



IBM Global Business Services  
Long-Term Preservation  
Center of Excellence – The Netherlands

## DigCCurr 2009 Digital Curation Practice, Promise and Prospects

### IBM Demo Session

April 1st 2009  
Dr. Raymond J. van Diessen  
Tel: +31-6-51617487  
Email: [Raymond\\_vanDiessen@nl.ibm.com](mailto:Raymond_vanDiessen@nl.ibm.com)



# Preservation Manager Demo

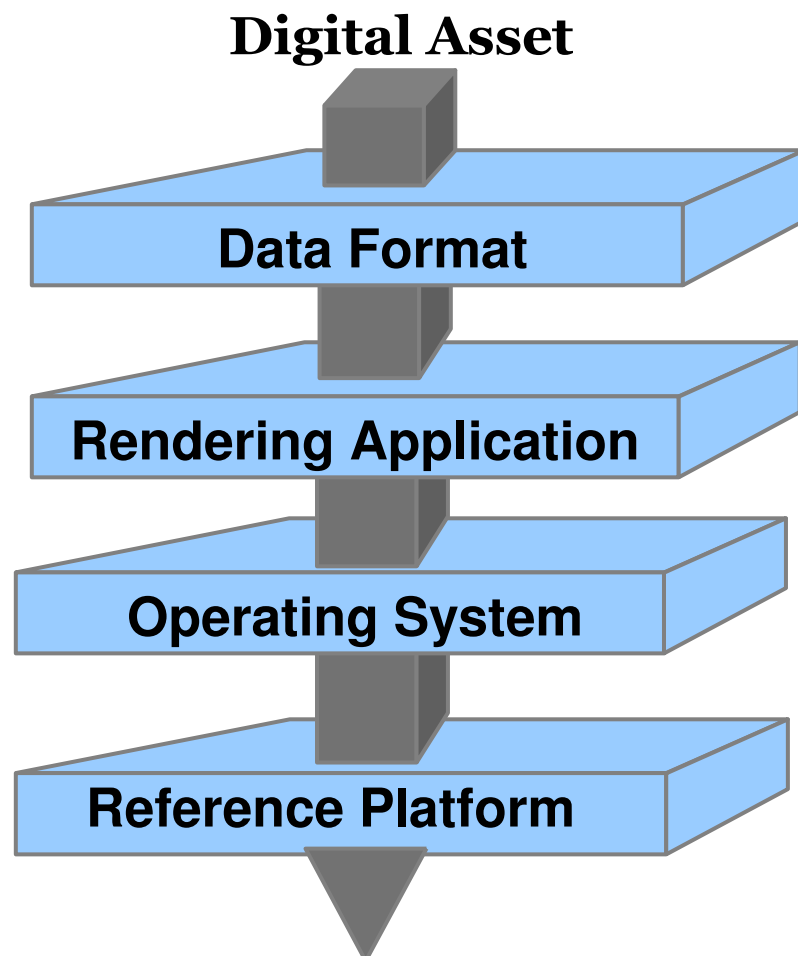
## Objectives

*The Preservation Manager provides the services for monitoring the technical environments needed to ensure accessibility of the digital objects over time:*

- Identification of digital objects endangered of becoming inaccessible due to technology changes (hardware and software)
- Guiding the preservation actions to be taken, i.e. implementation of specific migration and emulation strategies
- Supplying technical metadata to generate / validate the technical environments during delivery / rendering of a digital object

## Preservation Manager Core Concepts

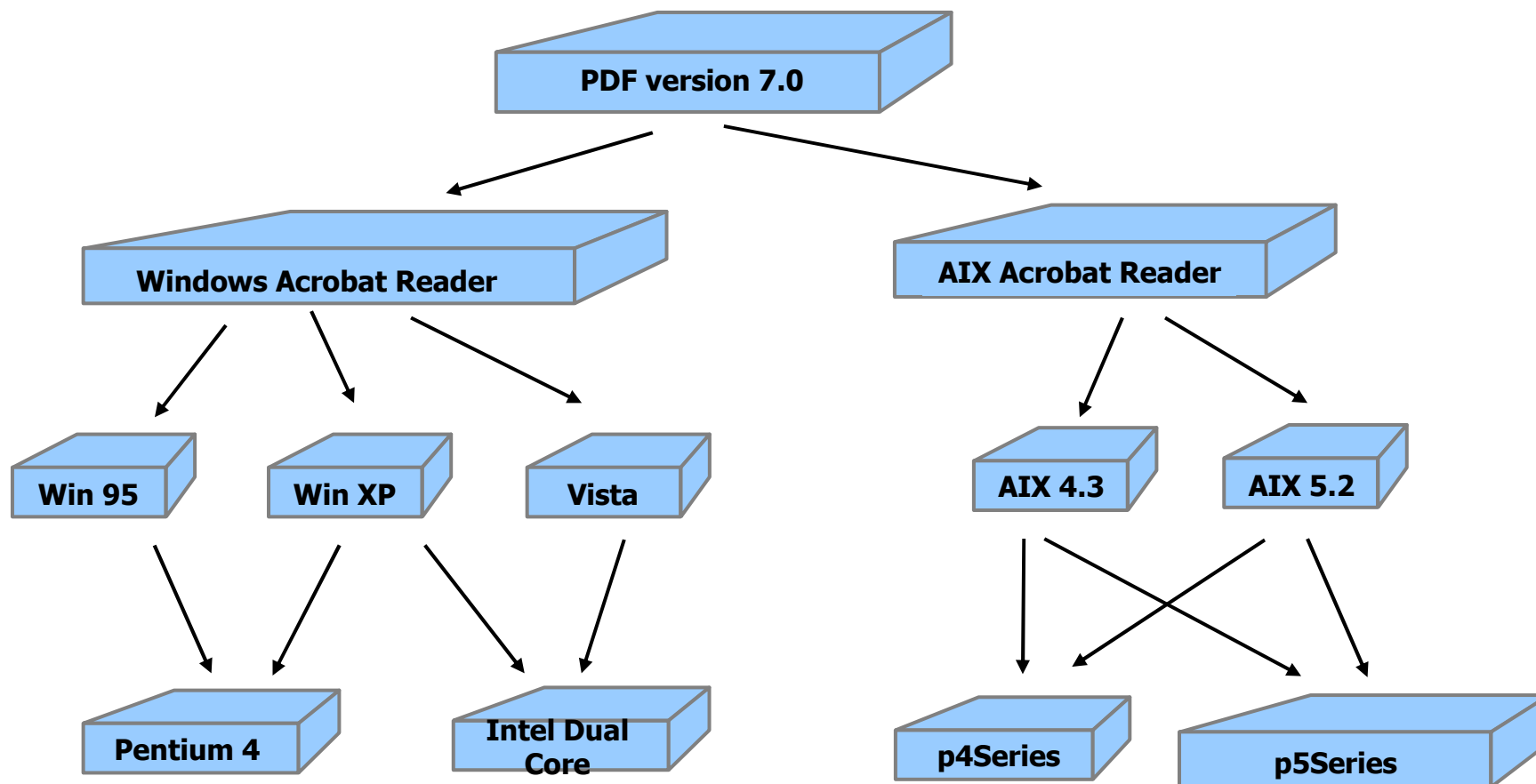
*On an abstract level the complete hardware and software stack needed to render a digital asset contains four layers*



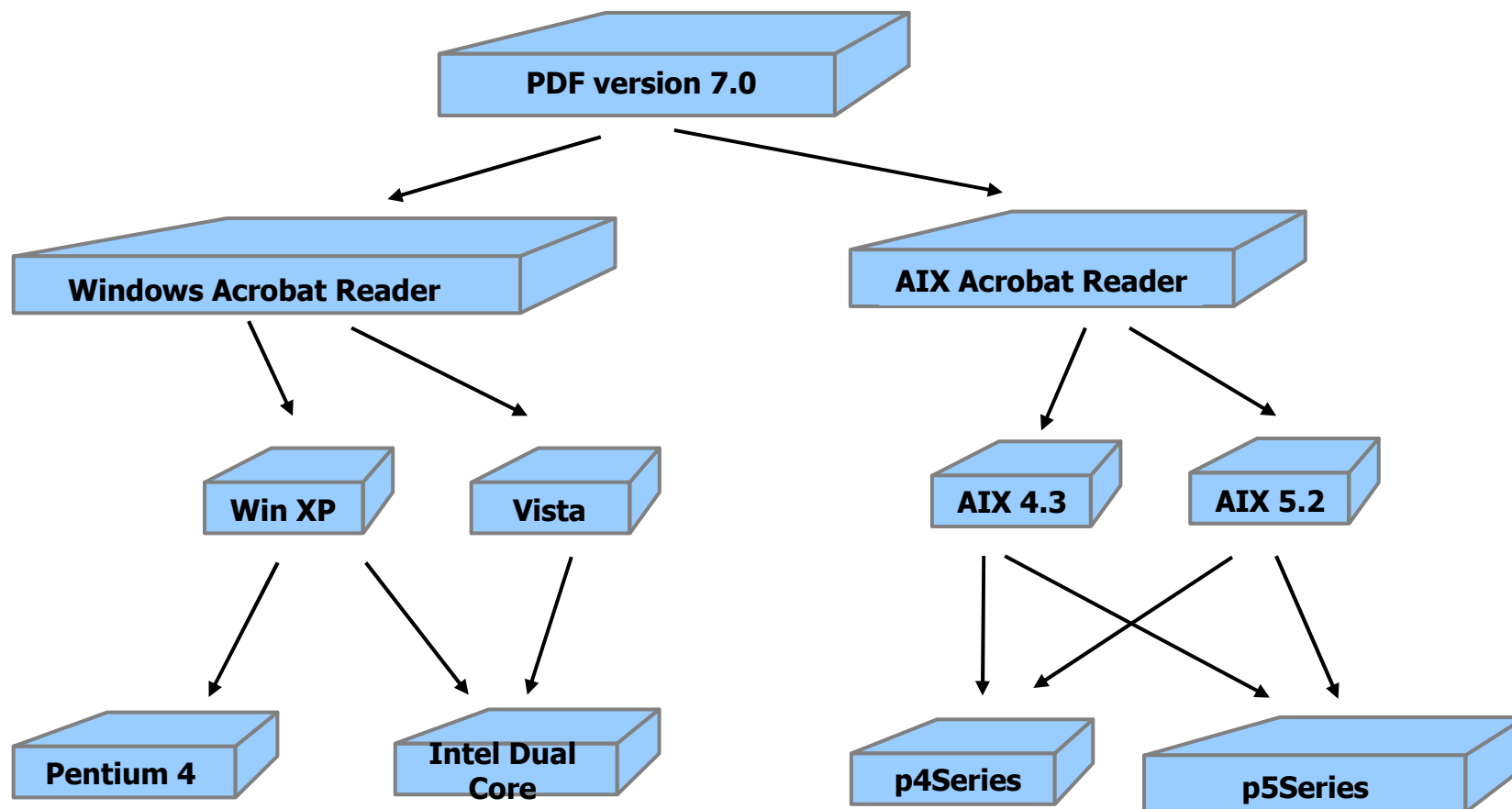
- The [Data Format](#) identifies the structure and some of the meaning of a bit-stream.
- The structure and meaning of the bit-stream are further defined within the application logic of a specific [Rendering Application](#)
- The [Operating System](#) contains shared functionalities that all rendering applications need like access to a printer, screen, file interaction, font rendering etc.
- The [Reference Platform](#) specifies how the bits and bytes are transformed to a physical representation like pages or images

**Requested Information Object**

## View Paths

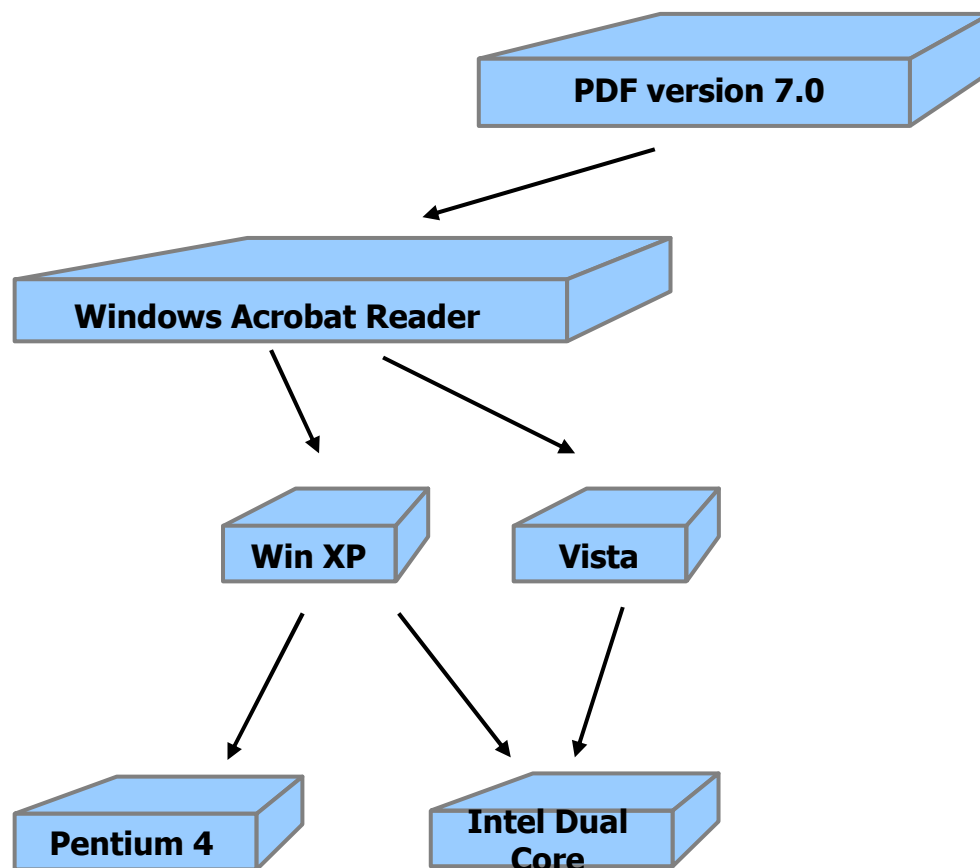


## View Paths



So... **Windows<sup>95</sup>** gets obsolete..

## View Paths



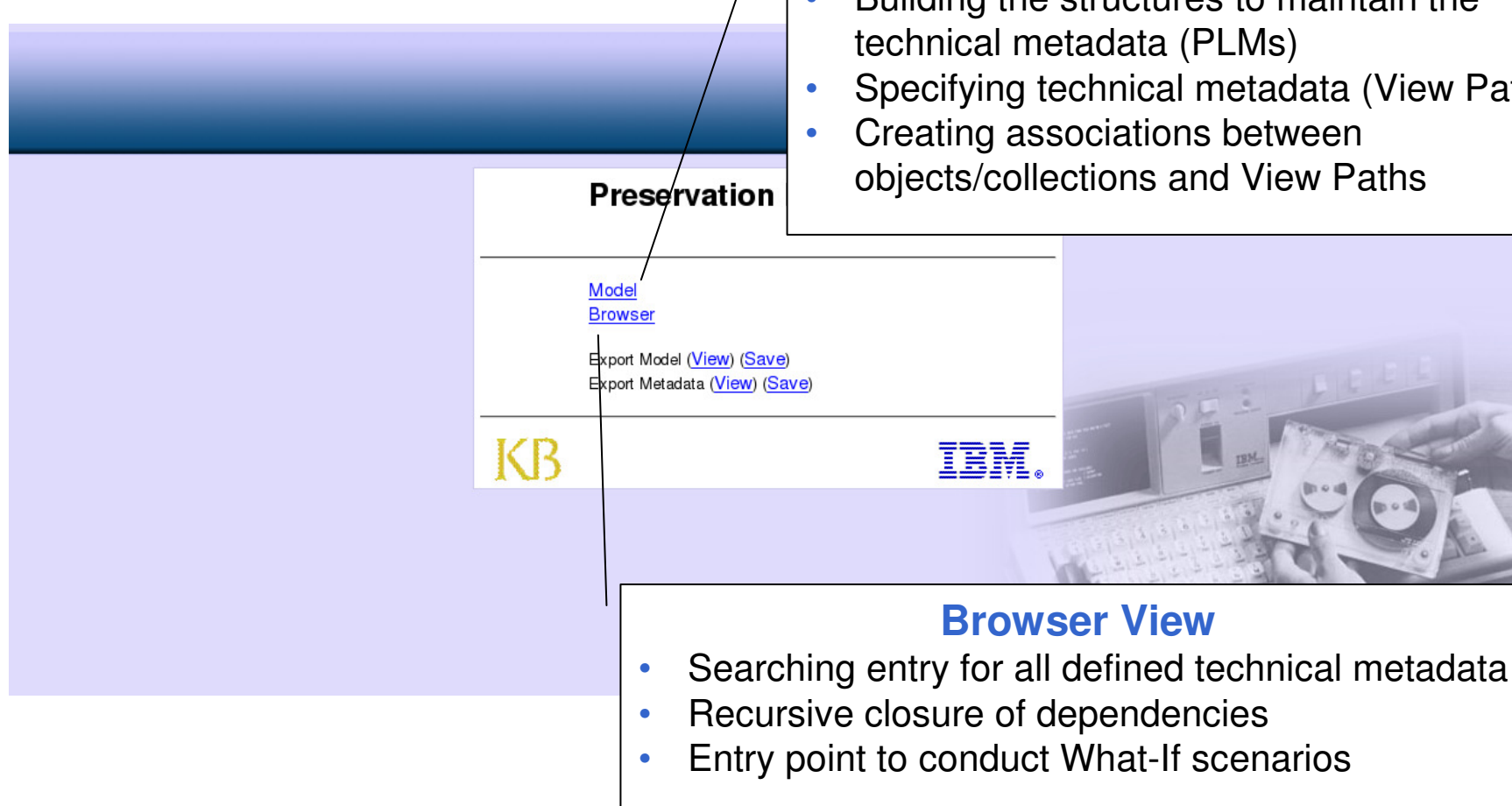
So... **Windows<sup>95</sup>** gets obsolete..

And... **Acrobat for AIX** is no longer supported..



# Preservation Manager

*The Preservation Manager knows two modes of operation: model view and browser view*





## Model View

*The technical metadata is defined “Lego Style” through the reuse of more basic building block*

### ✓ Attribute

Preservation Manager Model: Attributes

Status = Under Construction

Attributes

- Applications
- CPU
- CPU speed
- Color depth
- Device drivers
- Display resolution
- End of support
- General usage end
- General usage start
- Harddisk
- Internet connection
- Keyboard
- Libraries
- Memory
- Mouse
- Name
- Pixels per Inch
- Release date
- Remarks
- Removable media
- Sound card
- Supplier
- Version
- Video card
- Visual display unit

Name: Display resolution

Description: The display resolution of a digital television or computer display typically refers to the number of distinct pixels in each dimension that can be displayed. It can be an ambiguous term especially as the displayed resolution is controlled by all different factors in cathode ray tube (CRT) and flat panel or projection displays using fixed picture-element (pixel) arrays.

One use of the term "display resolution" applies to fixed-pixel-array displays such as plasma display panels (PDPs), liquid crystal displays (LCDs), digital light processing (DLP) projectors, or similar technologies, and is simply the physical number of columns and rows of pixels creating the display (e.g., 1280×1024). A consequence of having a fixed grid display is that for multiformat video inputs all displays need a "scaling-engine" (a digital video processor that includes a memory array) to match the incoming picture format to the display.

Type: Enumeration

Metric:

Value:

Add > Import

320\*200 (CGA)  
640\*480 (VGA)  
800\*600 (SVGA)  
1024\*768 (XGA)  
1280\*1024 (SXGA)  
1400\*1050 (SXGA+)  
1600\*1200 (UGA)  
1920\*1080 (HD 1080)  
1920\*1200 (WUXGA)

Delete

Save New

## Model View

*The technical metadata is defined “Lego Style” through the reuse of more basic building block*

✓Layer

Preservation Manager Model: Layers

Status = Under Construction

Layers

- Application
- Desktop Personal Computer
- Operating System
- Prerequisites

Name: Application

Description:

Application software is any tool that functions and is operated by means of a computer, with the purpose of supporting or improving the software user's work. In other words, is the subclass of computer software that employs the capabilities of a computer directly and thoroughly to a task that the user wishes to perform. This should be contrasted with system software (infrastructure) or middleware (computer services/ processes integrators) is involved in integrating a computer's various capabilities, but typically does not directly apply them in the performance of tasks that benefit the user. In this context the term application refers to both the application software and its implementation.[1]

A simple, if imperfect analogy in the world of hardware would be the relationship of an electric light bulb (an application) to an electric power generation plant (a system). The power plant merely generates electricity, not itself of any real use until harnessed to an application like the electric light that performs a service that benefits the user.

Typical examples of software applications are word processors, spreadsheets, and media players,

Attributes:

- Applications
- CPU
- CPU speed
- Color depth
- Device drivers
- Display resolution
- End of support
- General usage end
- General usage start
- Harddisk

Add >

< Remove

☒ Required

☐ Multi Valued

Name \*

Version \*

Supplier \*

Release date \*

Up

Down

\* - Required

+ - Multi Valued

Save

New

## Model View

*The technical metadata is defined “Lego Style” through the reuse of more basic building block*

✓PLM

Preservation Manager Model: PLM's

PLM's

Application on Personal Computer

Name: Application on Personal Computer

System Id: PLM020090111090350903KB0017835

Description: The PLM defines the stack to describe an application on a Personal Computer; Layer Prerequisites was removed from this PLM.

Layers

Application  
Desktop Personal Computer  
Operating System  
Prerequisites

Add >  
< Remove

Application  
Prerequisites  
Operating System  
Desktop Personal Computer

Up  
Down

Save Copy New Delete

## Model View

*The technical metadata is defined “Lego Style” through the reuse of more basic building block*

File types

Portable Document Format (PDF) (1.2)  
Portable Document Format (PDF) (1.3)  
Portable Document Format (PDF) (1.7)

Preservation Manager Model: File types

Status = Under Construction

Name: Portable Document Form

File type Id (DIAS): 001

Format type: PDF

Version: 1.2

Base coding: application/pdf

Description: Portable Document Format (PDF) is a file format created by Adobe Systems in 1993 for document exchange. PDF is used for representing two-dimensional documents in a manner independent of the application software, hardware, and operating system.[1]  
Each PDF file encapsulates a complete description of a fixed-layout 2-D document (and, with Acrobat 3-D, embedded 3-D documents) that includes the text, fonts, images, and 2-D vector graphics which

Save New

✓File Type

## Model View

*The technical metadata is defined “Lego Style” through the reuse of more basic building block*

Preservation Manager Model: View paths

Status = Under Construction

View paths

- PDF 1.2 Viewer
- PDF 1.3 Viewer
- PDF 1.7 Viewer
- PDF 1.7 Viewer (Linux)

Name: PDF 1.2 Viewer

Description: Portable Document Format (PDF) is a file format created by Adobe Systems in 1993 for document exchange. PDF is used for representing two-dimensional documents in a manner independent of the application software, hardware, and operating system.[1]

Each PDF file encapsulates a complete description of a fixed-layout 2-D document (and, with Acrobat 3-D, embedded 3-D documents) that includes the text, fonts, images, and 2-D vector graphics which comprise the documents.

PDF is an open standard that was officially published on July 1, 2008 by the ISO as ISO 32000-1:2008.

PDF version 1.2 was release in 1996.  
Wikipedia, Jan 2009

Review Date: 2010-01-21 (YYYY-MM-DD)

PLM: Application on Personal Computer

Application

Name \*: Adobe Acrobat Reader for Windows

Version \*: 3.01

Supplier \*: Adobe Systems Incorporated

Release date \*: 1997-06-10 dd/mm/yyyy

Prerequisites

Applications +

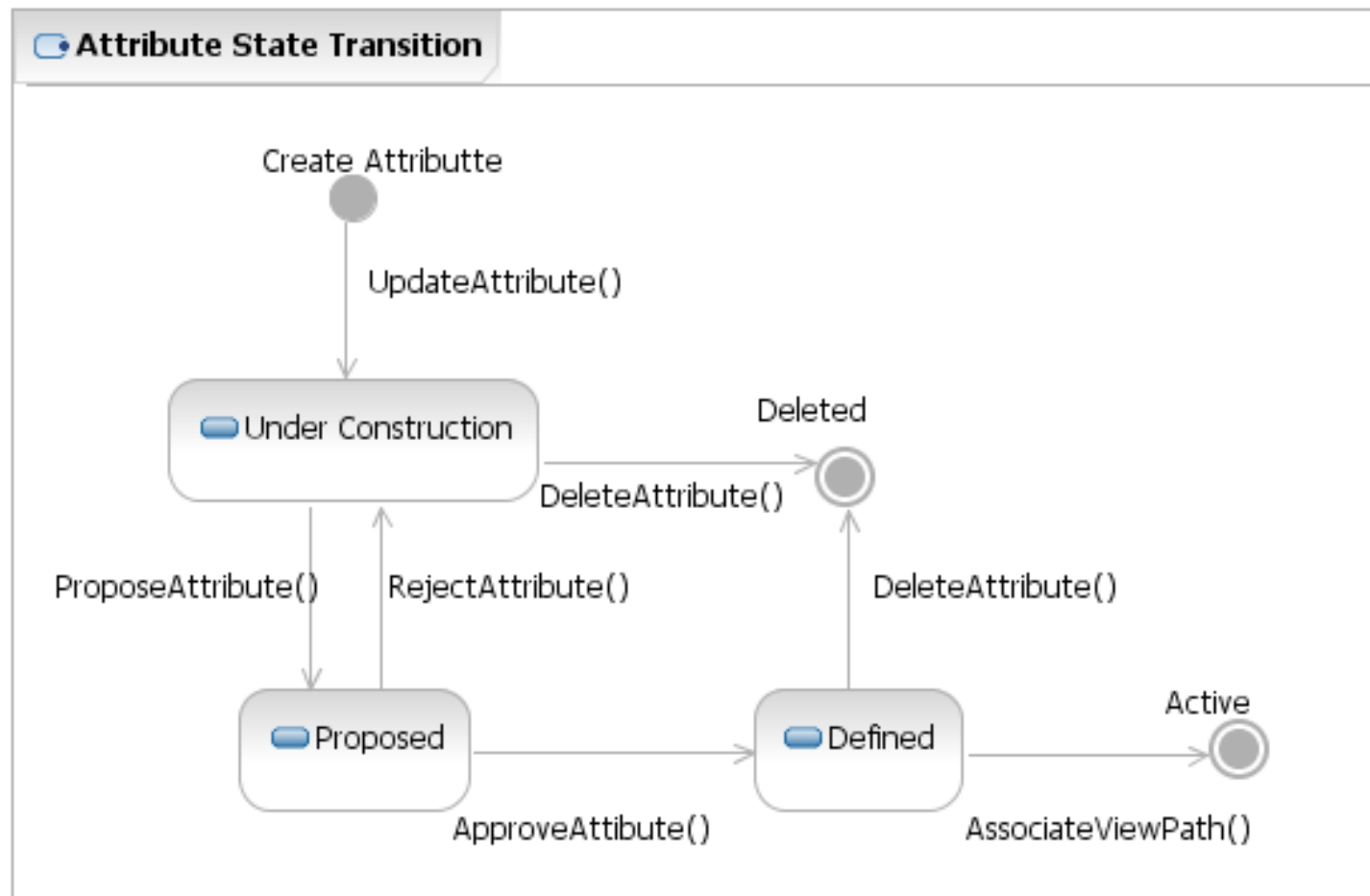
Add

✓View Path



## Quality Control

*Quality control within the Preservation Manager is being managed by the status indicator of each building block*



(above is also applicable for other building block: layer, PLM, view path and file type)

## Browser View

*The Browser view provides a structured navigation process through the technical metadata and their dependencies*

**Preservation Manager Browser**


Search  object(s) with status  and creation date  dd/mm/yyyy

Search String

File types	View paths	PLM's	Layers	Attributes
Portable Document Format (PDF) (1.7)	PDF 1.7 Viewer PDF 1.7 Viewer (Linux)	Application on Personal Computer	Application Prerequisites Operating System Desktop Personal Computer	Name Version Supplier Release date Applications Device drivers Libraries End of support CPU CPU speed Memory Harddisk Video card Visual display unit Display resolution Color depth Sound card Keyboard Mouse Removable media

Object	File type
Id	FTY020090111111541931KB0014023
Name	Portable Document Format (PDF)
Description	Portable Document Format (PDF) is a file format created by Adobe Systems in 1993 for document exchange. PDF is used for representing two-dimensional documents in a manner independent of the application software, hardware, and operating system.[1] Each PDF file encapsulates a complete description of a fixed-layout 2-D document (and, with Acrobat 3-D, embedded 3-D documents) that includes the text, fonts, images, and 2-D vector graphics which comprise the documents. PDF is an open standard that was officially published on July 1, 2008 by the ISO as ISO 32000-1:2008. PDF 1.7 was release in 2006. Wikipedia, Jan 2009
Status	Active
Creation date	Sunday, January 11, 2009
DIAS file type Id	003
Format type	PDF
Version	1.7
Base Coding	application/pdf





## Usage Scenarios

*Preservation Manager be used in different type of configurations:*

### Usage Scenarios:

- Additional DIAS component
  - File type management
  - Technical metadata management
- File / Software Registry
  - File type management
  - Software configuration management

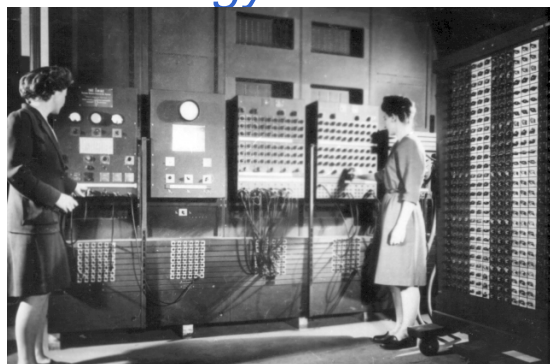
### IBM's Preservation Manager future value propositions:

- File / Software Registry implementation
  - IBM controlled formats and software stacks
  - Signaling and advisory services
- Combining with IBM's Preservation Data Store Technology
  - Specifying migration actions
  - ReplInfo Manager

# Universal Virtual Computer Demo

## Technology Innovation

*The speed of technology change is only increasing and with it technology obsolescence*



ENIAC (1946)



PDP-11 (1970)



IBM AT (1981)



Macintosh (1984)



IBM BLUE GENE/L (2005)



OQO (2002)

# Universal Virtual Computer

*The UVC has a very elementary instruction set (25) and it is a bit addressable*

op (8 bits)	segment (32 bits)	flag (1 bit)	register (31 bits)	segment (32 bits)	flag (1 bit)	register (31 bits)
----------------	----------------------	-----------------	-----------------------	----------------------	-----------------	-----------------------

## Move information between registers and memory

<b>Load</b>	Reg1 (dest), Reg2 (address), Reg3 (length)
<b>Store</b>	Reg1 (src), Reg2 (address), Reg3 (length)
<b>Lsign</b>	Reg1 (dest), Reg2 (address)
<b>Ssign</b>	Reg1 (src), Reg2 (address)

## Operations on registers

<b>Loadr</b>	Reg1 (dest), Reg2 (src)
<b>Psign</b>	Reg1 (positive)
<b>Nsign</b>	Reg1 (negative)
<b>Loadc</b>	Reg1 (dest), Reg2 (length), bit string
<b>Rlen</b>	Reg1 (length), Reg2 (src)

## Numeric instructions

<b>Add</b>	Reg1 (dest), Reg2 (arg)
<b>Subt</b>	Reg1 (dest), Reg2 (arg)
<b>Mult</b>	Reg1 (dest), Reg2 (arg)
<b>Div</b>	Reg1 (quotient), Reg2 (arg), Reg3 (remainder)

## Comparison instructions

<b>Grt</b>	Reg1 (arg1), Reg2 (arg2)
<b>Equ</b>	Reg1 (arg1), Reg2 (arg2)

## Logical instructions

<b>Not</b>	Reg1 (dest),
<b>Or</b>	Reg1 (dest), Reg2 (arg1)
<b>And</b>	Reg1 (dest), Reg2 (arg1)

## Instructions that alter the flow of execution

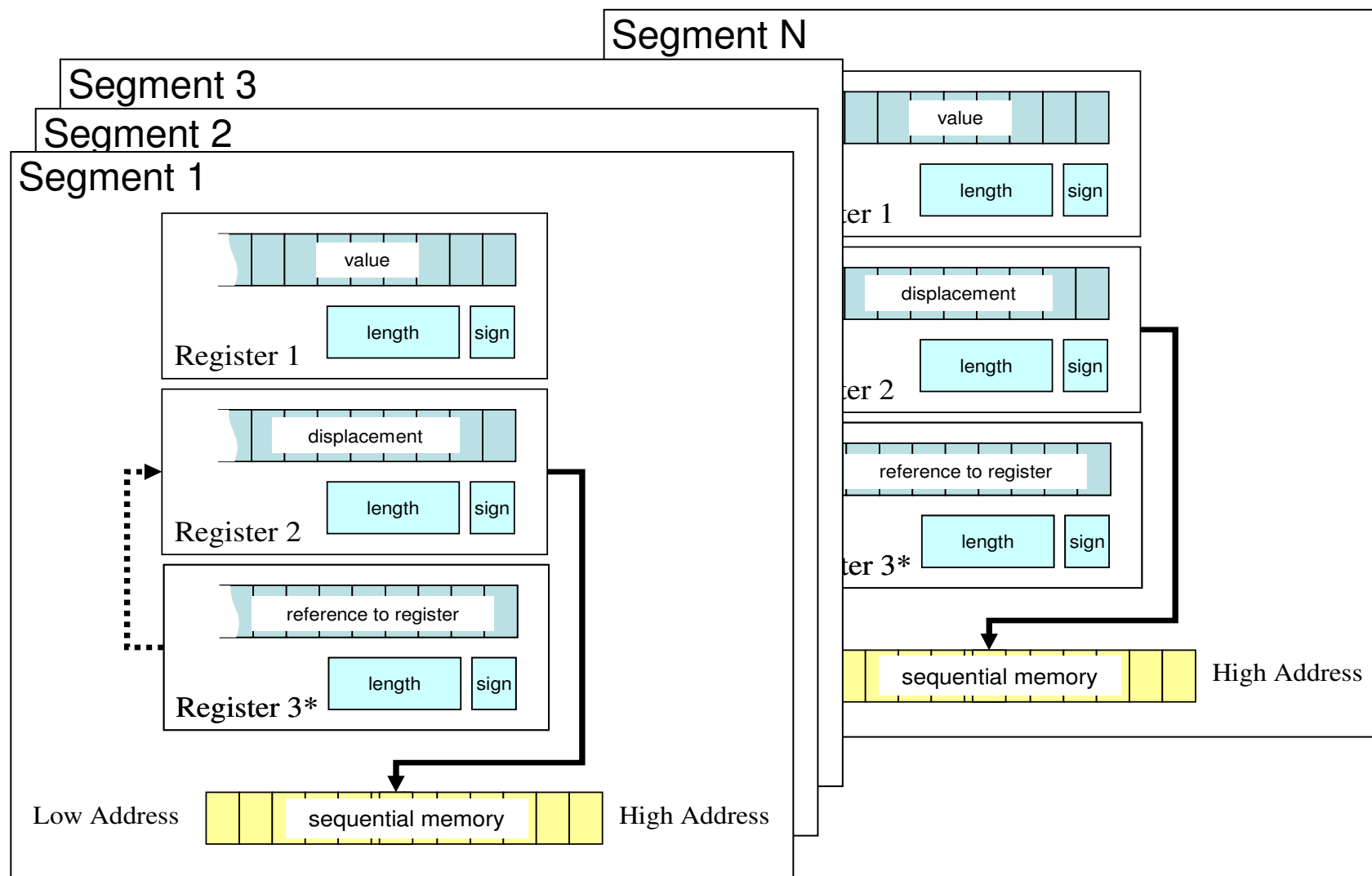
<b>Br</b>	Reg1 (address)
<b>Brc</b>	Reg1 (address)
<b>Break</b>	
<b>Call</b>	Reg1 (section), Reg2 (address), Reg3 (argument)
<b>Stop</b>	

## Communication with the outside world (I/O)

<b>In</b>	Reg1 (msg type), Reg2 (length), Reg3 (address)
<b>Out</b>	Reg1 (msg type), Reg2 (length), Reg3 (address)

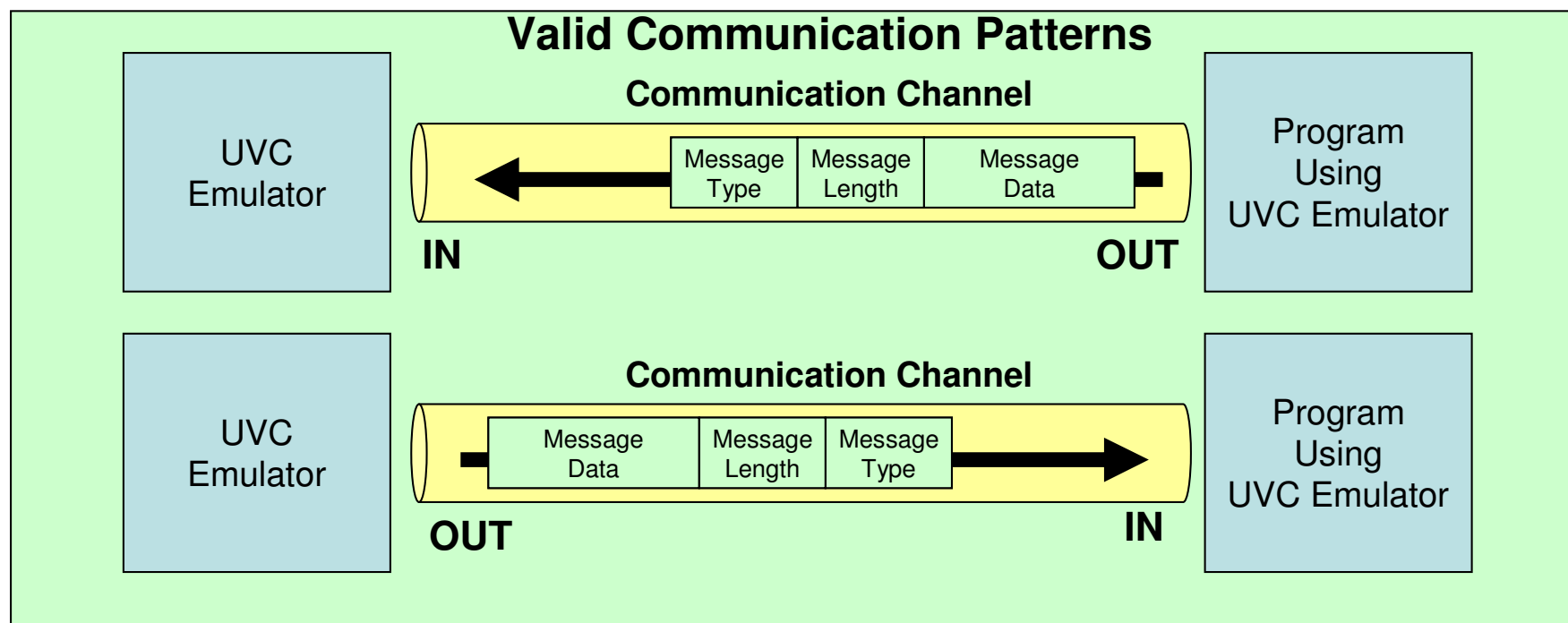
# Virtualization

*UVC is a segment-based machine with almost unlimited resources*



## Communication

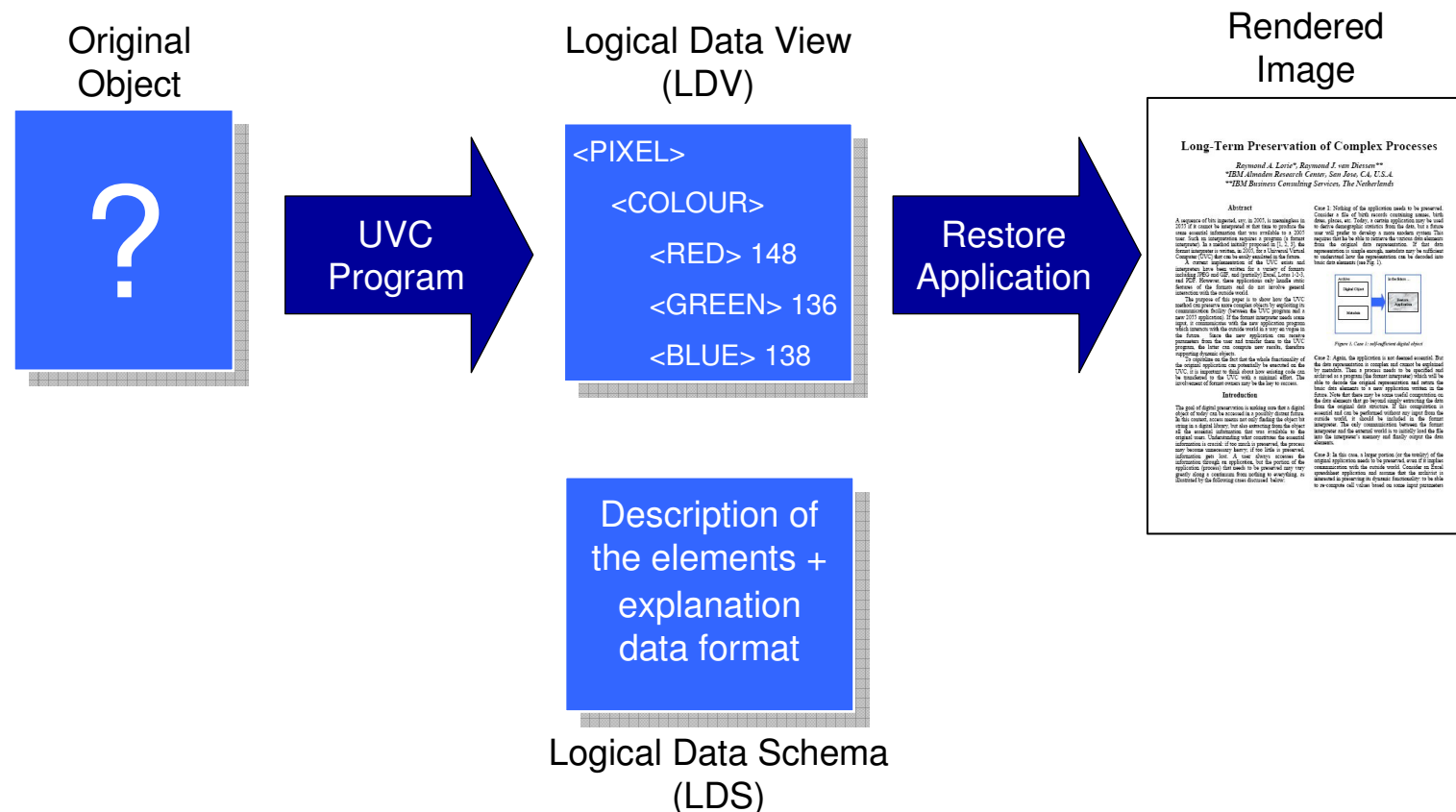
*Interaction with the outside world is established through the defined communication channel*





# UVC Data Preservation

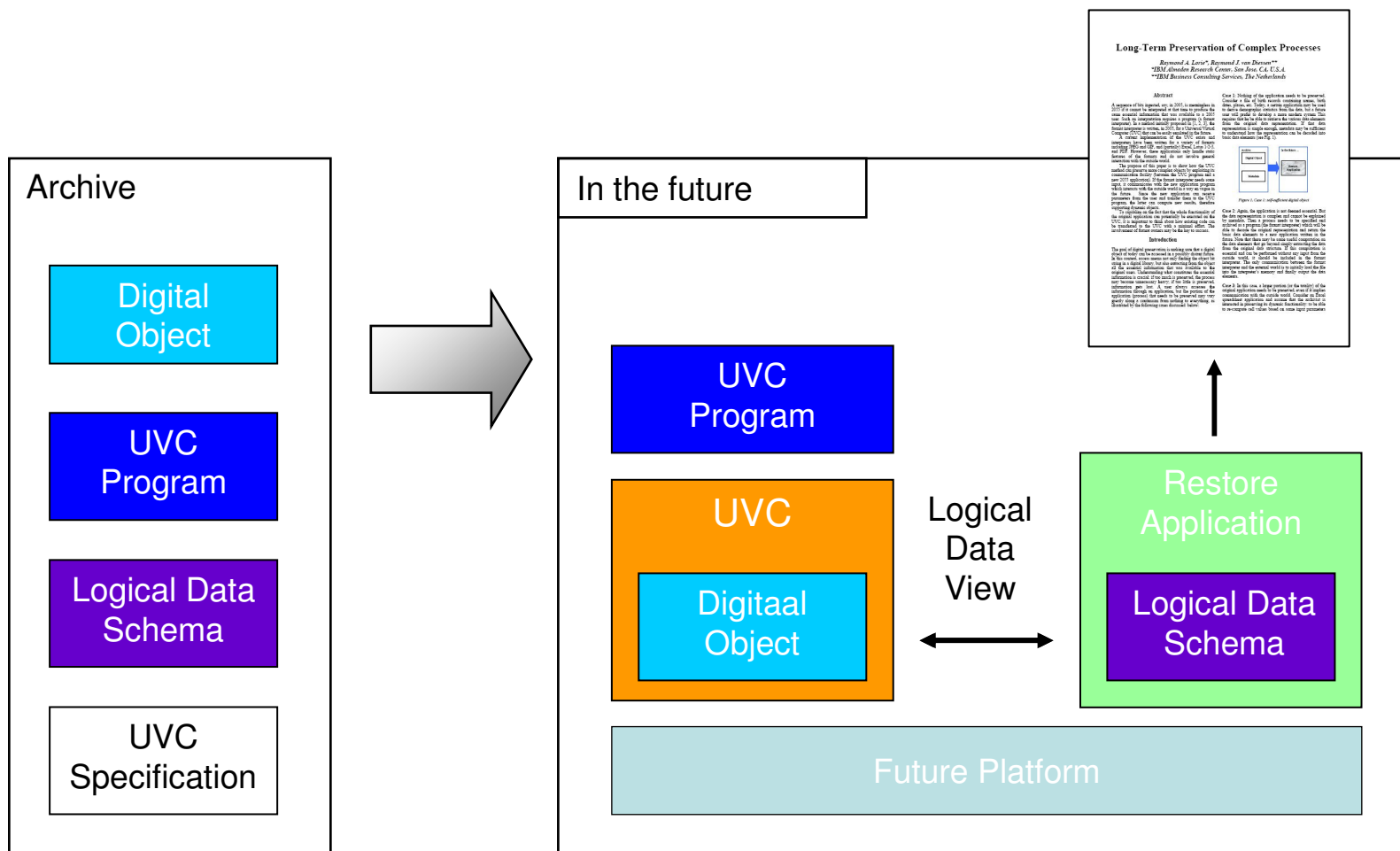
*By migrating “proprietary” formats into machine-independent descriptions, the informational content can be long-term preserved*



*“Long-Term Archiving of Digital Information”, Raymond Lorie, IBM Almaden Research Center, RJ 10185 (95059), March 28, 2000*

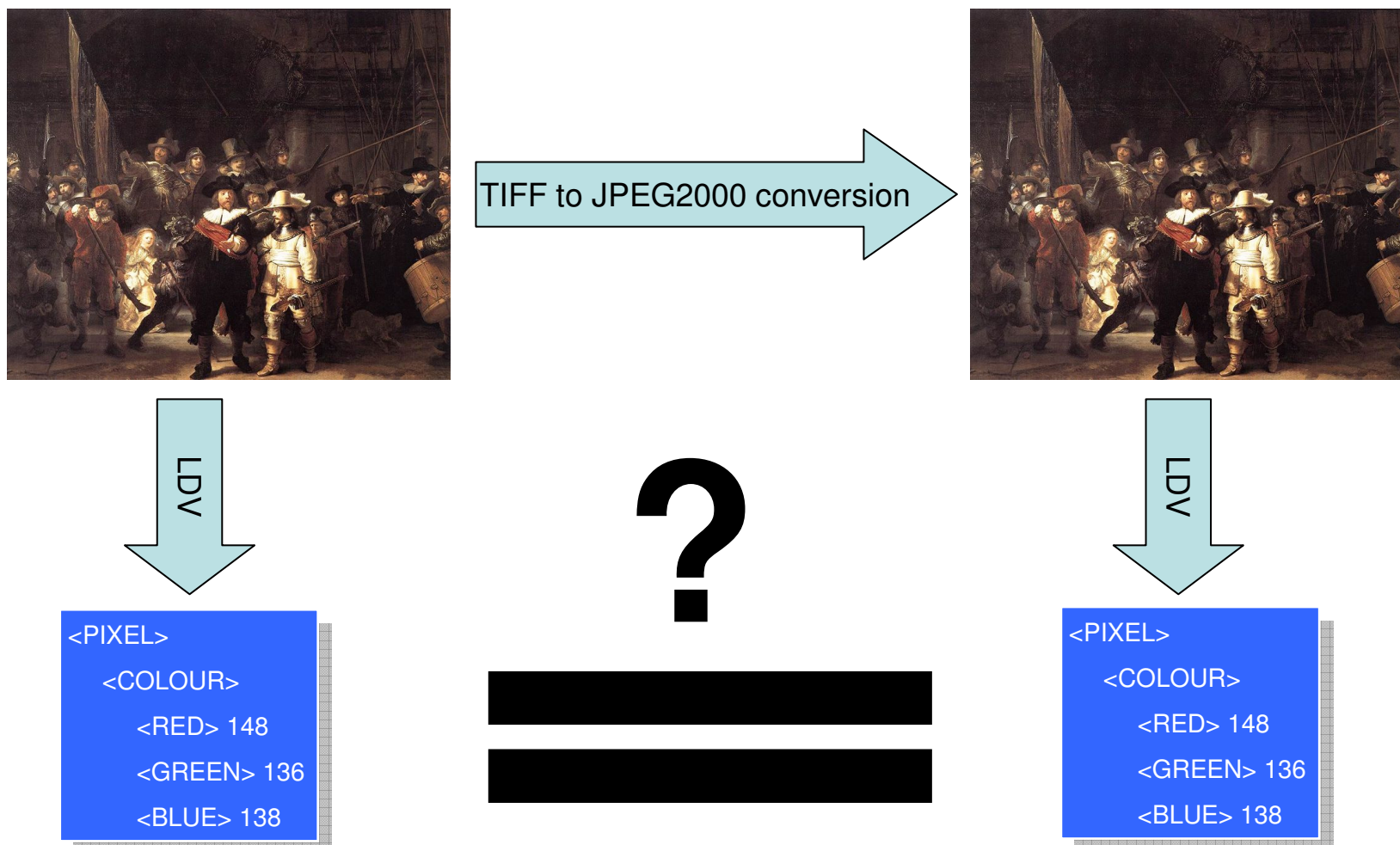


*Overview of IBM's UVC long-term Data Preservation process based on a combination of emulation (UVC) and migration (LDV)*



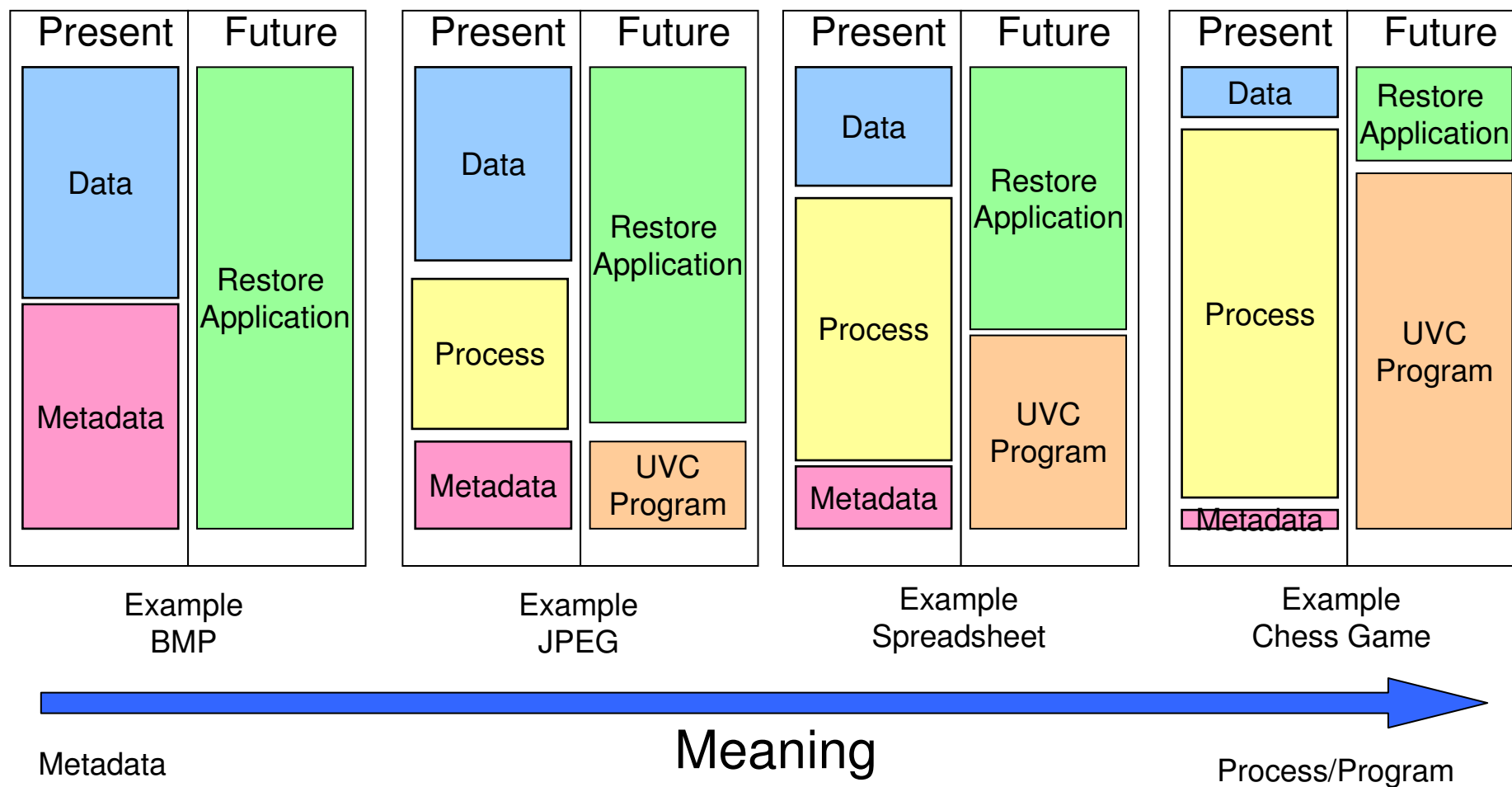
## Migration Validation

*The UVC Data Preservation approach could also be used to automatically validate format migrations*



## UVC Program Preservation

*The UVC vision intended to gradually also support the emulation of existing program, i.e. UVC Program Emulation*



# Planets UVC Spreadsheet Proof of Concept

The screenshot displays a VMware Player window titled 'Ubuntu' running a virtual machine. The desktop background is a landscape image. Three application windows are open:

- demo\_spreadsheet - Sharp Tools Spreadsheet:** A spreadsheet window with a menu bar (File, Edit, Table, Chart, Help) and a toolbar. It contains a table with renovation expenses.
- uvc@uvc-vmware: ~/uvc\_package/spreadsheet:** A terminal window showing the command `cat demo_spreadsheet.csv` and its output, which lists the same renovation expenses as the first window.
- UVC Spreadsheet:** A spreadsheet window with a menu bar (File) and a toolbar. It contains a table with renovation expenses, identical to the other two windows.

The terminal window output is as follows:

```
uvc@uvc-vmware: ~/uvc_package/spreadsheet$ cat demo_spreadsheet.csv
Expenses renovation    Roof    12000.0
      Kitchen    7800.0
      Bathroom    3200.0
      Garden    2100.0

      Total:    =C1+C2+C3+C4
```

The VMware Player window shows the system status bar at the bottom with the text 'To direct input to this virtual machine, press Ctrl+G.' and the VMware Player logo.





# Thank You