Historisch

Kulturwissenschaftliche

Informationsverarbeitung

Session:
Characterisation of Digital Content

Digital Preservation – The Planets Way
Sofia, 16 – 18 September 2009

Volker Heydegger and Jan Schnasse

Overview

- Part 1: Characterising Digital Content: The eXtensible Characterisation Languages

- Part 2: Demonstration of XCL Tools:

  Evaluation of Format Conversion

Historisch
Kulturwissenschaftliche
Informationsverarbeitung

Characterising Digital Content: The eXtensible Characterisation Languages

Digital Preservation – The Planets Way
Sofia, 16 – 18 September 2009

Volker Heydegger

Overview

- ❑ Characterisation: Why and What
- ❑ About File Formats
- ❑ XCL: Goals
- ❑ XCL: Architecture
- ❑ XCL by Example

1

Characterisation

Why characterisation?

“Characterisation is an essential precursor to preservation. It provides the information required to make preservation planning decisions about digital objects, and to validate the results of preservation actions. “

(A. Brown: Developing Practical Approaches to Active Preservation, IJDC, 2007)

Characterisation

Why characterisation?

“Characterisation is an essential precursor to preservation. It provides the information required to make preservation planning decisions about digital objects, and to validate the results of preservation actions. “

(A. Brown: Developing Practical Approaches to Active Preservation, IJDC, 2007)

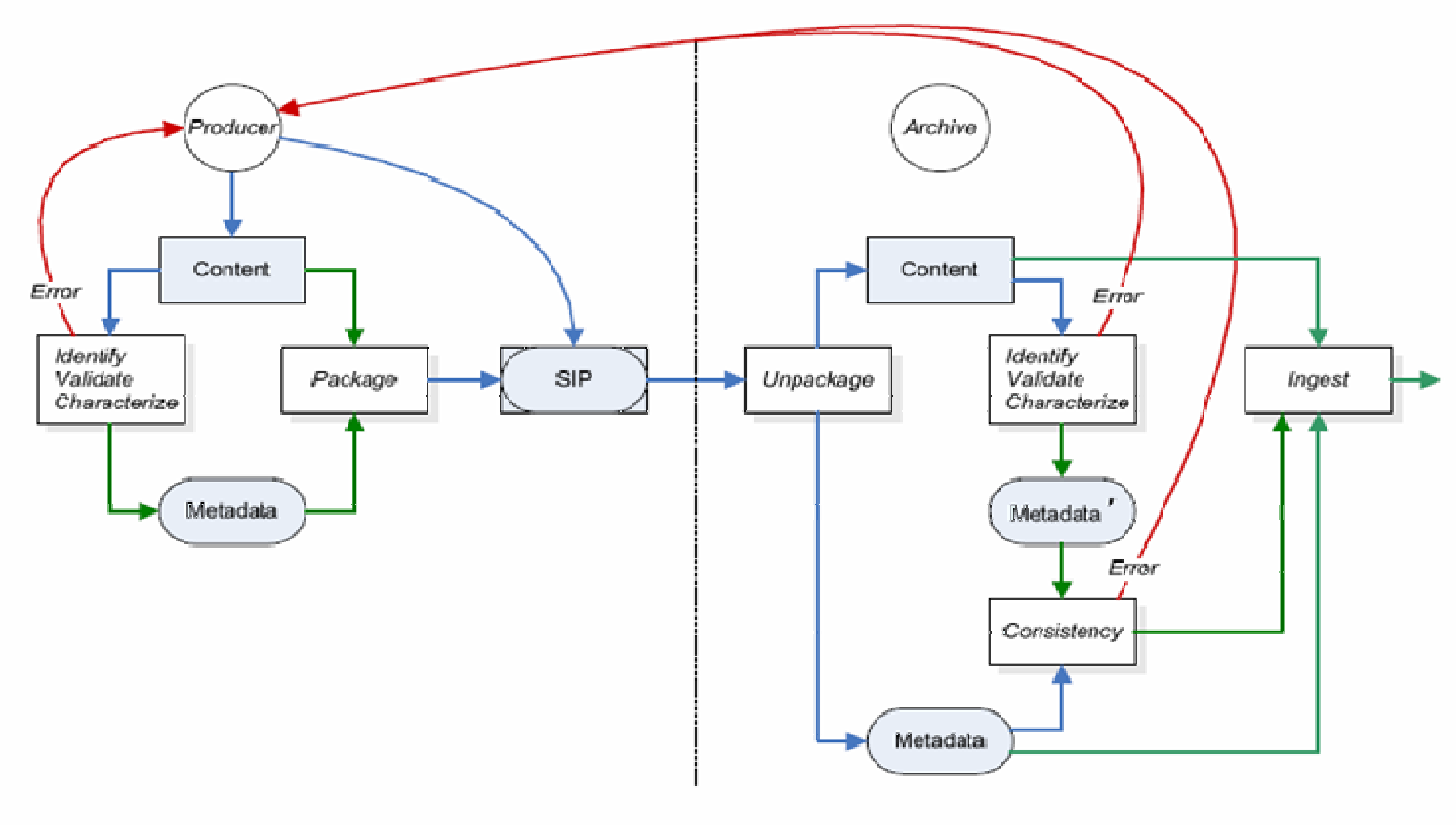
Characterisation

Why characterisation?

“Characterisation is an essential precursor to preservation. It provides the information required to make preservation planning decisions about digital objects, and to validate the results of preservation actions. “

(A. Brown: Developing Practical Approaches to Active Preservation, IJDC, 2007)

Why characterisation?



Source: S. Abrams: Automated Characterization in Preservation Workflows. *Tools and Trends*: *International Conference on Digital Preservation* Koninklijke Bibliotheek, 1-2 November 2007

Characterisation

What is subject to characterisation?

> “One essential process in digital preservation is to perform format characterization to extract technical metadata associated with each digital object in the preservation archival collection. The technical metadata are important attributes for understanding and managing the digital archival collections, especially for format monitoring and researching format transformation procedures.”

(C.C.H. Chou: Format Identification, Validation, Characterization and Transformation in DAITSS, [?2007])

Characterisation

What is subject to characterisation?

> “One essential process in digital preservation is to perform format characterization to extract technical metadata associated with each digital object in the preservation archival collection. The technical metadata are important attributes for understanding and managing the digital archival collections, especially for format monitoring and researching format transformation procedures.”

(C.C.H. Chou: Format Identification, Validation, Characterization and Transformation in DAITSS, [?2007])

Characterisation

What is subject to characterisation?

> “One essential process in digital preservation is to perform format characterization to extract technical metadata associated with each digital object in the preservation archival collection. The technical metadata are important attributes for understanding and managing the digital archival collections, especially for format monitoring and researching format transformation procedures.”

(C.C.H. Chou: Format Identification, Validation, Characterization and Transformation in DAITSS, [?2007])

About File Formats

What is a format?

011100110001110100011010…

- ❑ On a very basic level (storage level) digital content is nothing but binary data
- ❑ On the software level, digital content is stored as formatted data, i.e. as *meaningful* sequences of bytes

  → (*File) Format*
- ❑ On the most human-perceivable level it appears in a rendered form

How many file formats?

- PRONOM: ~ 550
- www.wotsit.org: ~ 900
- www.fileformat.info: 567
- www.fileinfo.com: > 3000 (file extensions)

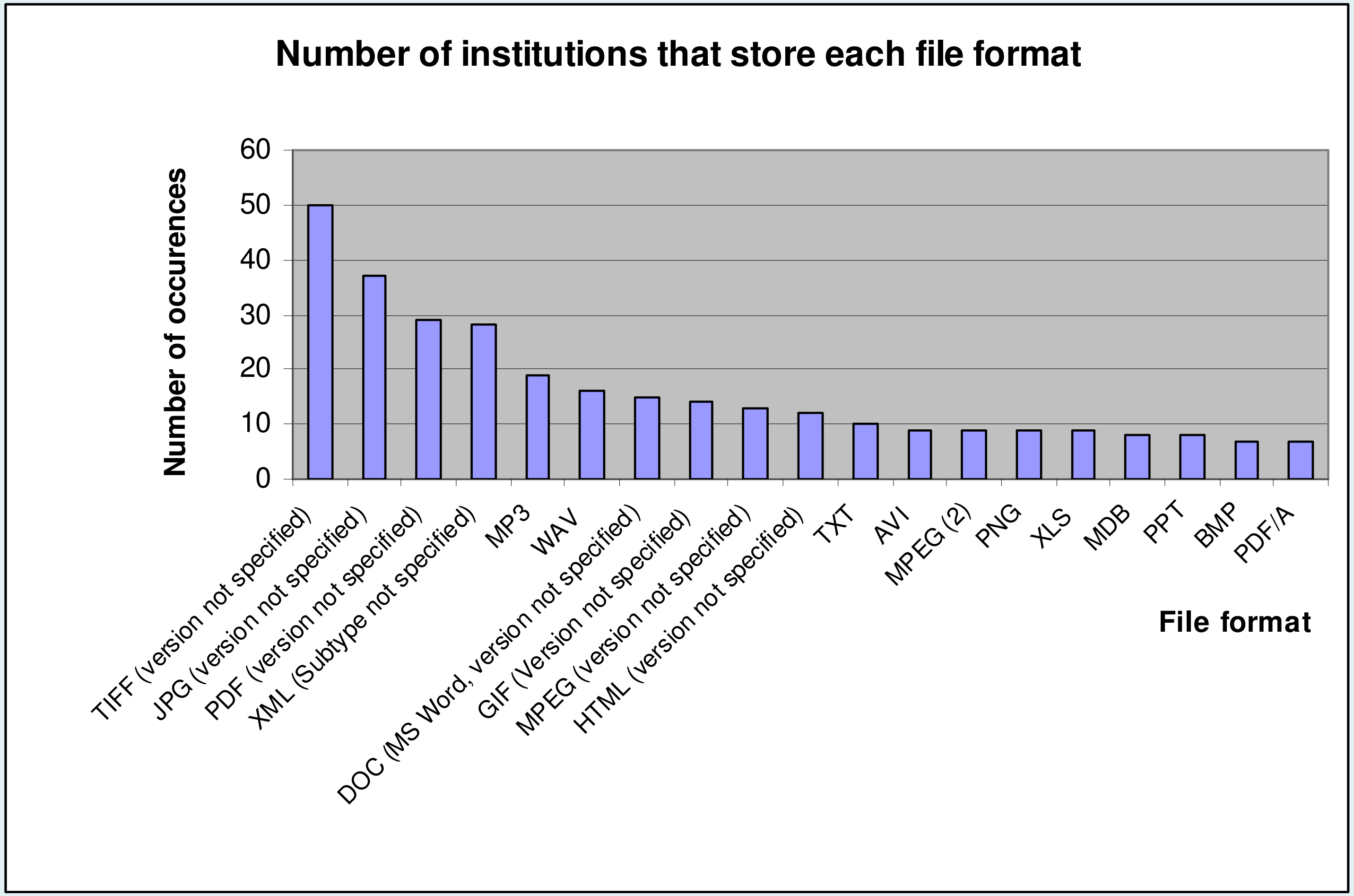
How many file formats can we find in institutions?

Planets internal study: “Gap analysis in tool provision”

- 76 institutions from 13 countries
- 137 different file formats (124 excl. versions)

Source: Planets internal report:
Gap analysis in tool provision (third version).

How many file formats are used more often?



Source: Planets internal report:
Gap analysis in tool provision (third version).

Suitability of formats for preservation (1)

High confidence	Medium confidence	Low confidence
❖TIFF (uncompressed) ❖ PNG (*.png)	❖ BMP (*.bmp) ❖ JPEG/JFIF (*.jpg) ❖JPEG2000 (prefer lossless or uncompressed) (*.jp2) ❖TIFF (compressed) ❖GIF (*.gif)	❖MrSID (*.sid) ❖TIFF (in Planar format) ❖FlashPix (*.fpx) ❖PhotoShop (*.psd) ❖All other raster image formats not listed here

Source: http://www.fcla.edu/digitalArchive/pdfs/recFormats.pdf

Suitability of formats for preservation (2)

High confidence	Medium confidence	Low confidence
❖ Plain text (encoding: ISO8859-1 - 9, UTF-8, UTF-16 with BOM) ❖ XML (includes XSD/XSL/XHTML, etc.; with included or accessible schema and character encoding explicitly specified) ❖ PDF/A-1 (ISO 19005-1)	❖ Cascading Style Sheets (*.css) ❖ DTD (*.dtd) ❖ PDF (*.pdf) (embedded fonts) ❖ Rich Text Format 1.x (*.rtf) ❖ HTML 4.x (include a DOCTYPE declaration) ❖ SGML (*.sgml) ❖ Open Office (*.sxw/*.odt) ❖ Office Open XML (*.docx)	❖PDF (*.pdf) (encrypted) ❖ Microsoft Word (*.doc) ❖ WordPerfect (*.wpd) ❖ DVI (*.dvi) ❖ All other text formats not listed here

Source: http://www.fcla.edu/digitalArchive/pdfs/recFormats.pdf

Suitability of formats for preservation (3)

High confidence	Medium confidence	Low confidence
❖AIFF (PCM) (*.aif, *.aiff) ❖ WAV (PCM) (*.wav)	❖SUN Audio (uncompressed) (*.au) ❖Standard MIDI (*.mid, *.midi) ❖Ogg Vorbis (*.ogg) ❖Free Lossless Audio Codec (*.flac) ❖ Advance Audio Coding (*.mp4, *.m4a, *.aac) ❖ MP3 (MPEG-1/2, Layer 3)(*.mp3)	❖AIFC (compressed) (*.aifc) ❖ NeXT SND (*.snd) ❖ RealNetworks 'Real Audio, (*.ra, *.rm, *.ram) ❖ Windows Media Audio ❖(*.wma) ❖WAV (compressed) (*.wav) ❖All other audio formats not listed here

Source: http://www.fcla.edu/digitalArchive/
pdfs/recFormats.pdf

Criteria for suitability

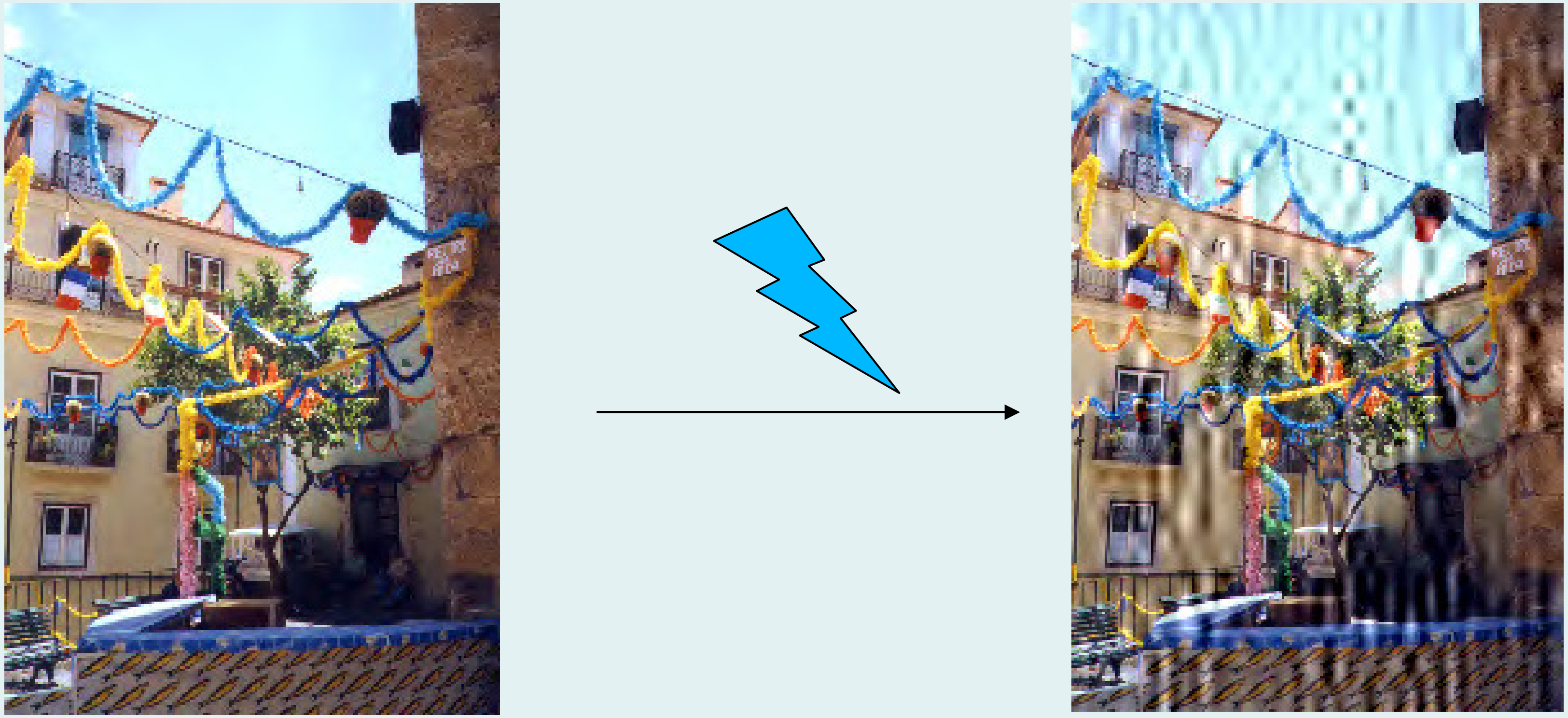
- Openess
- Adoption
- Complexity
- Technical protection mechanism
- Self-documentation
- Robustness
- Dependencies

(J. Rog, C. van Wijk: Evaluating File Formats for Long-term Preservation, iPres 2007)

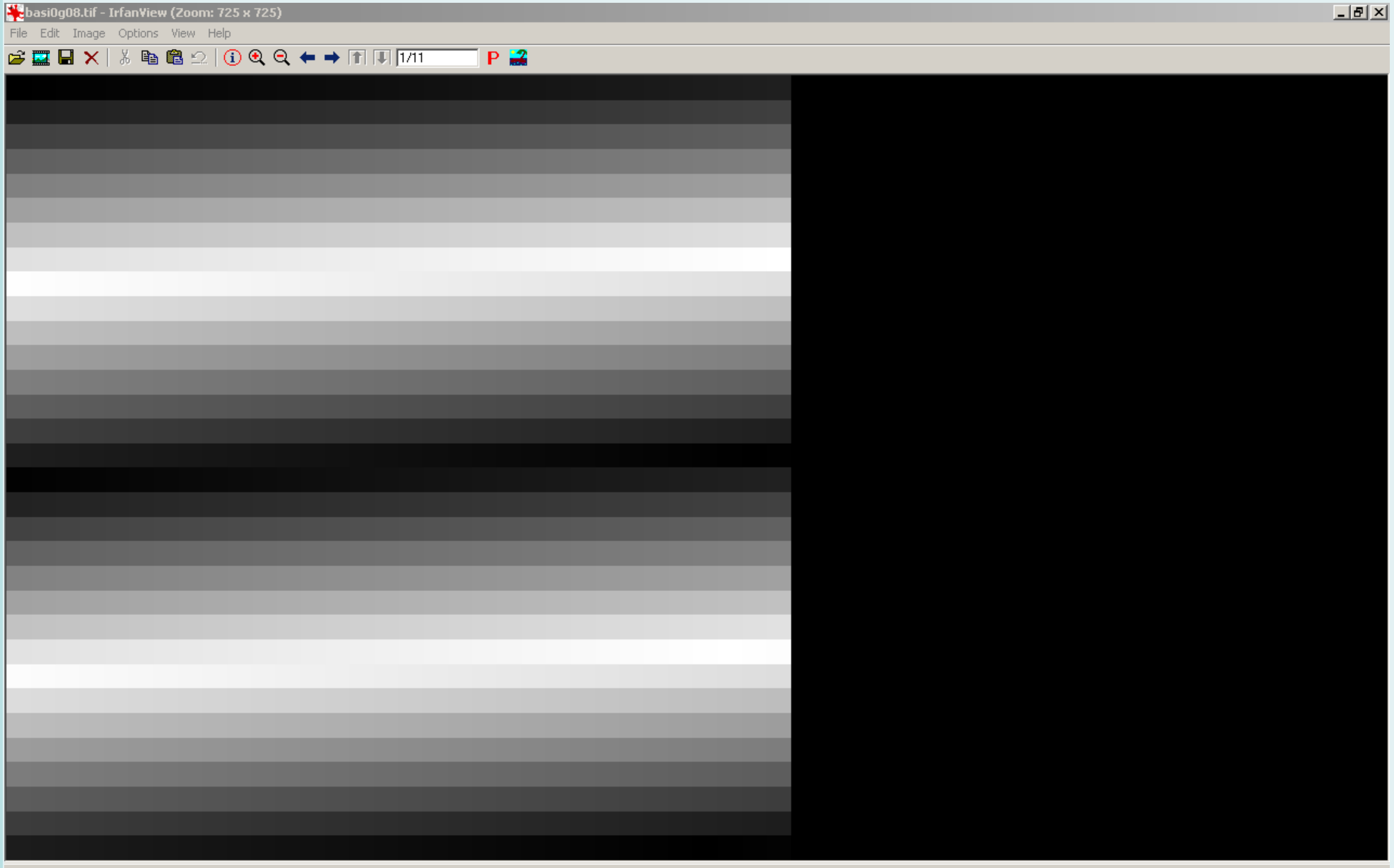
Robustness of Formats

Robustness

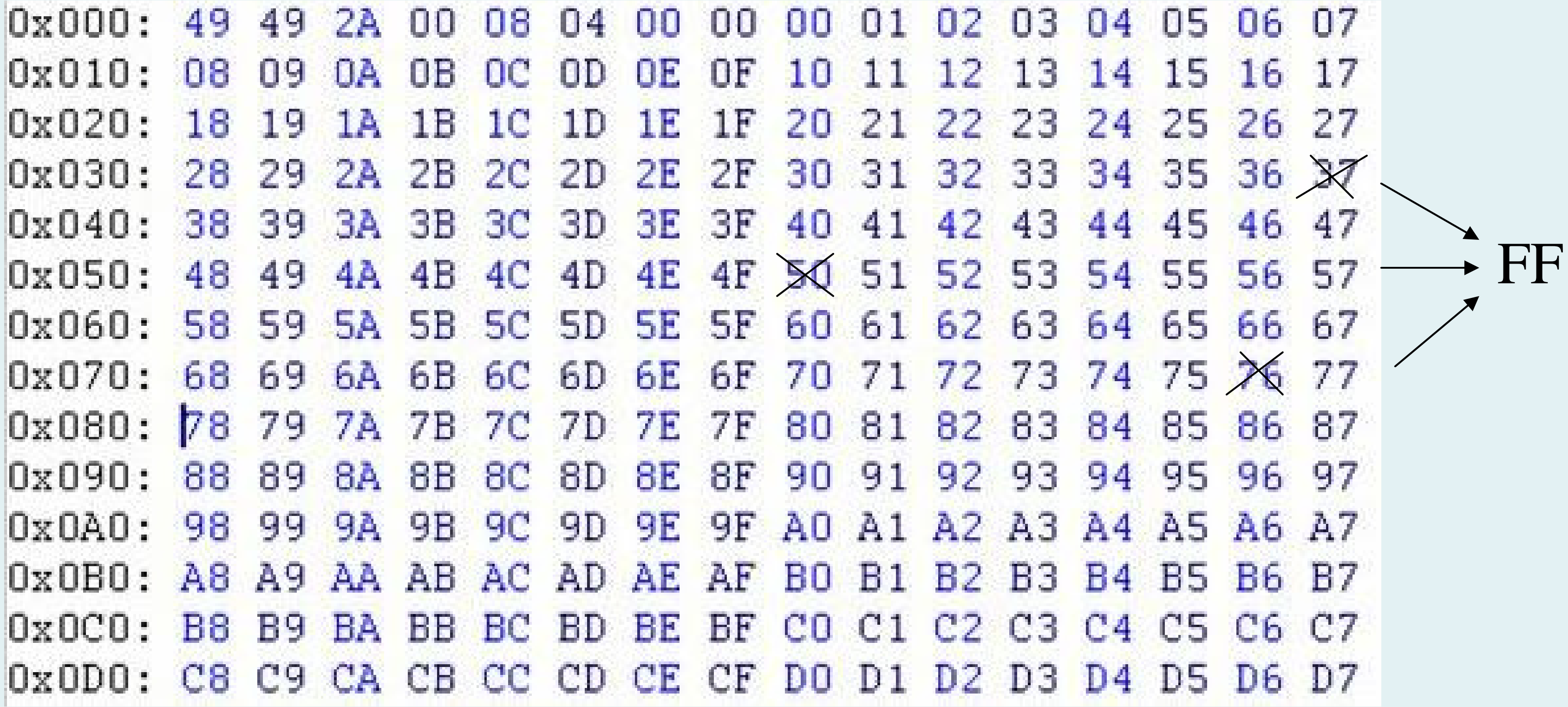
::= resilience of file formats against bit-stream corruption



What happens if data is corrupted in files?



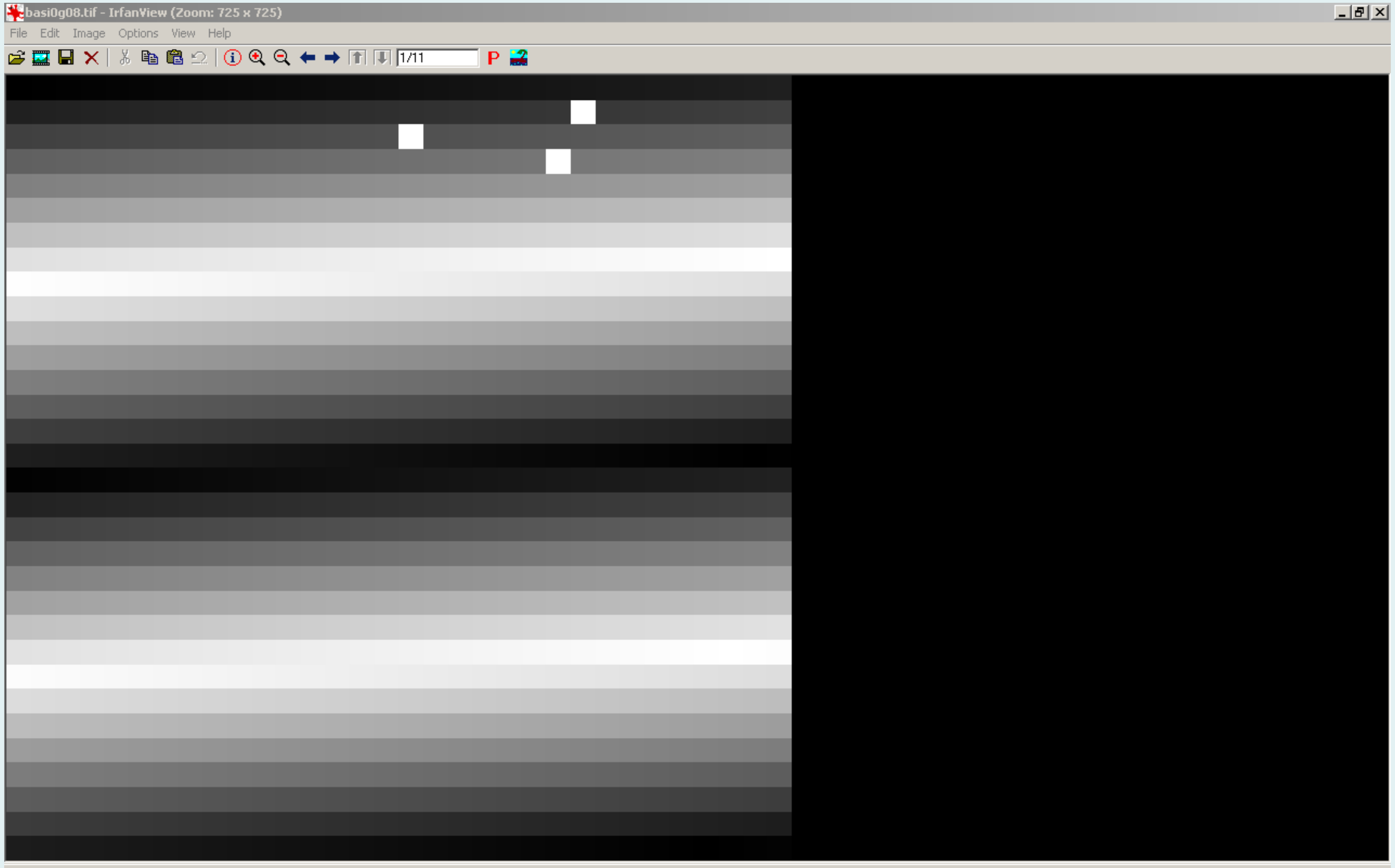
Testimage: Tiff, greyscale, 32x32 pixel, 8 bit per pixel

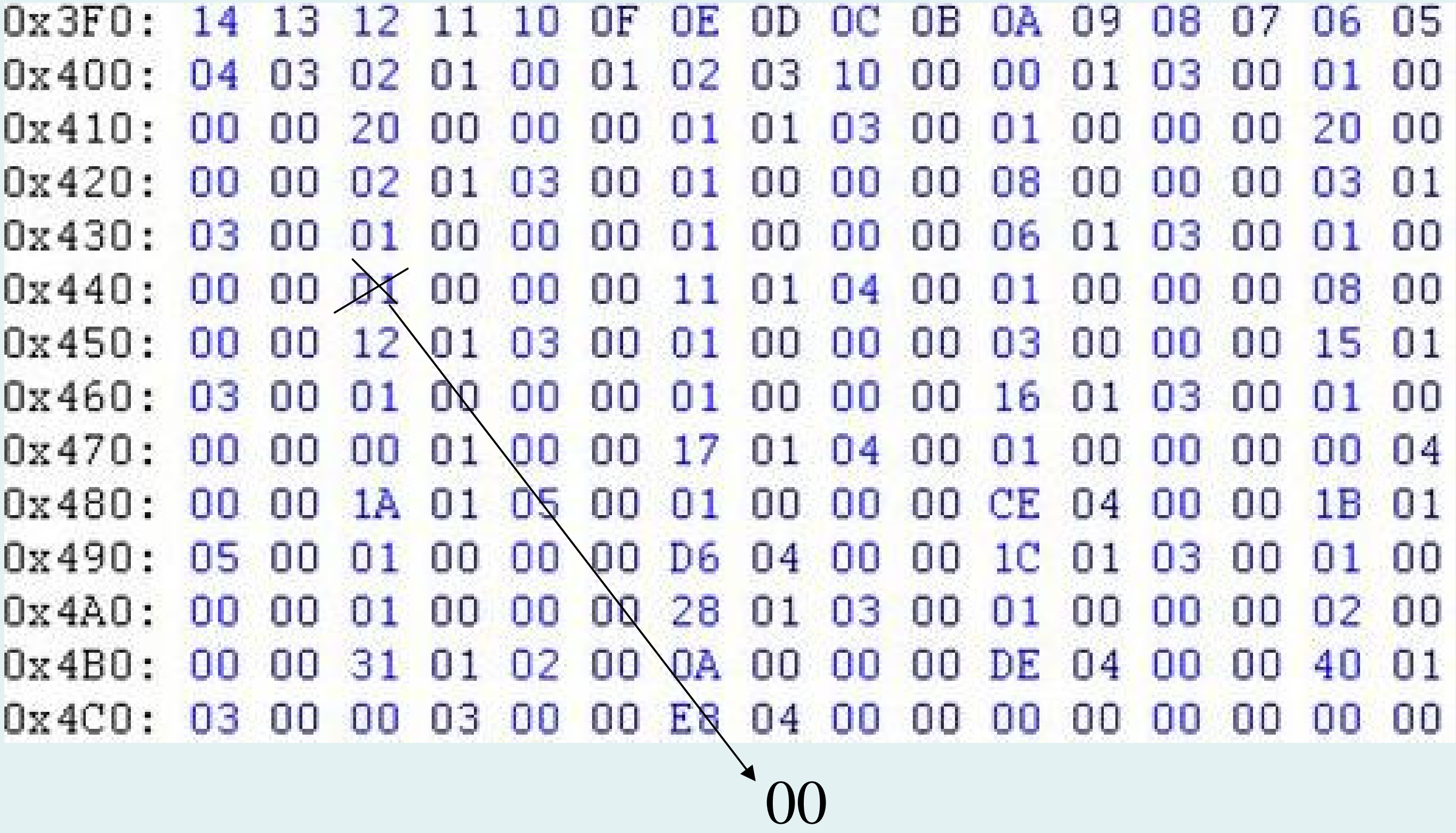


```
0x000: 49 49 2A 00 08 04 00 00 00 01 02 03 04 05 06 07
0x010: 08 09 0A 0B 0C 0D 0E 0F 10 11 12 13 14 15 16 17
0x020: 18 19 1A 1B 1C 1D 1E 1F 20 21 22 23 24 25 26 27
0x030: 28 29 2A 2B 2C 2D 2E 2F 30 31 32 33 34 35 36 ~~37~~
0x040: 38 39 3A 3B 3C 3D 3E 3F 40 41 42 43 44 45 46 47
0x050: 48 49 4A 4B 4C 4D 4E 4F ~~50~~ 51 52 53 54 55 56 57
0x060: 58 59 5A 5B 5C 5D 5E 5F 60 61 62 63 64 65 66 67
0x070: 68 69 6A 6B 6C 6D 6E 6F 70 71 72 73 74 75 ~~76~~ 77
0x080: 78 79 7A 7B 7C 7D 7E 7F 80 81 82 83 84 85 86 87
0x090: 88 89 8A 8B 8C 8D 8E 8F 90 91 92 93 94 95 96 97
0x0A0: 98 99 9A 9B 9C 9D 9E 9F A0 A1 A2 A3 A4 A5 A6 A7
0x0B0: A8 A9 AA AB AC AD AE AF B0 B1 B2 B3 B4 B5 B6 B7
0x0C0: B8 B9 BA BB BC BD BE BF C0 C1 C2 C3 C4 C5 C6 C7
0x0D0: C8 C9 CA CB CC CD CE CF D0 D1 D2 D3 D4 D5 D6 D7
```

First 224 bytes of testfile

Information loss: 1 byte data = = 1 Pixel

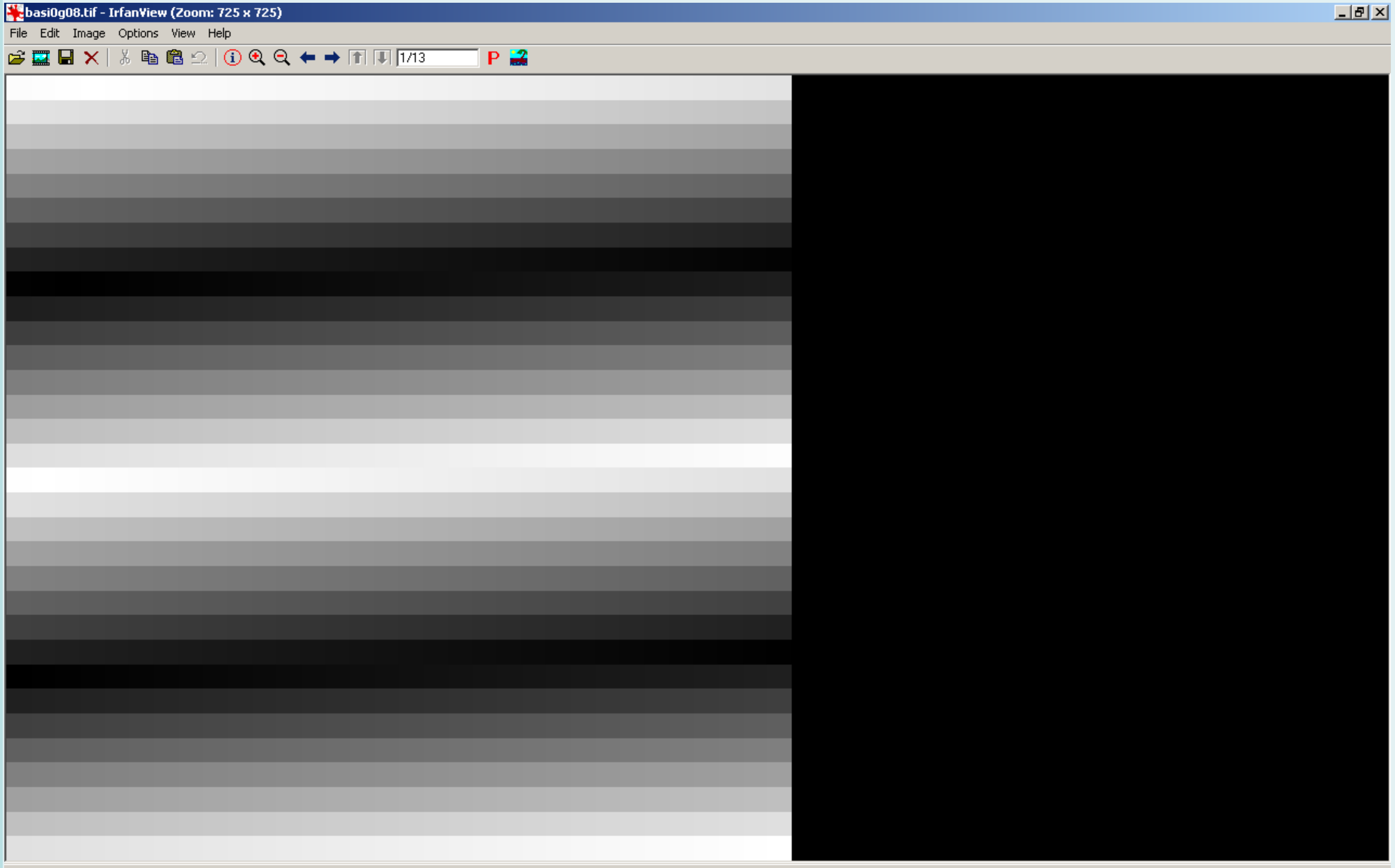


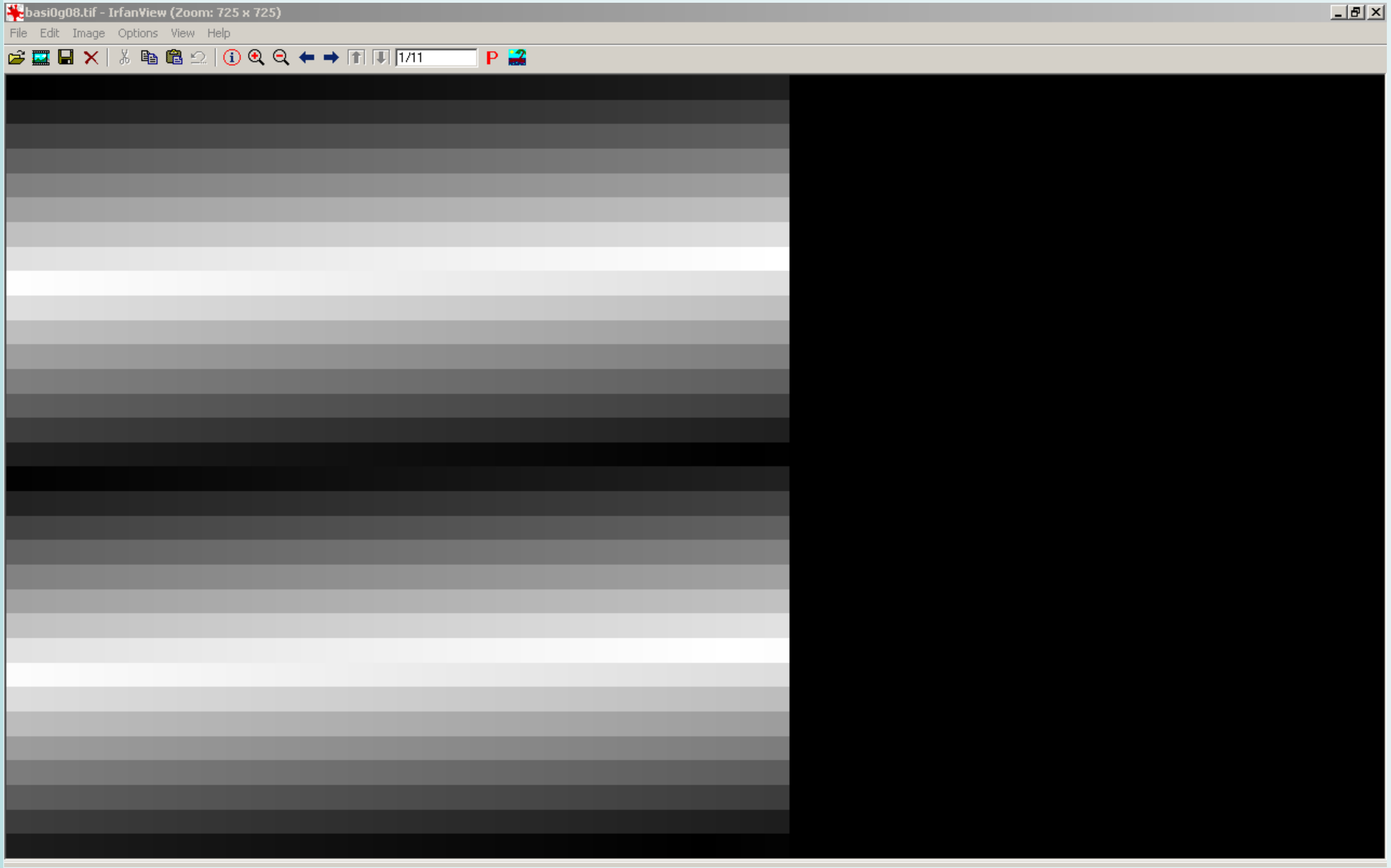


```
0x3F0: 14 13 12 11 10 0F 0E 0D 0C 0B 0A 09 08 07 06 05
0x400: 04 03 02 01 00 01 02 03 10 00 00 01 03 00 01 00
0x410: 00 00 20 00 00 00 01 01 03 00 01 00 00 00 20 00
0x420: 00 00 02 01 03 00 01 00 00 00 08 00 00 00 03 01
0x430: 03 00 01 00 00 00 01 00 00 00 06 01 03 00 01 00
0x440: 00 00 01 00 00 00 11 01 04 00 01 00 00 00 08 00
0x450: 00 00 12 01 03 00 01 00 00 00 03 00 00 00 15 01
0x460: 03 00 01 00 00 00 01 00 00 00 16 01 03 00 01 00
0x470: 00 00 00 01 00 00 17 01 04 00 01 00 00 00 00 04
0x480: 00 00 1A 01 05 00 01 00 00 00 CE 04 00 00 1B 01
0x490: 05 00 01 00 00 00 D6 04 00 00 1C 01 03 00 01 00
0x4A0: 00 00 01 00 00 00 28 01 03 00 01 00 00 00 02 00
0x4B0: 00 00 31 01 02 00 0A 00 00 00 DE 04 00 00 40 01
0x4C0: 03 00 00 03 00 00 E8 04 00 00 00 00 00 00 00 00
```

Part of the TIFF Image File Directory, Tag: Photometric Interpretation

1 bit changes == 100% information changed




basi0g08.tif - IrfanView (Zoom: 725 x 725)
File
Edit
Image
Options
View
Help
1/11
P

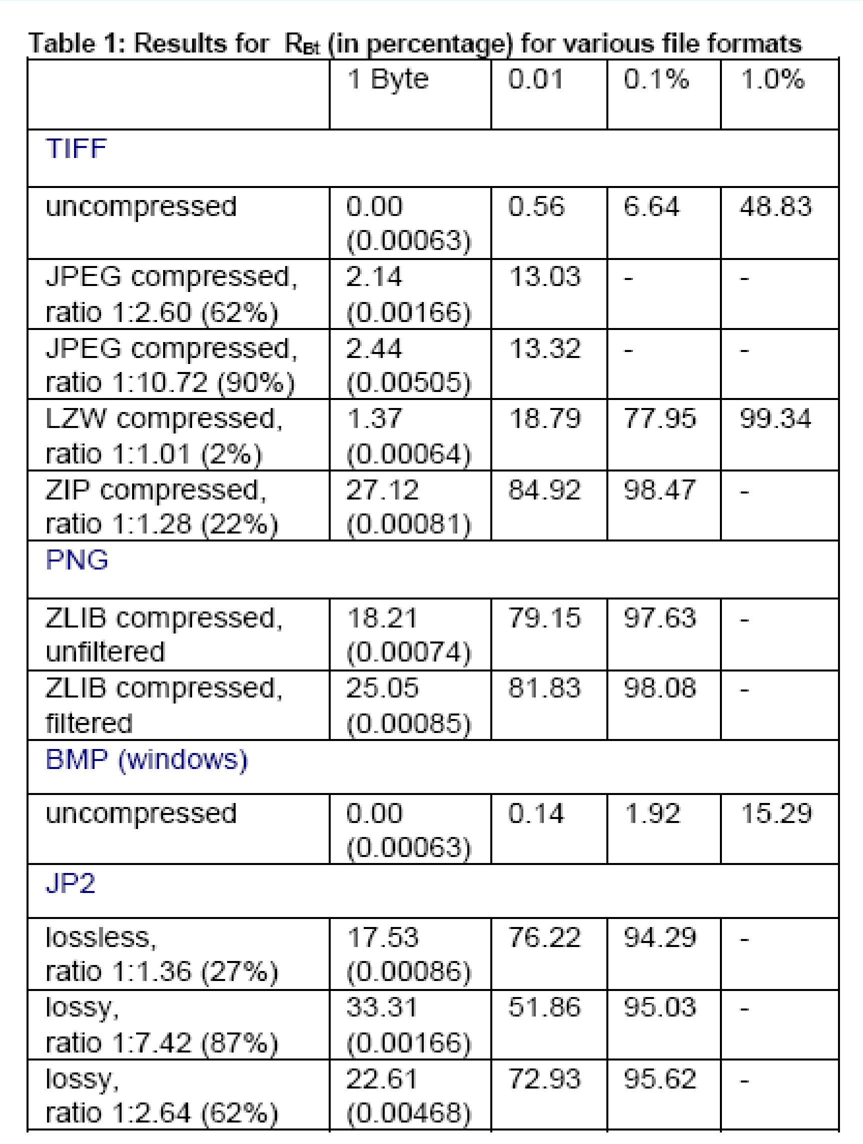


Table 1: Results for R_{Bt} (in percentage) for various file formats

	1 Byte	0.01	0.1%	1.0%
TIFF				
uncompressed	0.00 (0.00063)	0.56	6.64	48.83
JPEG compressed, ratio 1:2.60 (62%)	2.14 (0.00166)	13.03	-	-
JPEG compressed, ratio 1:10.72 (90%)	2.44 (0.00505)	13.32	-	-
LZW compressed, ratio 1:1.01 (2%)	1.37 (0.00064)	18.79	77.95	99.34
ZIP compressed, ratio 1:1.28 (22%)	27.12 (0.00081)	84.92	98.47	-
PNG				
ZLIB compressed, unfiltered	18.21 (0.00074)	79.15	97.63	-
ZLIB compressed, filtered	25.05 (0.00085)	81.83	98.08	-
BMP (windows)				
uncompressed	0.00 (0.00063)	0.14	1.92	15.29
JP2				
lossless, ratio 1:1.36 (27%)	17.53 (0.00086)	76.22	94.29	-
lossy, ratio 1:7.42 (87%)	33.31 (0.00166)	51.86	95.03	-
lossy, ratio 1:2.64 (62%)	22.61 (0.00468)	72.93	95.62	-

Categories of characteristics

What is subject to characterisation?

“One essential process in digital preservation is to perform format characterization to extract technical metadata associated with each digital object in the preservation archival collection. The technical metadata are important attributes for understanding and managing the digital archival collections, especially for format monitoring and researching format transformation procedures.”

(C.C.H. Chou: Format Identification, Validation, Characterization and Transformation in DAITSS, [?2007])

Non-technical characteristics (“associated metadata”)

What’s the name of the object?

Which software created the object?

Who holds the intellectual rights for the object?

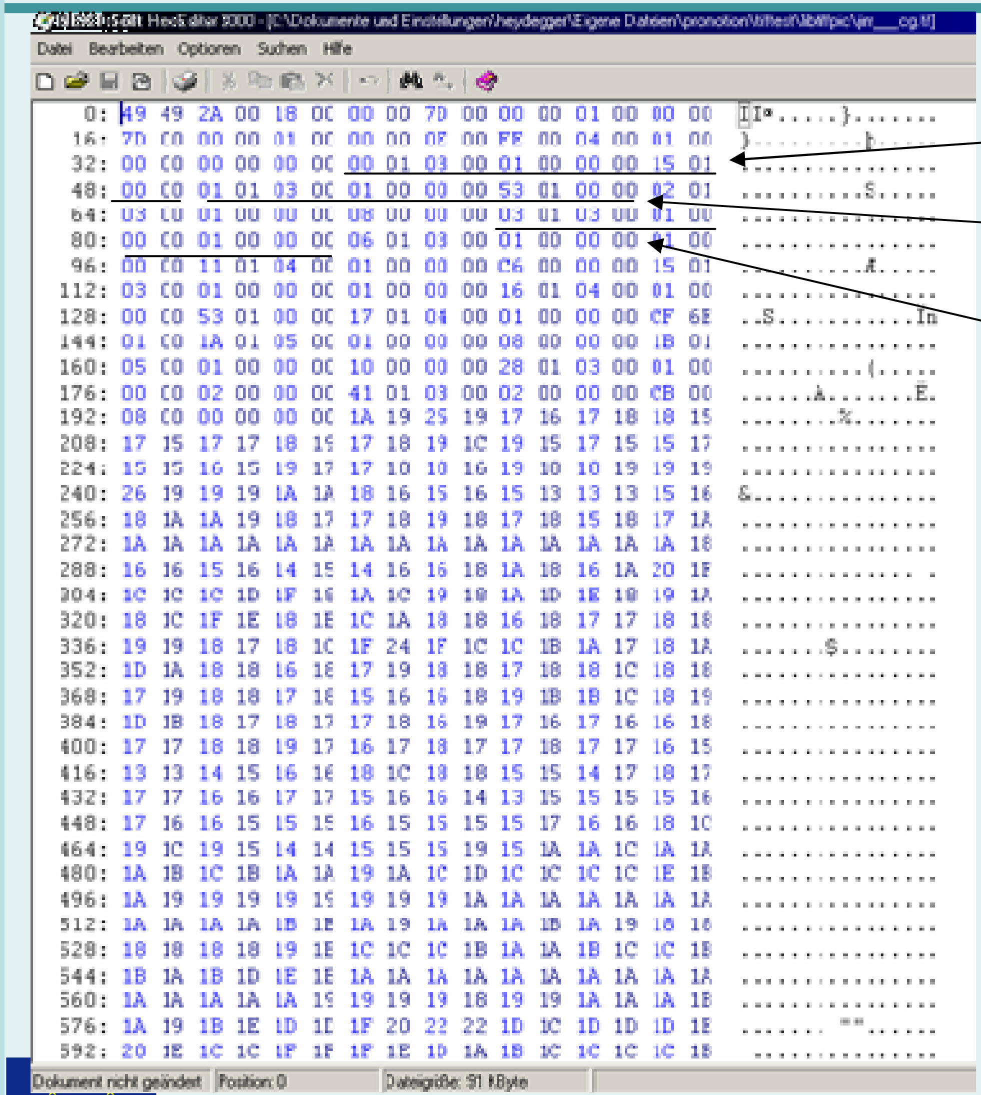
When was the object modified for the last time?

Which collection does the object belong to?

Where is the object located in our repository?

…

Technical characteristics


Hex Editor 2000 - [C:\Dokumente und Einstellungen\heydegger\Eigene Dateien\promotion\tiftest\libtiffpic\jim___cg.tif]
Datei
Bearbeiten
Optionen
Suchen
Hilfe
Dokument nicht geändert
Position: 0
Dateigröße: 91 KByte

```
  0: 49 49 2A 00 18 0C 00 00 7D 00 00 00 01 00 00 00   II*.....}.......
 16: 7D C0 00 00 01 0C 00 00 0F 00 FE 00 04 00 01 00   }.........þ.....
 32: 00 C0 00 00 00 0C 00 01 03 00 01 00 00 00 15 01   ................
 48: 00 C0 01 01 03 0C 01 00 00 00 53 01 00 00 02 01   ..........S.....
 64: 03 C0 01 00 00 0C 08 00 00 00 03 01 03 00 01 00   ................
 80: 00 C0 01 00 00 0C 06 01 03 00 01 00 00 00 01 00   ................
 96: 00 C0 11 01 04 0C 01 00 00 00 C6 00 00 00 15 01   ..........Æ.....
112: 03 C0 01 00 00 0C 01 00 00 00 16 01 04 00 01 00   ................
128: 00 C0 53 01 00 0C 17 01 04 00 01 00 00 00 CF 6E   ..S...........Ïn
144: 01 C0 1A 01 05 0C 01 00 00 00 08 00 00 00 1B 01   ................
160: 05 C0 01 00 00 0C 10 00 00 00 28 01 03 00 01 00   ..........(.....
176: 00 C0 02 00 00 0C 41 01 03 00 02 00 00 00 CB 00   ......A.......Ë.
192: 08 C0 00 00 00 0C 1A 19 25 19 17 16 17 18 18 15   ........%.......
208: 17 15 17 17 18 19 17 18 19 1C 19 15 17 15 15 17   ................
224: 15 15 16 15 19 17 17 10 10 16 19 10 10 19 19 19   ................
240: 26 19 19 19 1A 1A 18 16 15 16 15 13 13 13 15 16   &...............
256: 18 1A 1A 19 18 17 17 18 19 18 17 18 15 18 17 1A   ................
272: 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 18   ................
288: 16 16 15 16 14 15 14 16 16 18 1A 18 16 1A 20 1F   ............... .
304: 1C 1C 1C 1D 1F 16 1A 1C 19 18 1A 1D 1E 18 19 1A   ................
320: 18 1C 1F 1E 18 1E 1C 1A 18 18 16 18 17 17 18 18   ................
336: 19 19 18 17 18 1C 1F 24 1F 1C 1C 1B 1A 17 18 1A   .......$........
352: 1D 1A 18 18 16 16 17 19 18 18 17 18 18 1C 18 18   ................
368: 17 19 18 18 17 16 15 16 16 18 19 1B 1B 1C 18 19   ................
384: 1D 1B 18 17 18 17 17 18 16 19 17 16 17 16 16 18   ................
400: 17 17 18 18 19 17 16 17 18 17 17 18 17 17 16 15   ................
416: 13 13 14 15 16 16 18 1C 18 18 15 15 14 17 18 17   ................
432: 17 17 16 16 17 17 15 16 16 14 13 15 15 15 15 16   ................
448: 17 16 16 15 15 15 16 15 15 15 15 17 16 16 18 1C   ................
464: 19 1C 19 15 14 14 15 15 15 19 15 1A 1A 1C 1A 1A   ................
480: 1A 1B 1C 1B 1A 1A 19 1A 1C 1D 1C 1C 1C 1C 1E 1B   ................
496: 1A 19 19 19 19 19 19 19 19 1A 1A 1A 1A 1A 1A 1A   ................
512: 1A 1A 1A 1A 1B 1E 1A 19 1A 1A 1A 1B 1A 19 18 18   ................
528: 18 18 18 18 19 1E 1C 1C 1C 1B 1A 1A 1B 1C 1C 1B   ................
544: 1B 1A 1B 1D 1E 1E 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A   ................
560: 1A 1A 1A 1A 1A 19 19 19 19 18 19 19 1A 1A 1A 1B   ................
576: 1A 19 1B 1E 1D 1D 1F 20 22 22 1D 1C 1D 1D 1D 1E   ........""......
592: 20 1E 1C 1C 1F 1F 1F 1E 1D 1A 1B 1C 1C 1C 1C 1B   ................
```

Image width: 277

Image length: 339

Compression: uncompressed

ImageLength

The number of rows of pixels in the image.

Tag = 257 (101.H)

Type = SHORT or LONG

N = 1

No default. See also ImageWidth.

ImageWidth

The number of columns in the image, i.e., the number of pixels per row.

Tag = 256 (100.H)

Type = SHORT or LONG

N = 1

No default. See also ImageLength.

Categories of characteristics

- Significant characteristics:

= Those properties which are essential for keeping the integrity of the object

→ Significant properties are always of technical nature

Lessons learnt so far

- Characterisation is an essential part within an overall preservation framework.
- File Format is the central concept for representation of digital content.
- A Format describes the characteristics of objects.
- There is a huge amount of formats but only a couple of them are actually suitable for preservation.

XCL: Goals

- Support preservation planning framework
- Support a specific preservation action task: Evaluation of file format conversion
- Develop a more abstract model for extraction of characteristics (syn. properties) from files
- Develop tools which use this model in order to enable characterisation in an efficiently, i.e. in an automated way

XCL: Goals

- In practice:

  - Develop an „eXtensible Characterisation Definition Language“ (XCDL), able to describe the content of digital objects (=1 + n more files), processible by a software tool for further analysis.

  - Develop an „eXtensible Characterisation Extraction Language“ (XCEL), able to describe any machine readable format in a formal language, processible by a software tool for extraction of content as XCDL.

XCL: Goals

- Support preservation planning framework
- Support a specific preservation action task: Evaluation of file format conversion
- Develop a more abstract model for extraction of characteristics (syn. properties) from files
- Develop tools which use this model in order to enable characterisation in an efficiently, i.e. in an automated way

CEI
CENTRAL EUROPEAN INITIATIVE

planets

Why automate?

Assumption:
Preservation is only feasible, if the content of two digital objects can be compared without human intervention.

Why automate?

1 million objects: use *five minutes* for each

== 416 666.7 hours

== 52 803.4 8-hour days for a Human

CEI
CENTRAL EUROPEAN INITIATIVE

planets

Why automate?

1 million objects: use *one second* for each.

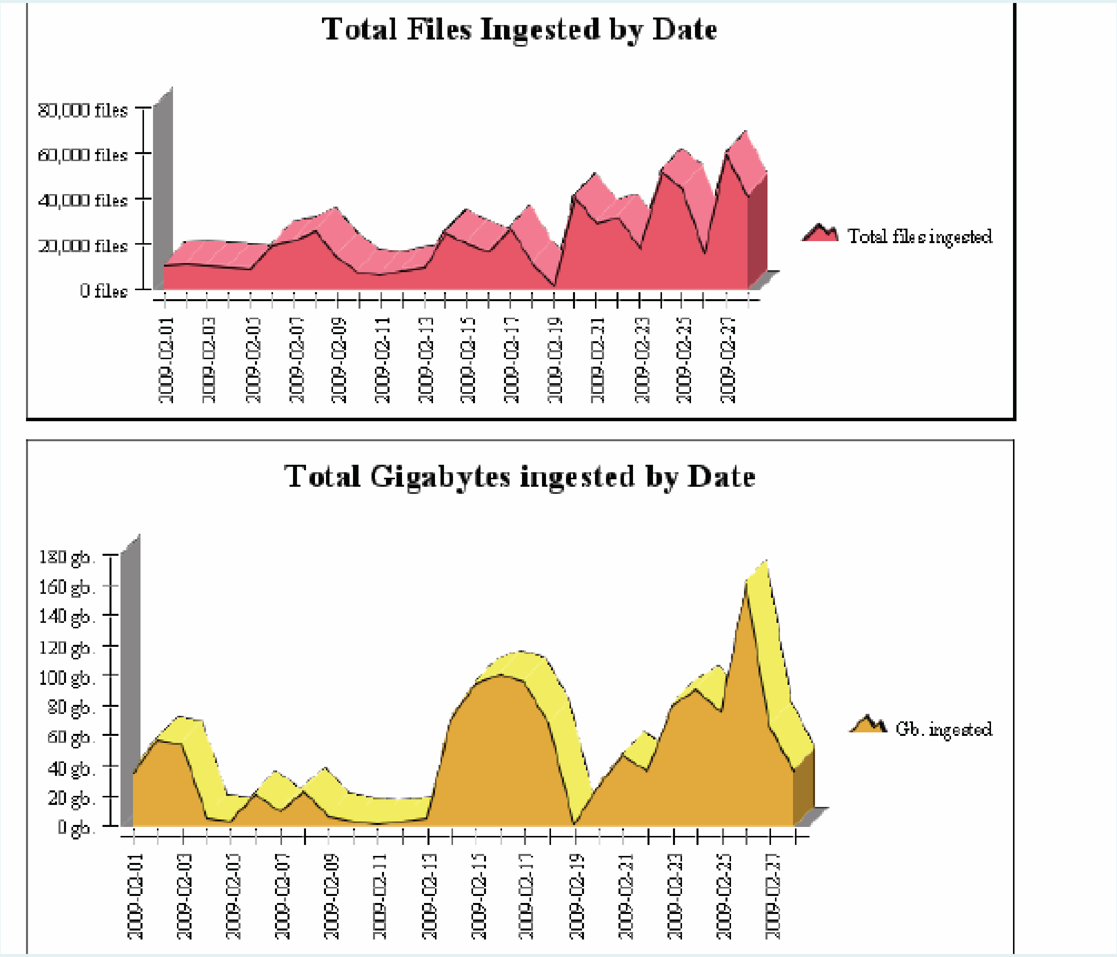
== 16666.7 minutes == 277.8 hours

== 11.57 working days of a computer

== 34.7 8-hour days for a Human

== 7 working weeks

Why automate?

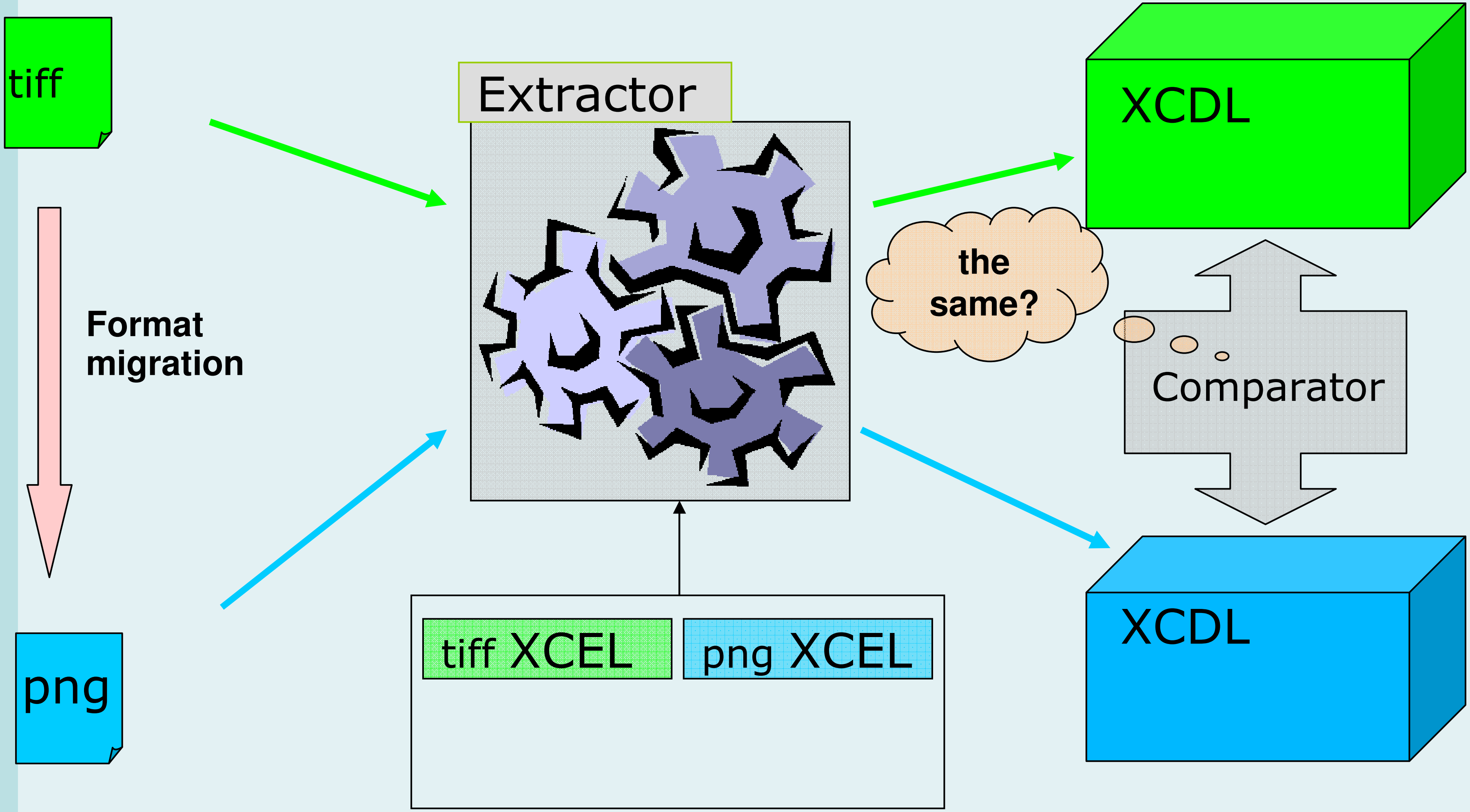


Source: http://www.fcla.edu/digitalArchive/pdfs/reports/ingest_stats_February_09.pdf

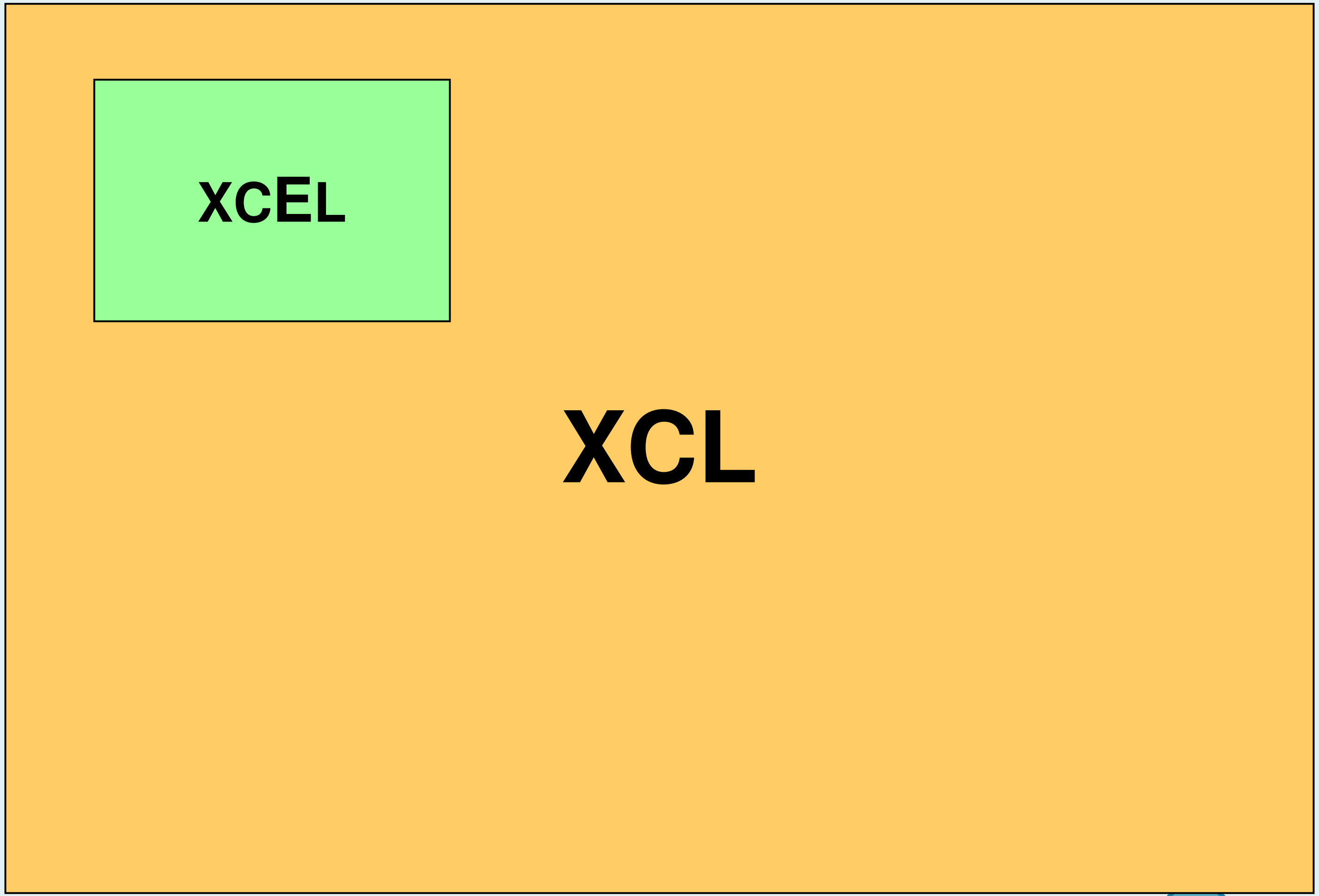
XCL: Goals

- Support preservation planning framework
- Support a specific preservation action task: Evaluation of file format conversion
- Develop a more abstract model for extraction of characteristics (syn. properties) from files
- Develop tools which use this model in order to enable characterisation in an efficiently, i.e. in an automated way

XCL: Main application: Evaluation of format conversion

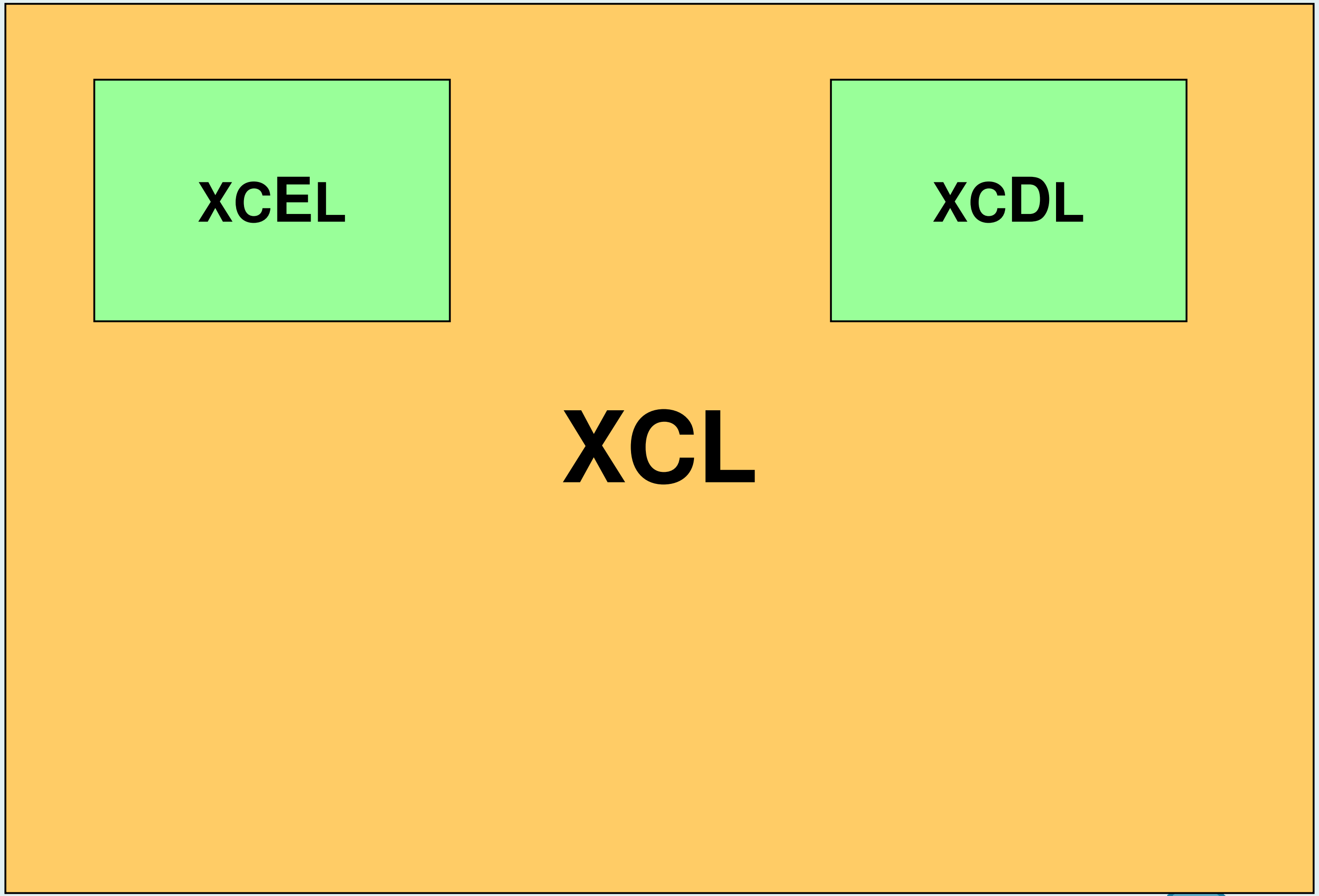


XCL: Architecture



XCEL

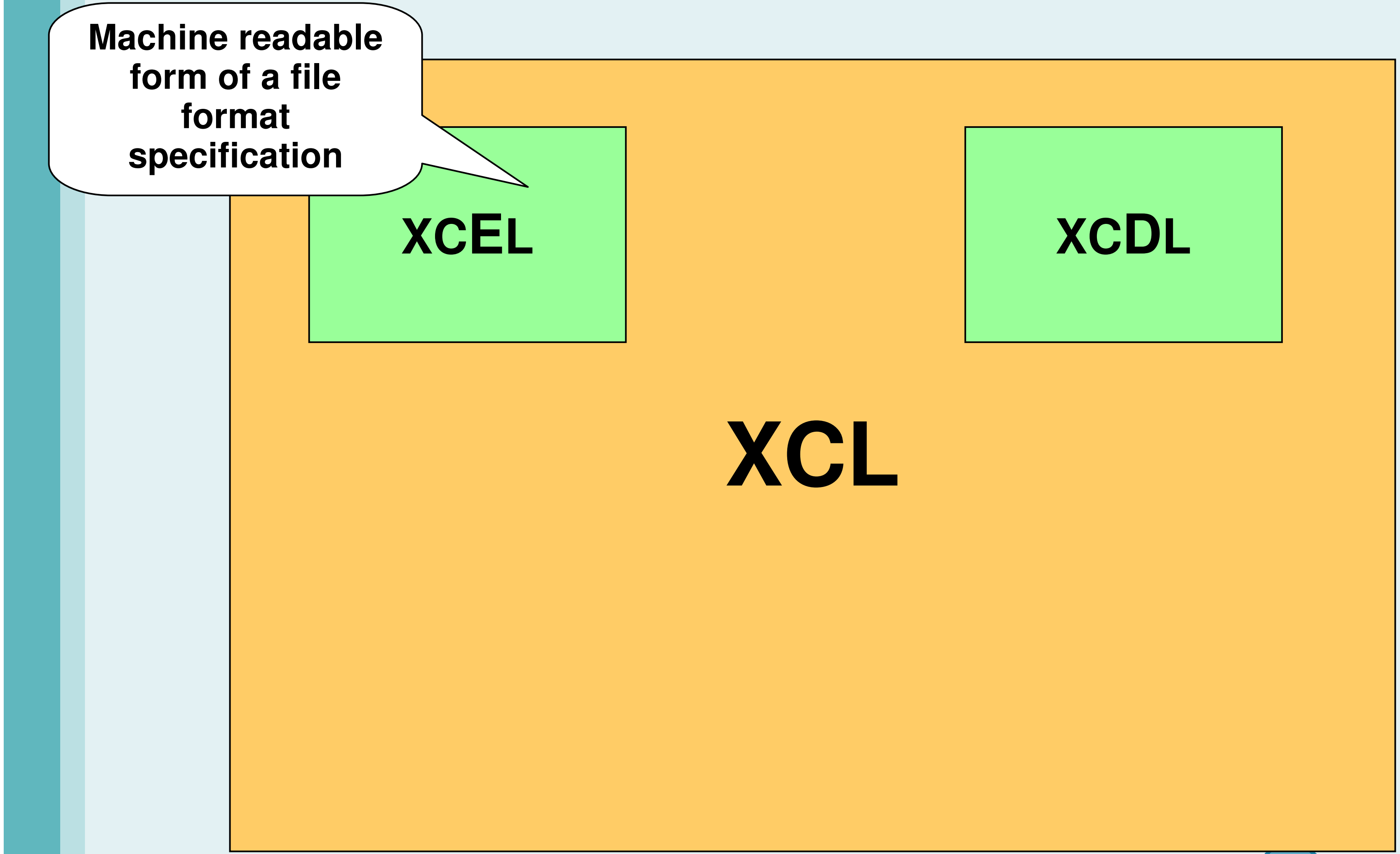
XCL



XCEL

XCDL

XCL

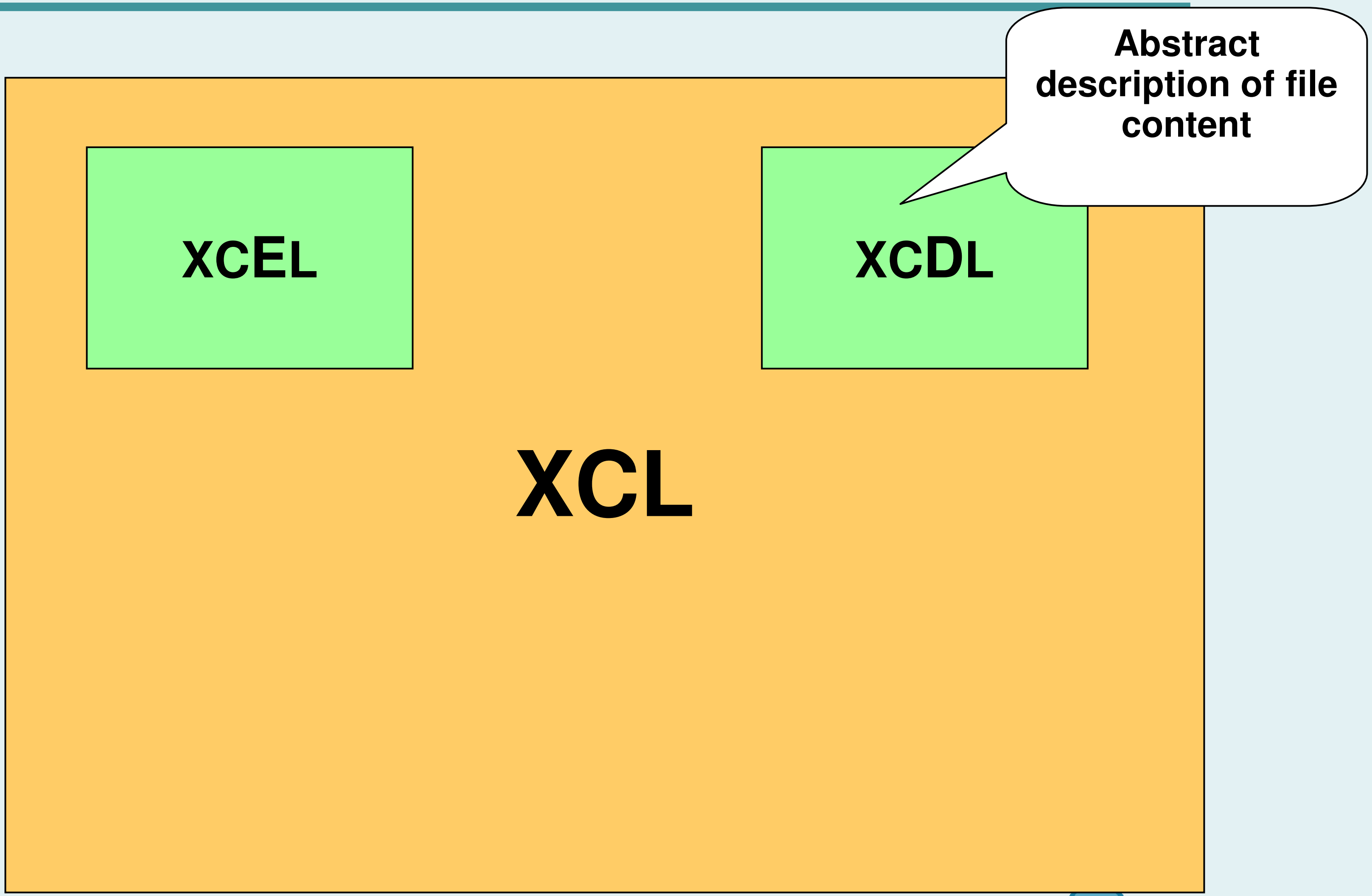


Machine readable form of a file format specification

XCEL

XCDL

XCL

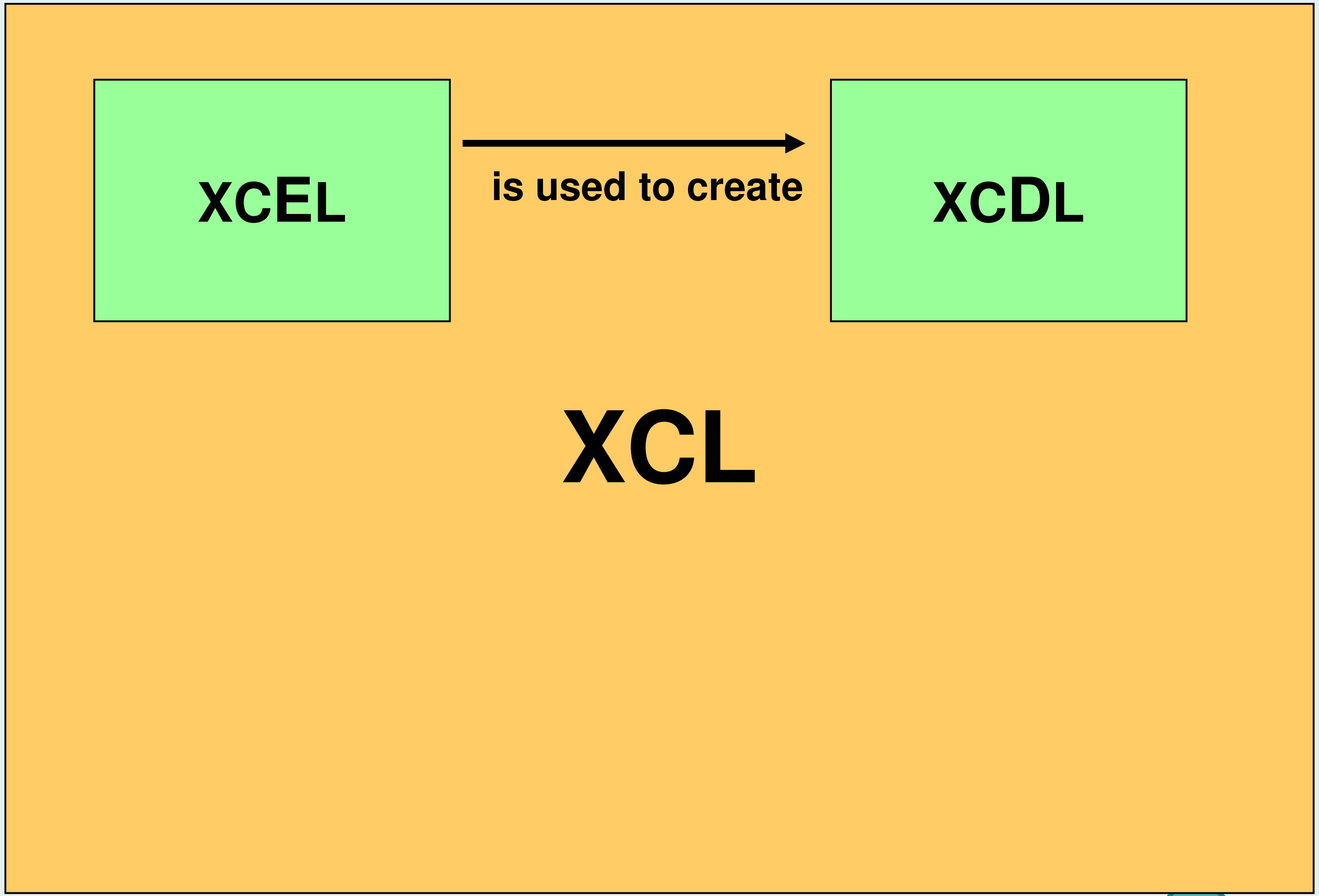


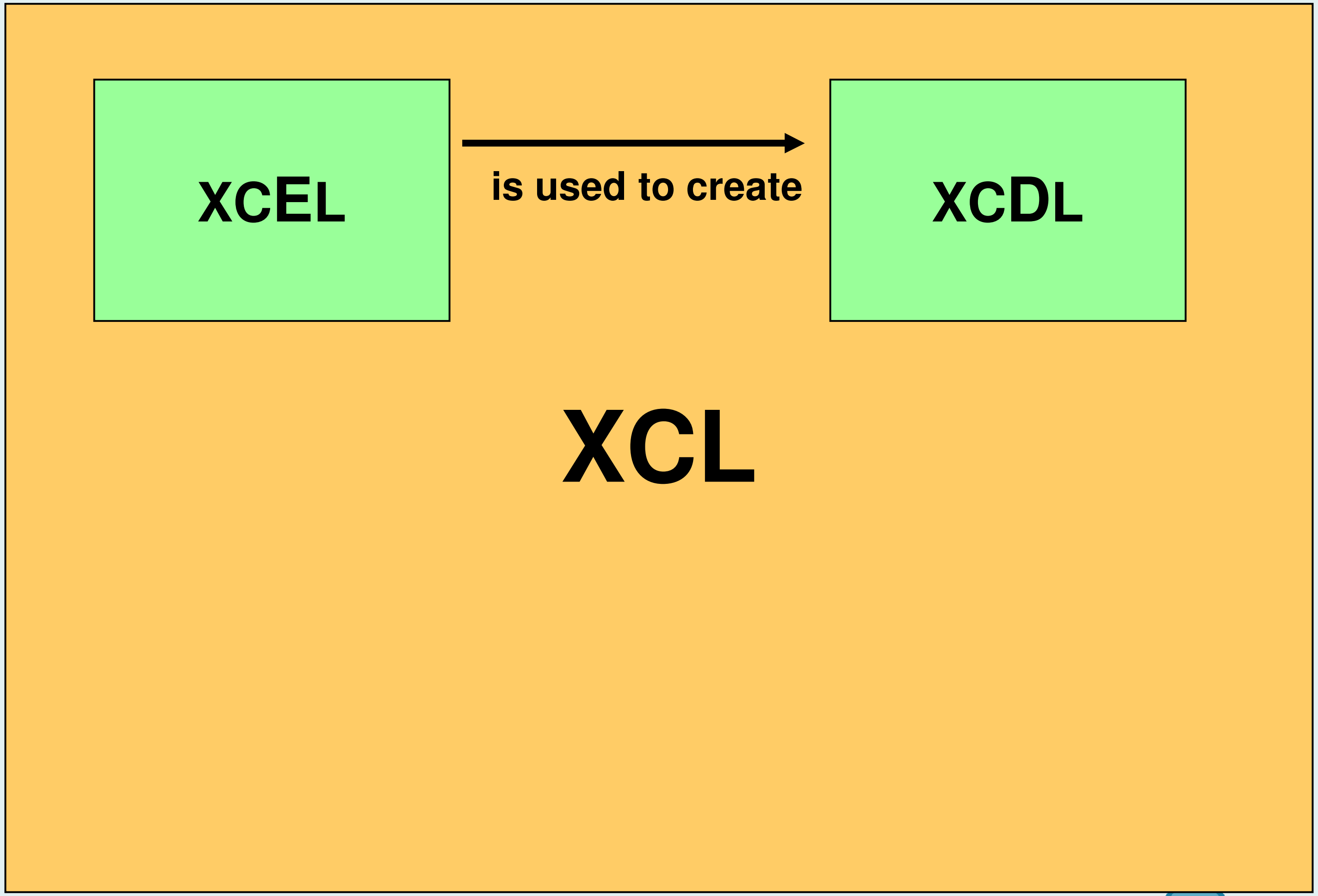
XCEL

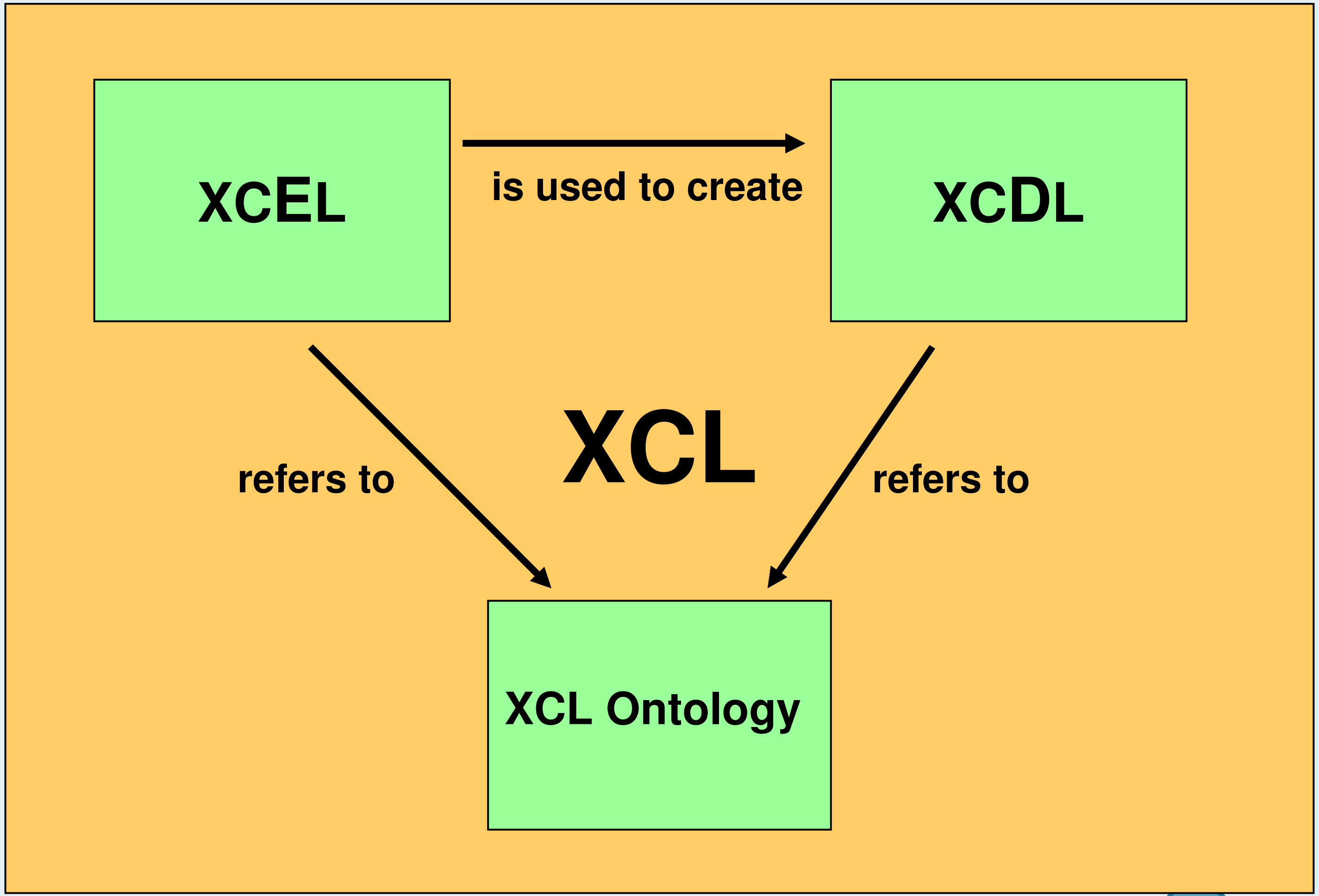
XCDL

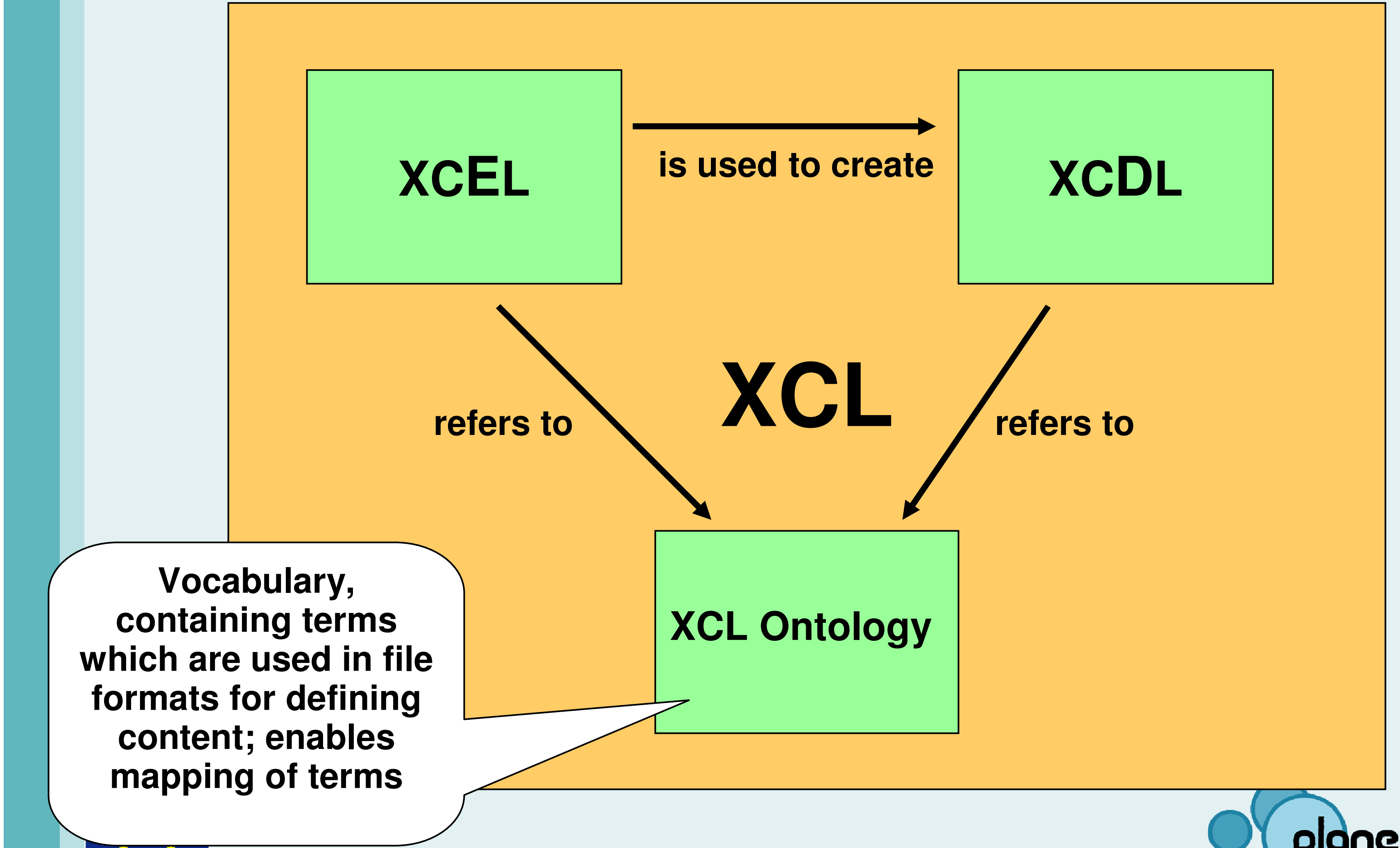
XCL

Abstract description of file content


XCEL
is used to create
XCDL
XCL


XCEL
is used to create
XCDL
XCL


XCEL
is used to create
XCDL
XCL
refers to
refers to
XCL Ontology



XCEL

is used to create

XCDL

refers to

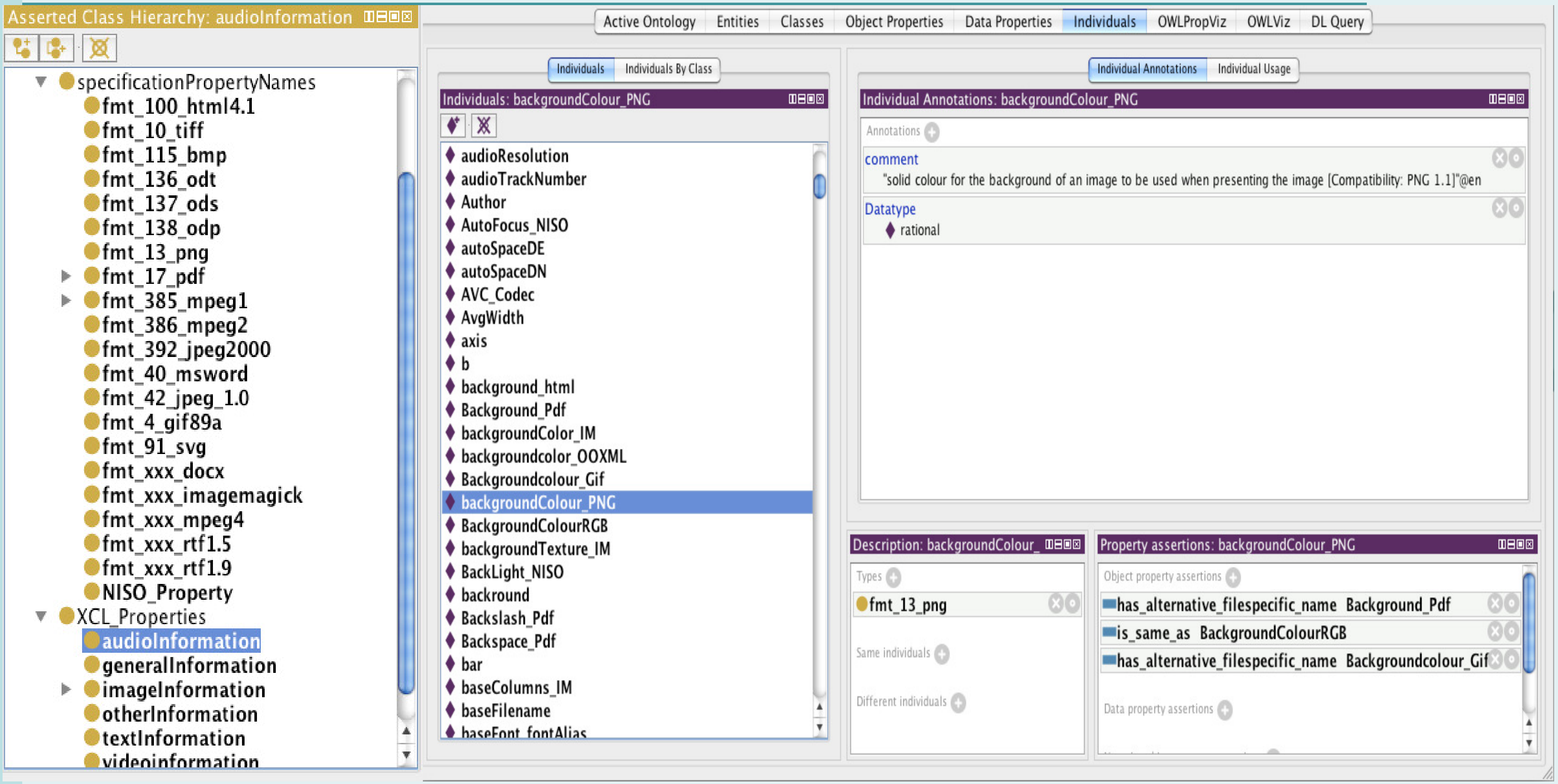
XCL

refers to

XCL Ontology

Vocabulary, containing terms which are used in file formats for defining content; enables mapping of terms

The Ontology



Asserted Class Hierarchy: audioInformation

- specificationPropertyNames
  - fmt_100_html4.1
  - fmt_10_tiff
  - fmt_115_bmp
  - fmt_136_odt
  - fmt_137_ods
  - fmt_138_odp
  - fmt_13_png
  - fmt_17_pdf
  - fmt_385_mpeg1
  - fmt_386_mpeg2
  - fmt_392_jpeg2000
  - fmt_40_msword
  - fmt_42_jpeg_1.0
  - fmt_4_gif89a
  - fmt_91_svg
  - fmt_xxx_docx
  - fmt_xxx_imagemagick
  - fmt_xxx_mpeg4
  - fmt_xxx_rtf1.5
  - fmt_xxx_rtf1.9
  - NISO_Property
- XCL_Properties
  - audioInformation
  - generalInformation
  - imageInformation
  - otherInformation
  - textInformation
  - videoinformation

Individuals: backgroundColour_PNG

- audioResolution
- audioTrackNumber
- Author
- AutoFocus_NISO
- autoSpaceDE
- autoSpaceDN
- AVC_Codec
- AvgWidth
- axis
- b
- background_html
- Background_Pdf
- backgroundColor_IM
- backgroundcolor_OOXML
- Backgroundcolour_Gif
- backgroundColour_PNG
- BackgroundColourRGB
- backgroundTexture_IM
- BackLight_NISO
- backround
- Backslash_Pdf
- Backspace_Pdf
- bar
- baseColumns_IM
- baseFilename
- baseFont_fontAlias

Individual Annotations: backgroundColour_PNG

Annotations

comment
"solid colour for the background of an image to be used when presenting the image [Compatibility: PNG 1.1]"@en

Datatype
rational

Description: backgroundColour_

Types
fmt_13_png

Same individuals

Different individuals

Property assertions: backgroundColour_PNG

Object property assertions

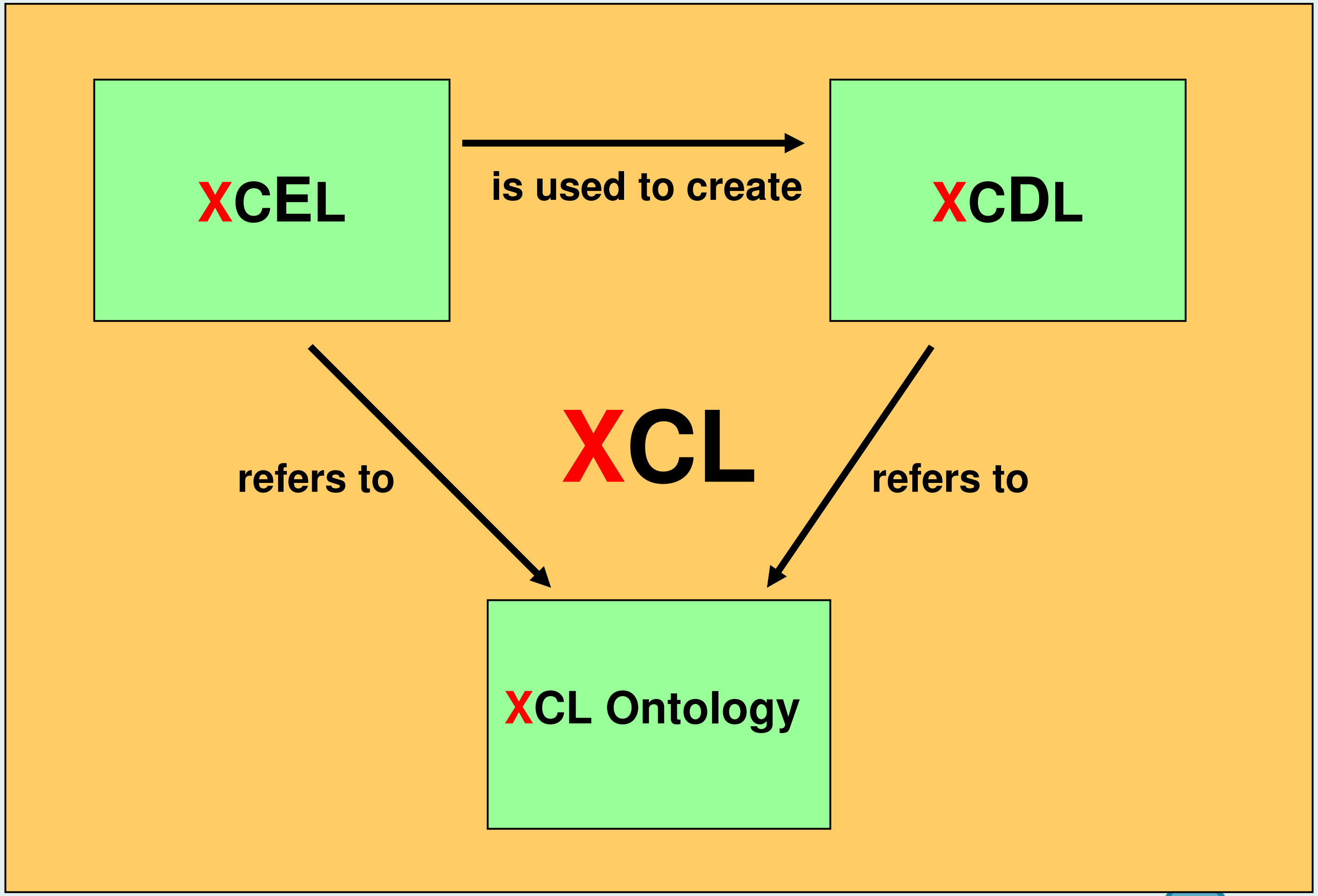
has_alternative_filespecific_name Background_Pdf

is_same_as BackgroundColourRGB

has_alternative_filespecific_name Backgroundcolour_Gif

Data property assertions

XML as backbone language



eXtensible Characterisation Extraction Language (XCEL)

- Describing how properties of digital objects are stored

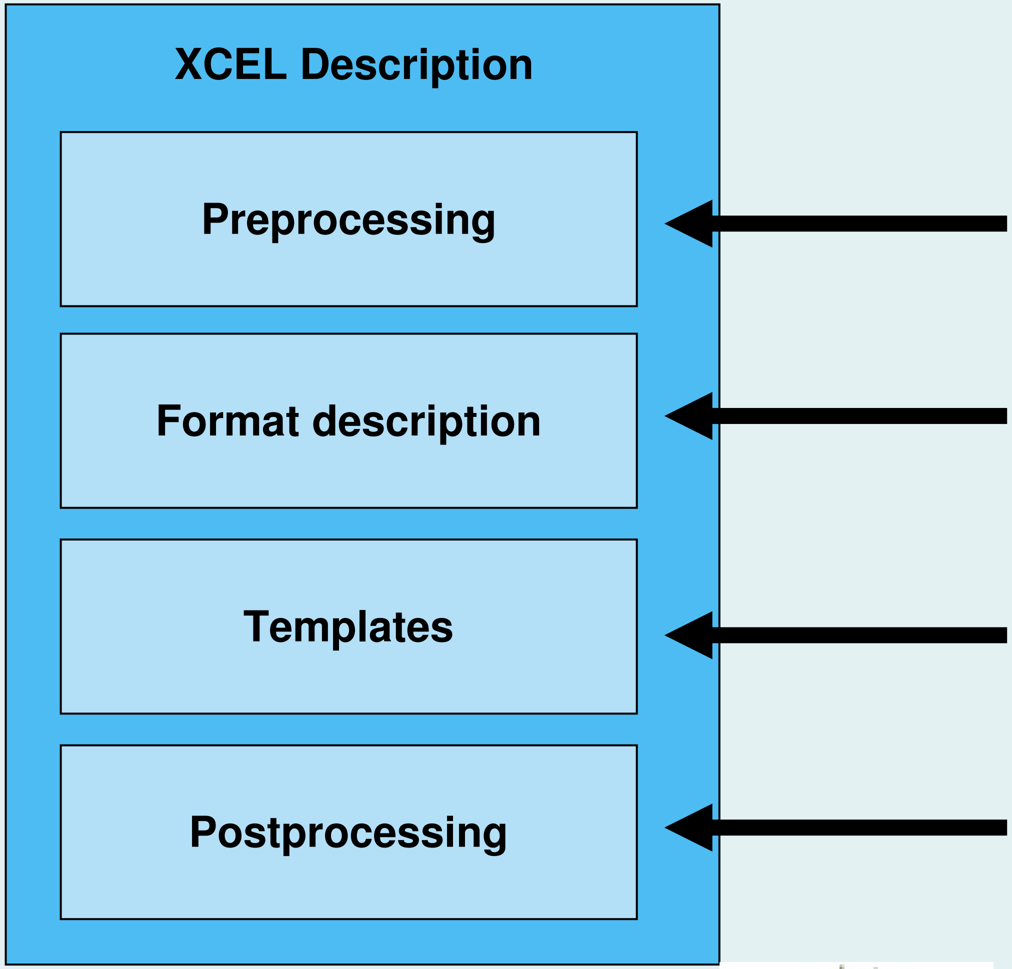
- File format specification tagged in XML, according to the XCEL language definitions

- Interpretable through an XCEL interpreter (Extractor), able to extract characteristics

CEI
CENTRAL EUROPEAN INITIATIVE

planets

XCEL: Global Architecture



XCEL Description

Preprocessing

Format description

Templates

Postprocessing

Configuration tasks, affecting the behaviour of the XCEL interpreter

Description of the structure of the object

Description of recuring structures

Actions on the result of the format description processing

XCEL: Basic Structuring Elements

There are just a few elements sufficient enough to describe a file format:

processing

valueInterpretation

nonValidValues

item

param

valueLabel

symbol

value

eXtensible Characterisation Definition Language (XCDL)

- Describes the content of a file /set of files in an abstract way.
- Designed for decription of the content of *any* file format.
- Designed as a means to describe only parts *or* all of the content.

CENTRAL EUROPEAN INITIATIVE

XCDL: Basic Structuring Elements

Again, there are just a few elements sufficient enough to describe the content of a digital object:

- object
- dataRef
- type
- normData
- property
- valueSet
- value
- propertySet
- labValue

Benefits of the XCL approach

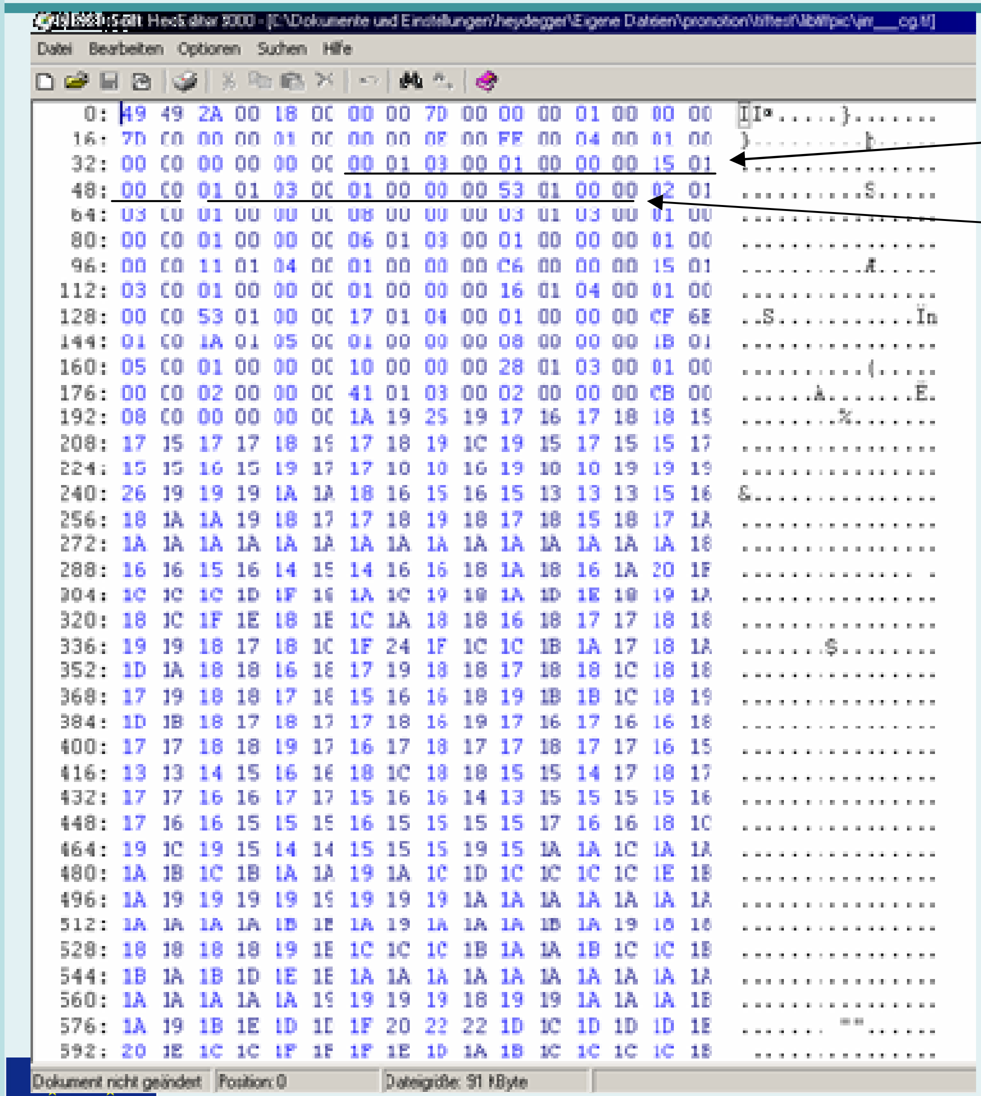
- XCL is a generic solution, uses an abstract model, provides a unique vocabulary

→Extensible: XCL is based on XML

→XCEL provides a means for description of any file format

→ XCDL is a language with which all sort of content can be expressed

XCL by Example


Hex-Editor 2000 - [C:\Dokumente und Einstellungen\heydegger\Eigene Dateien\promotion\tiftest\libtiffpic\jim___cg.tif]
Datei
Bearbeiten
Optionen
Suchen
Hilfe

```
  0: 49 49 2A 00 18 0C 00 00 7D 00 00 00 01 00 00 00   II*.....}.......
 16: 7D C0 00 00 01 0C 00 00 0F 00 FE 00 04 00 01 00   }.........þ.....
 32: 00 C0 00 00 00 0C 00 01 03 00 01 00 00 00 15 01   ................
 48: 00 C0 01 01 03 0C 01 00 00 00 53 01 00 00 02 01   ..........S.....
 64: 03 C0 01 00 00 0C 08 00 00 00 03 01 03 00 01 00   ................
 80: 00 C0 01 00 00 0C 06 01 03 00 01 00 00 00 01 00   ................
 96: 00 C0 11 01 04 0C 01 00 00 00 C6 00 00 00 15 01   ..........Æ.....
112: 03 C0 01 00 00 0C 01 00 00 00 16 01 04 00 01 00   ................
128: 00 C0 53 01 00 0C 17 01 04 00 01 00 00 00 CF 6E   ..S...........Ïn
144: 01 C0 1A 01 05 0C 01 00 00 00 08 00 00 00 1B 01   ................
160: 05 C0 01 00 00 0C 10 00 00 00 28 01 03 00 01 00   ..........(.....
176: 00 C0 02 00 00 0C 41 01 03 00 02 00 00 00 CB 00   ......A.......Ë.
192: 08 C0 00 00 00 0C 1A 19 25 19 17 16 17 18 18 15   ........%.......
208: 17 15 17 17 18 19 17 18 19 1C 19 15 17 15 15 17   ................
224: 15 15 16 15 19 17 17 10 10 16 19 10 10 19 19 19   ................
240: 26 19 19 19 1A 1A 18 16 15 16 15 13 13 13 15 16   &...............
256: 18 1A 1A 19 18 17 17 18 19 18 17 18 15 18 17 1A   ................
272: 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A 18   ................
288: 16 16 15 16 14 15 14 16 16 18 1A 18 16 1A 20 1F   .............. .
304: 1C 1C 1C 1D 1F 16 1A 1C 19 18 1A 1D 1E 18 19 1A   ................
320: 18 1C 1F 1E 18 1E 1C 1A 18 18 16 18 17 17 18 18   ................
336: 19 19 18 17 18 1C 1F 24 1F 1C 1C 1B 1A 17 18 1A   .......$........
352: 1D 1A 18 18 16 16 17 19 18 18 17 18 18 1C 18 18   ................
368: 17 19 18 18 17 16 15 16 16 18 19 1B 1B 1C 18 19   ................
384: 1D 1B 18 17 18 17 17 18 16 19 17 16 17 16 16 18   ................
400: 17 17 18 18 19 17 16 17 18 17 17 18 17 17 16 15   ................
416: 13 13 14 15 16 16 18 1C 18 18 15 15 14 17 18 17   ................
432: 17 17 16 16 17 17 15 16 16 14 13 15 15 15 15 16   ................
448: 17 16 16 15 15 15 16 15 15 15 15 17 16 16 18 1C   ................
464: 19 1C 19 15 14 14 15 15 15 19 15 1A 1A 1C 1A 1A   ................
480: 1A 1B 1C 1B 1A 1A 19 1A 1C 1D 1C 1C 1C 1C 1E 1B   ................
496: 1A 19 19 19 19 19 19 19 19 1A 1A 1A 1A 1A 1A 1A   ................
512: 1A 1A 1A 1A 1B 1E 1A 19 1A 1A 1A 1B 1A 19 18 18   ................
528: 18 18 18 18 19 1E 1C 1C 1C 1B 1A 1A 1B 1C 1C 1B   ................
544: 1B 1A 1B 1D 1E 1E 1A 1A 1A 1A 1A 1A 1A 1A 1A 1A   ................
560: 1A 1A 1A 1A 1A 19 19 19 19 18 19 19 1A 1A 1A 1B   ................
576: 1A 19 1B 1E 1D 1D 1F 20 22 22 1D 1C 1D 1D 1D 1E   ....... ""......
592: 20 1E 1C 1C 1F 1F 1F 1E 1D 1A 1B 1C 1C 1C 1C 1B    ...............
```

Dokument nicht geändert | Position: 0 | Dateigröße: 91 KByte

Image width: 277

Image length: 339

ImageLength

The number of rows of pixels in the image.

Tag = 257 (101.H)

Type = SHORT or LONG

N = 1

No default. See also ImageWidth.

ImageWidth

The number of columns in the image, i.e., the number of pixels per row.

Tag = 256 (100.H)

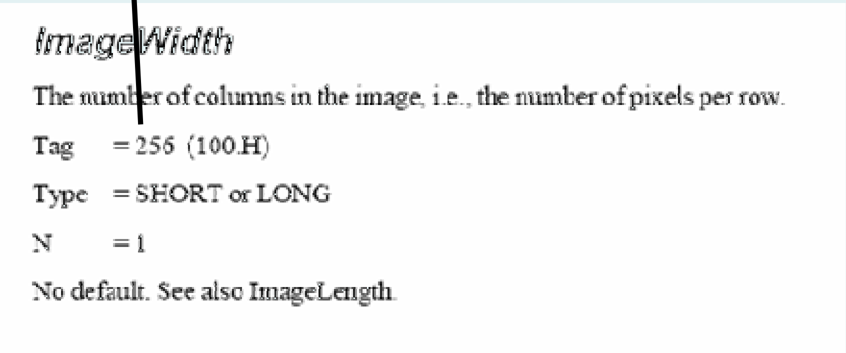
Type = SHORT or LONG

N = 1

No default. See also ImageLength.

XCEL representation

```
<!-- Tag 256: ImageWidth (XCL: imageWidth) -->
 <item xsi:type="structuringItem" identifier="IFDE_256"
  optional="true">
  <symbol interpretation="uint16" length="2" value="256"/>
  <item xsi:type="structuringItem" order="choice">
    <item xsi:type="structuringItem" order="sequence">
      <!– Data type (value ‚3‘ means uint16)-->
      <symbol interpretation="uint16" length="2"  value="3"/>
      <!– number of values (N)->
      <symbol interpretation="uint32" length="4" value="1"/>
      <!-- the value and name of property -->
      <symbol interpretation="uint16" length="2"
      name="imageWidth"/>
      <!-- wasted space-->
      <symbol interpretation="uint16" length="2"/>
      […]
    </item>
  </item>
 </item>
```



ImageWidth

The number of columns in the image, i.e., the number of pixels per row.

Tag = 256 (100.H)

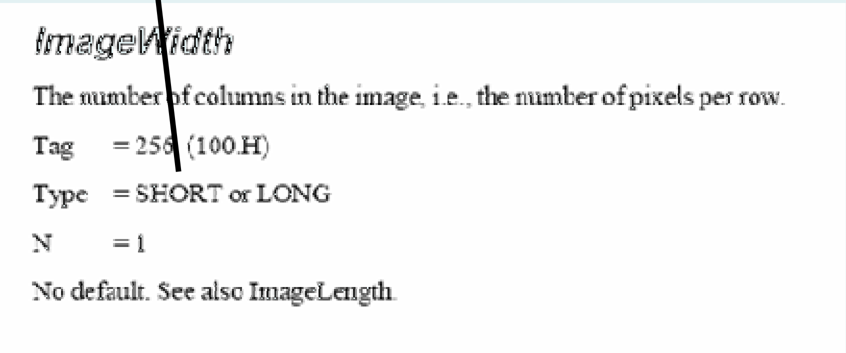
Type = SHORT or LONG

N = 1

No default. See also ImageLength.

XCEL representation

```
<!-- Tag 256: ImageWidth (XCL: imageWidth) -->
  <item xsi:type="structuringItem" identifier="IFDE_256"
   optional="true">
   <symbol interpretation="uint16" length="2" value="256"/>
   <item xsi:type="structuringItem" order="choice">
     <item xsi:type="structuringItem" order="sequence">
       <!– Data type (value ‚3‘ means uint16)-->
       <symbol interpretation="uint16" length="2"  value="3"/>
       <!– number of values (N)->
       <symbol interpretation="uint32" length="4" value="1"/>
       <!-- the value and name of property -->
       <symbol interpretation="uint16" length="2"
       name="imageWidth"/>
       <!-- wasted space-->
       <symbol interpretation="uint16" length="2"/>
       […]
     </item>
   </item>
  </item>
```



ImageWidth

The number of columns in the image, i.e., the number of pixels per row.

Tag = 256 (100.H)

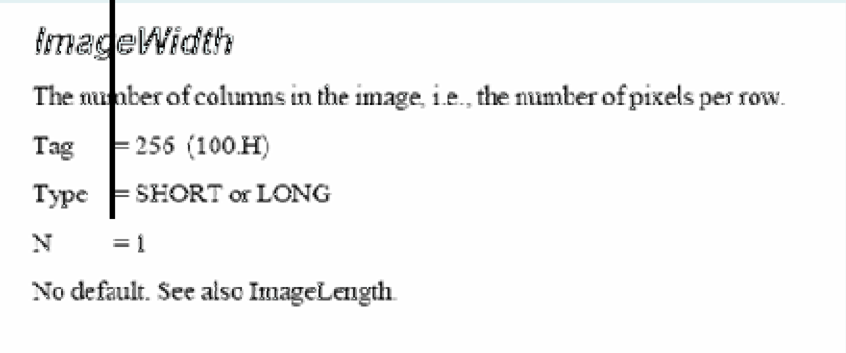
Type = SHORT or LONG

N = 1

No default. See also ImageLength.

XCEL representation

```
<!-- Tag 256: ImageWidth (XCL: imageWidth) -->
  <item xsi:type="structuringItem" identifier="IFDE_256"
   optional="true">
   <symbol interpretation="uint16" length="2" value="256"/>
   <item xsi:type="structuringItem" order="choice">
     <item xsi:type="structuringItem" order="sequence">
       <!– Data type (value ‚3‘ means uint16)-->
       <symbol interpretation="uint16" length="2"  value="3"/>
       <!– number of values (N)->
       <symbol interpretation="uint32" length="4" value="1"/>
       <!-- the value and name of property -->
       <symbol interpretation="uint16" length="2"
       name="imageWidth"/>
       <!-- wasted space-->
       <symbol interpretation="uint16" length="2"/>
       […]
     </item>
   </item>
  </item>
```



ImageWidth

The number of columns in the image, i.e., the number of pixels per row.

Tag = 256 (100.H)

Type = SHORT or LONG

N = 1

No default. See also ImageLength.

XCDL representation

```
…
<property id="p5">
    <name id="id30" >imageWidth</name>
    <valueSet id="i_i1_s4" >
        <labValue>
            <val>277</val>
            <type>int</type>
        </labValue>
    </valueSet>
 </property>
...
```

XCEL entry:

```
<!-- the value and name of property -->
    <symbol interpretation="uint16" length="2"
    name="imageWidth"/>
```

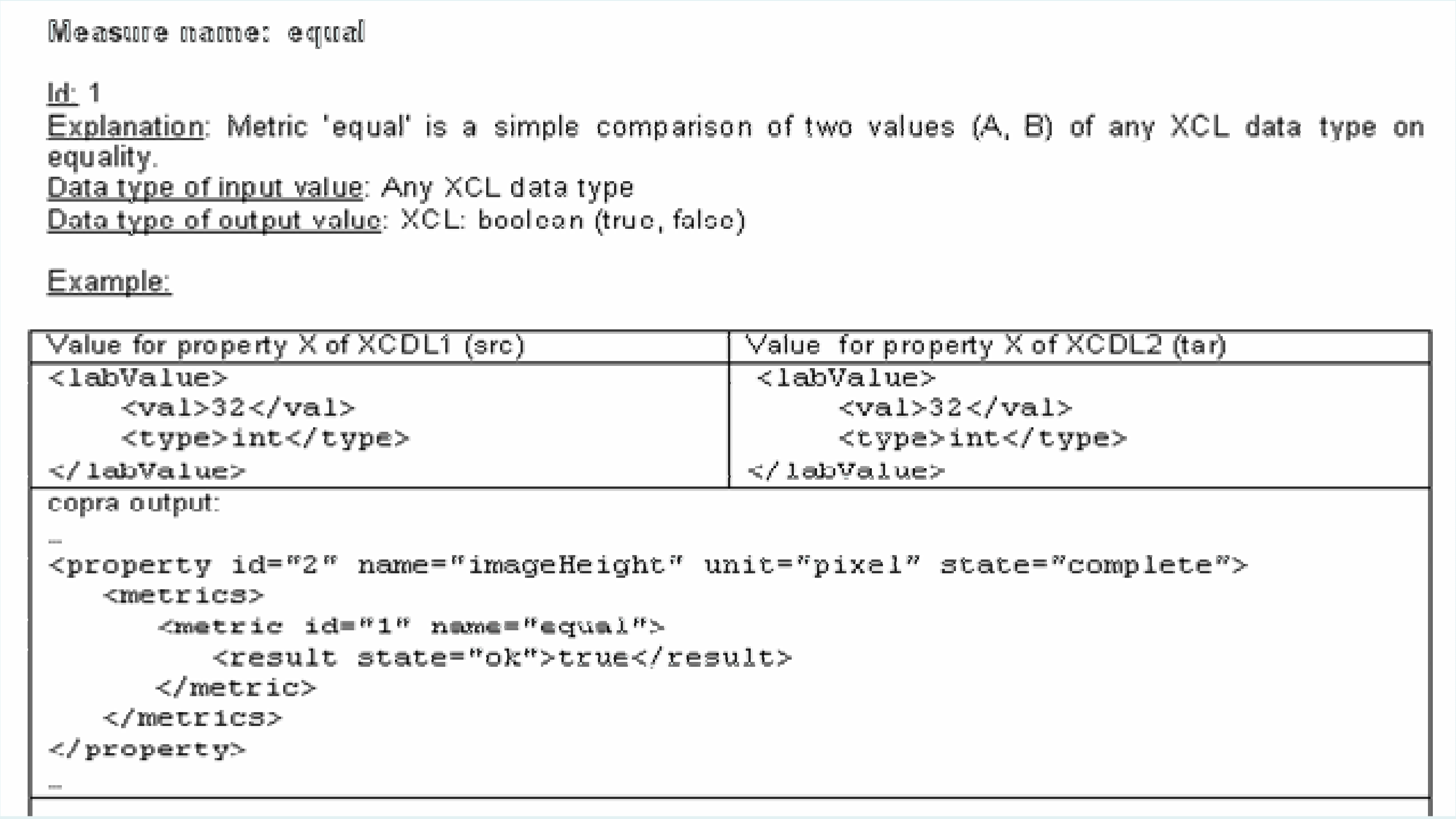
XCDL representation

```
…
<property id="p5">
     <name id="id30" >imageWidth</name>
     <valueSet id="i_i1_s4" >
         <labValue>
             <val>277</val>
             <type>int</type>
         </labValue>
     </valueSet>
  </property>
...
```

XCEL entry:

```
<!– Data type (value ‚3‘ means uint16)-->
      <symbol interpretation="uint16"
        length="2"  value="3"/>
```

XCDL representations can now be compared…



Measure name: equal

Id: 1
Explanation: Metric 'equal' is a simple comparison of two values (A, B) of any XCL data type on equality.
Data type of input value: Any XCL data type
Data type of output value: XCL: boolean (true, false)

Example:

Value for property X of XCDL1 (src)	Value for property X of XCDL2 (tar)

`<labValue>`
`    <val>32</val>`
`    <type>int</type>`
`</labValue>`
	`<labValue>`
	`    <val>32</val>`
	`    <type>int</type>`
	`</labValue>`

copra output:

```
…
<property id="2" name="imageHeight" unit="pixel" state="complete">
   <metrics>
      <metric id="1" name="equal">
         <result state="ok">true</result>
      </metric>
   </metrics>
</property>
…
```

Thank you for your attention!

Any questions?