

- Chip data are distributed on extended network via web



### PASSIVE DEVICE RF.ID 13.56 MHZ.

### LIBRARY STANDARD INDOOR

Riv.Bibliogr.11bis(20).Cons. Milkau, Fritz,1859-1934

Centralkataloge und Titeldrucke; geschichtliche erörterungen und praktiske Vorschläge im hinblick auf die Herstellung eines Gesamtkatalogs der preussischen Wiessenschaftlichen Bibliotheken;

(20)

Pubblicazione: 1898.



### SECURITY SOFTWARE

### VIDEOCLIP SEARCH BY NAME, ID CARD, GATE, SENSOR, VIDEOCAMERA

	🔰 Identificazior	e							X
	Filtro ID Tessera Zona	Cognome e Nome Izzo Varco	IDBib	Segnatura	Da: 08/08/2014 Telecamera	A: • 08/08/2014	Accesso	Profilo TUTTI	•
	- Risultati	O IDBib	SEGNATURA	LIVELLO		TESSERA	REFERENTE	PROFILO	
					1220, Comonol			01	
Pergamon <sup>®</sup>	Records: 1	1				<u>Esp</u>	orta Dettaglio	Vedi Filmat	• D

### VATICAN LIBRARY TREASURES



### NEED TO DIGITIZE

Vatican Library needs to protect and share around the world testimonials of human

history



# HOW AND IN WHICH FORMAT?

### GUARANTEE INTEROPERABILITY OVER TIME



Illustrations of Divine Comedy by Botticelli





Koran in Kufic

Bilingual version of Iliad

## THIS FORMAT "FITS" BETTER FOR BAV



	TIFF	<u>FITS</u>
Architecture	32 bit	64 bit
Max File size	4 GB	Unlimited
3D Visualisation	None	Yes
Format	Proprietary	Open Source
Updated	1998	6 monthly

### F.I.T.S. Flexible Image Transport System)

The table of colors (color look-up table or colormap), uniquely combines a number stored in the array to a color image.

In fact, in the digital image is decomposed into a many "points", and each point has its own color, as in the example below:

Tabella 1	Tabella
	00
	01
	10
	11
	••

Tabella 3

01	01	01	01	10	10	10	10
01	01	01	01	10	10	10	10
01	01	01	01	10	10	10	10
01	01	01	01	10	10	10	10
11	11	11	11	00	00	00	00
11	11	11	11	00	00	00	00
11	11	11	11	00	00	00	00
11	11	11	11	00	00	00	00

Table 1 represents the original image, the table 2 shows how each color is coded in numeric format, while Table 3 shows how the numerical matrix is stored. It 'clear that larger numbers guarantee a greater chance chromatic.

Currently in the project using the encoding called sRGB, with arrays as large as the image (in pixels) of three levels: one for each of the primary components (R, G, B). Eg. an image of 2048x3916 pixels in the file becomes an array 2048x3916x3. In each level is stored for each point the value of the primary component R, G or B.

In this encoding each color is identified by a triplet of values between 0 and 255 (for efficient coding in binary format)



Knowledge of Table 2 is essential, because the same numeric value may indicate different colors depending on the chosen encoding.

FITS format, unlike other formats born specifically to handle images, not inside any standard system to save a color look-up table, or color table.

What he first glance looked like a weakness, however, he has turned into one of his strong point, because just that this format allows flexibility unthinkable in other formats.

Indeed formats dedicated essentially to the images usually they have the look-up table incorporated in the specifications, and this sometimes makes them "rigid" and not able to follow the inevitable evolution of the sector.

How Sempio you can cite the GIF format, which can not handle more than 256 colors, or TIFF, which despite its undeniable qualities can not handle image files larger than about 2 GB, size if it once seemed completely unreachable now instead possible to obtain the new sensors of the apparatuses of acquisition.

The FITS, being born to handle numerical data (and not only) with the greatest freedom, does not have these constraints and is therefore able to adapt its structure to any changes in storage formats or changes in the type of data stored.

This flexibility is obtained by the use of certain bad that define a simple manner the type and size of the stored data.

The basic data types in FITS are the following:

Name	Size in Bits Example			
ASCII character	8	'Saturn'		
Integer				
Unsigned, one byte	8	127		
Unsigned, two bytes	16	32767		
Unsigned, four bytes	32	1451342181		
Single-precision real				
Fixed-point notation	32	3.14159		
Exponential notation	32	0.314159E+01		
Double-precision real				

Exponential notation 64 0.3141592653525D+01

The data type Unsigned 32-bit "limit" even the maximum size of points on the axes of the image matrix (to 4,294,967,295).

#### BITPIX

BITPIX è l'intero che indica il numero di bits usati per rappresentare ogni valore nella matrice dei dati. Nel caso di immagini indica il numero di bits per pixel.

#### **BITPIX Value Data**

Character or unsigned binary integer
16-bit two's complement binary integer
32-bit two's complement binary integer
32-bit floating point, single precision
64-bit floating point, double precision

#### NAXIS

NAXIS can go from 0 to 999, and the number of axles of the data matrix.

#### NAXISn

NAXIS*n* It indicates, with a non-negative integer, the length of each axis, that is, the number of elements of which it is composed. Each element of the array has size BITPIX

With a suitable combination of these bad you can easily save images with various types of colormap. The most common system is the space-color RGB (references). This theoretical model has several practical implementations, and in our case it was decided to use the sRGB (references to sRGB, AdobeRGB, etc.). In a concrete image of size N x M is stored in a matrix (N x M x 3), with layer 0 dedicated to the R component, the layer 1 to that G and the layer 2 to that B. Each single color component is saved in an entire 8-bit, and this allows a color change of approximately 16 ML of colors, ability is considered more than enough to represent so accurately the original colors. But if one day you want to be accurate even more pressure as color would be enough to use a matrix image with BITPIX = 16 to increase exponentially color accuracy possible.

Naturally sRGB encoding is not the only possible, there are several other AVAILABLE and much equipment, such as that CYMK (references) where the single color is decomposed into four components. In questo caso nel FITS usare una codifica CYMK vorrebbe dire semplicemente aggiungere un layer in più nella matrice dei dati, da N x M x 3 a N x M x 4 ed evidenziare questa scelta nell'Header del file. Se poi, per esempio, si diffondesse una look-up table a SEI componenti, questa attualmente sarebbe difficilmente gestibile dai formati grafici attuali, ma per il FITS vorrebbe dire definire una matrice immagine con NAXIS=6.

Clearly it is essential that in the FITS file where you saved the image to be shown what kind of colormapping was used, and for this purpose was introduced keyword colormap, which provides this indication. The value is used sRGB, which indicates the use of the homonymous standard. General information about the project.

- Why digitize
- The issues involved
- How did it come to FITS
- Not only FITS, the rest of the Project

### To decide whether to:

• write everything from scratch with the detailed reasons that led to the decision to start this project.

• put a link to existing site b.p. BAV or DigitaVaticana, perhaps by integrating and / or expanding the information they present Elements in my opinion to point out and highlight:

• For this project has not been made a simple translation of FITS keywords in the TIFF tags in use, but have been identified those judged necessary to correct long-term storage and image identification and sought the FITS keyword that best express them. Only when the parameter had no evidence has been introduced, the number as small as possible, the new keywords, use only "local". Can be cited as an example of this work the keywords that define who / how / when he scanned a document and who has verified its correct implementation and storage, keywords that have no equivalent in TIFF.

Clarify the question of the color space

• Explain that you plan to use only one header and one image file, even if it would be possible to enter more than one in every file. For a description of the keywords we could follow the pattern that follows the form that describes the keywords FITS standard and that I report with some modifications as an example. It could be divided into three parts:

existing keywords used in a "different" (but we must try not to put this word because it evokes deviations from the standard)
 new keywords

### 3. standard keywords used

### USE: BAV X Vatican Library

Create a space xi credits where we can bring together those who have participated in the realization effort. You state that you have very good web designer who created the sites beautiful and very functional, so you do not need any advice on this. Below I just made a diagram divided in "functional areas", just to see if there's anything there must be

V.A.L. F	ITS Keyword Dictionary: Version 1.0 - April 2015	
Generalità ———> STANDARD	This keyword dictionary defines keywords which may be used in Long Term Data Preservation project (ampliare) . Each keyword definition has the following fields: KEYWORD	Link Esterni
KEYWORDS	The name of FITS keyword(s). DEFAULT	Sito BAV
Keyword 2	no default the value "none" is given. When there are multiple primary keywords there will be a matching number of defaults. When both the primary and default have indices	Digita Vaticana
Keyword n	then the index value is the same. A common case is when the indexed primary keywords default to the unindexed version of the primary keyword. This occurs when the information is the same for all elements of an indexed class. If the unindexed	NTT Data
COMMON USED	keyword is also missing it would then default to the specified keyword for it. For example, in the case that there are multiple Tutors described in the header, the TUTOR	Others
KEYWORDS	keywords are defined to be TUTOR01, TUTOR02, etc.	
Keyword 1 Keyword 2	INDEX Index values for indexed keywords. When there are multiple keywords there will be multiple index values. If the keyword has multiple indices the values apply from left to	
Keyword n	right. There are two types of index values, a numeric range for 'd' indices and a character class for 'c' indices. Numeric ranges are of the form M-N where M and N are integers with M<=N. For example,	
NEW KEYWORDS	01-09 Integers from 1 to 9 > CONTINUE <	
Keyword 1		
keyword 2		
Keyword n		
All keywords in PDF	Credits:	

V.A.L. FITS Keyword Dictionary: Version 1.0 - April 2015					
	Last Revision ##-######				
Generalità	KEYWORD: COLORMAP	Link Esterni			
STANDARD	DEFAULT: RGB				
KEYWORDS	INDEX: none	Sito BAV			
Keyword 1 $\rightarrow$	VALUE: String				
Keyword 2	UNITS: None	Digita Vaticana			
Reyword 2	COMMENT: Codifica colore	Digita Vaticalia			
	EXAMPLE: RGB				
Keyword n	DESCRIPTION:	NTI Data			
	Definisce la codifica colore usata ecc ecc				
COMMON USED	> CONTINUE <	Others			
KEYWORDS					
Keyword 1					
Keyword 2					
Keyword n					
Reyword fi					
		_			
Kowword 1		-			
Keyword 2					
Keyword n					
	Crodits:				
All Reywords in PDF					

#### V.A.L. FITS Keyword Dictionary: Version 1.0 - April 2015

This keyword dictionary defines keywords which may be used in this project.

Each keyword definition has the following fields: KEYWORD The name of FITS keyword(s).

#### DEFAULT

Keyword or value which is applicable when the primary keyword is missing. If there is no default the value "none" is given. When there are multiple primary keywords there will be a matching number of defaults. When both the primary and default have indices then the index value is the same. A common case is when the indexed primary keywords default to the unindexed version of the primary keyword. This occurs when the information is the same for all elements of an indexed class. If the unindexed keyword is also missing it would then default to the specified keyword for it. For example, in the case that there are multiple Tutors described in the header, the TUTOR keywords are defined to be TUTOR1, TUTOR2, etc.

#### INDEX

Index values for indexed keywords. When there are multiple keywords there will be multiple index values. If the keyword has multiple indices the values apply from left to right. There are two types of index values, a numeric range for 'd' indices and a character class for 'c' indices. Numeric ranges are of the form M-N where M and N are integers with M <=N. For example,

1-9 Integers from 1 to 9

#### HDU

A recommendation for whether the keyword should be part of the primary or extension header-data unit

#### VALUE

The format of the value field (string, integer, ecc). The format descriptions are those used by istitutions like IRAF and are described below. The numeric formats also indicate a numeric FITS keyword value and other formats indicate string keyword values. Special types of value formats are indicated in parenthesis. These include 'date' for a FITS date string, 'checksum' for a FITS checksum string, 'datasum' for a FITS datasum string, and 'section' for an image section string.

#### UNITS

Expected or suggested units for the value. These are either explicit units or a keyword that specifies the unit.

#### COMMENT

The default comment in a FITS card. This is only a default and a more appropriate and detailed comment may be used.

#### EXAMPLE

An example keyword value.

DESCRIPTION A full description of the keyword and its use.

### **FITS Standard Keywords**

(From official FITS site http://heasarc.gsfc.nasa.gov/docs/fcg/standard\_dict.html)

KEYWORD: **SIMPLE**REFERENCE: <u>FITS Standard</u> STATUS: mandatoryHDU: primaryVALUE: logicalCOMMENT: does file conform to the Standard?DEFINITION: The SIMPLE keyword is required to be the first keyword in the primary header of all FITS files. The value field shall contain alogical constant with the value T if the file conforms to thestandard. This keyword is mandatory for the primary header and is notpermitted in extension headers. A value of F signifies that the filedoes not conform to this standard.

KEYWORD: **BITPIX**REFERENCE: <u>FITS Standard</u> STATUS: manditoryHDU: anyVALUE: integerRANGE: -64,-32,8,16,32COMMENT: bits per data valueDEFINITION: The value field shall contain an integer. The absolutevalue is used in computing the sizes of data structures. It shallspecify the number of bits that represent a data value.

KEYWORD: **NAXIS**REFERENCE: <u>FITS Standard</u> STATUS: mandatoryHDU: anyVALUE: integerRANGE: [0:999]COMMENT: number of axesDEFINITION: The value field shall contain a non-negative integer nogreater than 999, representing the number of axes in the associateddata array. A value of zero signifies that no data follow the header inthe HDU. In the context of FITS 'TABLE' or 'BINTABLE' extensions, thevalue of NAXIS is always 2.

KEYWORD: **NAXI Sn**REFERENCE: <u>FITS Standard</u> STATUS: mandatoryHDU: anyVALUE: integerRANGE: [0:]

COMMENT: size of the axisDEFINITION: The value field of this indexed keyword shall contain anon-negative integer, representing the number of elements along axis nof a data array. The NAXISn must be present for all values n = 1, ..., NAXIS, and for no other values of n. A value of zero for any of the NAXISn signifies that no data follow the header in the HDU. If NAXIS is equal to 0, there should not be any NAXISn keywords.

KEYWORD: **EXTEND**REFERENCE: <u>FITS Standard</u> STATUS: reservedHDU: primaryVALUE: logicalCOMMENT: may the FITS file contain extensions?DEFINITION: If the FITS file may contain extensions, a card image with the keyword EXTEND and the value field containing the logical value Tmust appear in the primary header immediately after the last NAXISncard image, or, if NAXIS=0, the NAXIS card image. The presence of thiskeyword with the value T in the primary header does not require that extensions be present.

**KEYWORD**: INSTRUME **REFERENCE:** FITS Standard STATUS: reserved DEFAULT: None INDEX: None HDU: any VALUE: String COMMENT: name of instrument. EXAMPLE: Hasselblad H3DII-50 DESCRIPTION: Brand and model of the acquisition device KEYWORD: DATE-OBS REFERENCE: FITS Standard DEFAULT: None INDEX: none HDU: Any VALUE: String Date and time of acquisition COMMENT: 2014-01-07T14:47:18 EXAMPLE: DESCRIPTION: Date and time of the image acquisition, in the format specified in the FITS Standard. The new Y2K compliant date format is 'yyyy-mm-dd' or 'yyyy-mm-ddTHH:MM:SS[.sss]'.

 KEYWORD:
 DATE

 REFERENCE:
 FITS Standard STATUS:
 reserved DEFAULT:
 None

 INDEX:
 None

 HDU:
 Any

VALUE: Date (YYYY-MM-DDTHH: MM: SS) COMMENT: Date and time of FITS file creation EXAMPLE: 2014-01-08T13:22:05 DESCRIPTION: The date of the creation of FITS file, in the format specified in The FITS Standard. The new Y2K compliant date format is 'yyyy-mm-dd' or 'yyyy-mm-ddTHH: MM: SS[.sss]' **KEYWORD**: AUTHOR REFERENCE: <u>FITS Standard</u>STATUS: reservedHDU: Any DEFAULT: None INDEX: None VALUE: String COMMENT: author of the data EXAMPLE: AUTHOR = 'Victor Vitensis, Niccolò Poggibonsi' DESCRIPTION: The value field shall contain a character string identifying who created the information in the data associated with the header. This keyword is appropriate when the data originate in a published paper or are compiled from many sources. It's possible the presence of CONTINUE tag, twhen a large numbers of authors is present. Example of use: AUTHOR = 'Victor Vitensis, Niccolò Poggibonsi, ' CONTINUE 'Bonifacio from Ragusa' / End AUTHOR **KEYWORD**: ORIGIN REFERENCE: FITS Standard STATUS: reservedHDU: Any DEFAULT: None INDEX: None VALUE: String COMMENT: Copyright notices EXAMPLE: Vatican Apostolic Library DESCRIPTION: If present indicates that reproduction of the images are subject a some type of copyright rules of the organization described **KEYWORD**: OBJECT **REFERENCE**: FITS Standard STATUS: reservedHDU: Any DEFAULT: None VALUE: String UNITS: None COMMENT: Item identification EXAMPLE: Barb.gr.238.pt.1 DESCRIPTION: Identifies the work, generally the volume or manuscript, to which belongs the image or the subject / object of the image. KEYWORD: **CTYPEn**REFERENCE: WCS-FITS STATUS: reserved DEFAULT: none INDEX: 1-n

HDU: image

VALUE: stringUNITS: nomeCOMMENT: name of the coordinate axisEXAMPLE: CTYPE2 = ' 'DEFINITION: The value field shall contain a character string, giving the name of the coordinate represented by axis n.

KEYWORD: CRPIXn REFERENCE: WCS-FITS STATUS: reserved DEFAULT : 0.0 INDEX: 1-n HDU: image VALUE: real COMMENT: coordinate system reference pixel EXAMPLE: 0.0 DEFINITION: The value field shall contain a floating point number, identifying the location of a reference point along axis n, in units of the axis index. This value is based upon a counter that runs from 1 to NAXISn with an increment of 1 per pixel. The reference point value need not be that for the center of a pixel nor lie within the actual data array. Use comments to indicate the location of the index point relative to the pixel KEYWORD: CRVALnREFERENCE: WCS-FITS STATUS : reserved DEFAULT : 0.0 INDEX: 1-n HDU: imageVALUE: realCOMMENT: coordinate system value at reference pixel EXAMPLE: 0.0 DEFINITION: The value field shall contain a floating point number, giving the value of the coordinate specified by the CTYPEn keyword at the reference point CRPIXn. Units must follow the prescriptions of section 5.3 of the FITS Standard. KEYWORD: CUNITnREFERENCE: WCS-FITSd STATUS : reserved INDEX: 1-n DEFAULT : 'mm imageVALUE: StringCOMMENT: Unit of 1st axisEXAMPLE: 'mm 'DEFINITION: measurement unit used for n axis KEYWORD: CDELTnREFERENCE: 'HDU: WCS-FITS STATUS: reserved INDEX: 1-n DEFAULT: 1 HDU: image

VALUE: realCOMMENT: coordinate increment along axisEXAMPLE: '0.0635 'DEFINITION: The value field shall contain a floating point numbergiving the partial derivative of the coordinate specified by the CTYPEn keywords with respect to the pixel index, evaluated at the reference point CRPIXn, in units of the coordinate specified by the CTYPEn keyword. These units must follow the prescriptions of section 5.3 of the FITS Standard.

KEYWORD: **PCn\_m**REFERENCE: <u>WCS-FITS</u> STATUS:

reservedHDU: primary & extensionVALUE:

realCOMMENT: coordinate system value at reference pixel

DEFINITION: Used for an eventual rotation matrix, used with CDELT1 and CDELT2 in

the final version of the FITS-WCS standard as an alternative to to the CD\_j matrix.

#### COMMON USED FITS KEYWORDS

 KEYWORD:
 CREATOR

 REFERENCE:
 HEASARC STATUS:
 Common Used DEFAULT:
 None

 INDEX:
 None
 None

 INDEX:
 None
 None

 HDU:
 pimary& extension
 Strip

 COMMENT:
 Software that created this FTTS file
 Software that created this FTTS file

 EXAMPLE:
 Value field shall contain a character string giving the current FTTS file.
 name, and optionally, the version of the program that originally

 Example: Weits USt Software Ver. 4.5
 Software Ver. 4.5
 Software Ver. 4.5

#### KEYWORD: **PROGRAM**

#### KEYWORD: LONGSTR

REFERENCE: <u>GSFC-NASA FITS Registerd Convention</u>STATUS: Common Used DEFAULT: OGIP 1.0 INDEX: None HDU: Primary & extension VALUE: String COMMENT: The OGIP Long String Convention may be used EXAMPLE: OGIP 1.0 DESCRIPTION: According to the HEASARC Convention of 07/27/1994 regarding the use of long strings in the values of FITS keywords it is recommended that the following keywords be added to the header of any HDU that uses this long string convention: LONGSTRN= 'OGIP 1.0' / The OGIP long string convention may be used. COMMENT This FITS file may contain long string keyword values that are COMMENT continued over multiple keywords. This convention uses the '&' COMMENT character at the end of a string which is then continued COMMENT on subsequent keywords whose name = 'CONTINUE'. The presence of the LONGSTRN keyword serves to indicate that long string keywords may be present in the FITS file. The value of this keyword gives the name and version number of the specific convention that is used, which in this case is the OGIP (Office of Guest Investigator Programs, at the HEASARC) long string convention, version 1.0. The value of this keyword is a string so that

it may be used to give the name of any other convention that the FITS community might adopt.

# SCANNING



Ancient manuscripts stored in Vatican Library are selected and carefully scanned by special devices

## TUNING AND ACQUISITION DEVICE

The preliminary stages of study and analysis obtained from the data received during the TEST BED, determined the selection of the Metis planetary devices that adopt CCD trilinear sensors, designed and manufactured by Kodak. These KLI sensors were designed about 20 years ago (and updated over time) for military and professional use only.. Recently, Kodak sold its division that produces CCD sensors to "TRUESENSE Imaging" which continues the development and updating using the latest technologies.

### CALIBRATION GUIDELINES

This guarantees careful selection according to production standards and careful disposal of materials at the limits of productivity. Basically, however, the two fundamental parameters, that as a solid foundation, support this project are definitely the choice of digital preservation format and strict observance in controlling technological obsolescence in reader compatibility of the conservation devices.

# CALIBRATION ACQUISITION DEVICES

The quality controls are inspired by the standard called UTT It is an open standard for the verification of digitized images from any type of scanner and digital camera system. Using the test target UTT are possible the following measurements:

- MTF (Modulation Transfer Function) in accordance with ISO 16607
- Color Reproduction (Color Spaces Supported CIE 1976, CIE 1994, CIE 200, etc)
- Noise level in accordance with ISO 12233

# SCANNING

No wear from the original copies manipulation





# SAVING



To ensure long-term storage, images are saved in multiple location using the most advanced technologies

# SOFTWARE DEFINED APPROACH

BUILDING ITAAS

Ability to handle this kind of heritage in a very simple and secure way

- Makes our IT infrastructure agile & innovative
- Increased efficiency and effectiveness
- Improved visibility across systems and Data Center



## DISTRIBUTION



The digitalized images are available on-line on Vatican portal for free search and visualization by worldwide

user

### ONLINE MANUSCRIPTS @ VATLIB.IT



### ONLINE MANUSCRIPTS @ VATLIB.IT



# PEOPLE & HIGH-TECHNOLOGY THE SECRET MIX TO BE INNOVATIVE

Vatican Library IT Team is working closely with several PARTNER

PARTNERS provides flexible powerful systems as well as very innovative solution design approach



"Continuous brainstorming attitude"



# WHAT'S THE NEXT PROJECT ?





### OPEN YOUR HEART AND YOUR MIND AND ANY DREAM WILL COME TRUE

BIBLIOTECA APOSTOLICA VATICANA LUCIANO AMMENTI