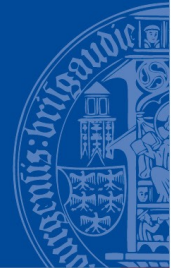


# Novel Workflows for Abstract Handling of Complex Interaction Processes in Digital Preservation

Klaus Rechert, University Freiburg

Albert-Ludwigs-Universität Freiburg



**UNI  
FREIBURG**



- The creation of most digital objects occurs solely in interactive graphical user interfaces.
- Many of these programs were designed as interactive applications, most without interfaces to automation.
- Cannot perform even simple operations like opening and saving a file in a different format in an unattended and fully automated task.
- Adding such functions is generally complex and costly, if at all possible.



- It becomes increasingly difficult to find suitable staff which is familiar with the ageing user environments of older computer systems.
- Human interaction in very repetitive tasks is error-prone and expensive.
- It is not always possible to rely on the availability of import functions of current programs or on migration tools.



Possible solution: Use a macro-recorder

Specialized tools to capture sequences of actions and replay these in an unattended manner.

But

Special tools require special knowledge.

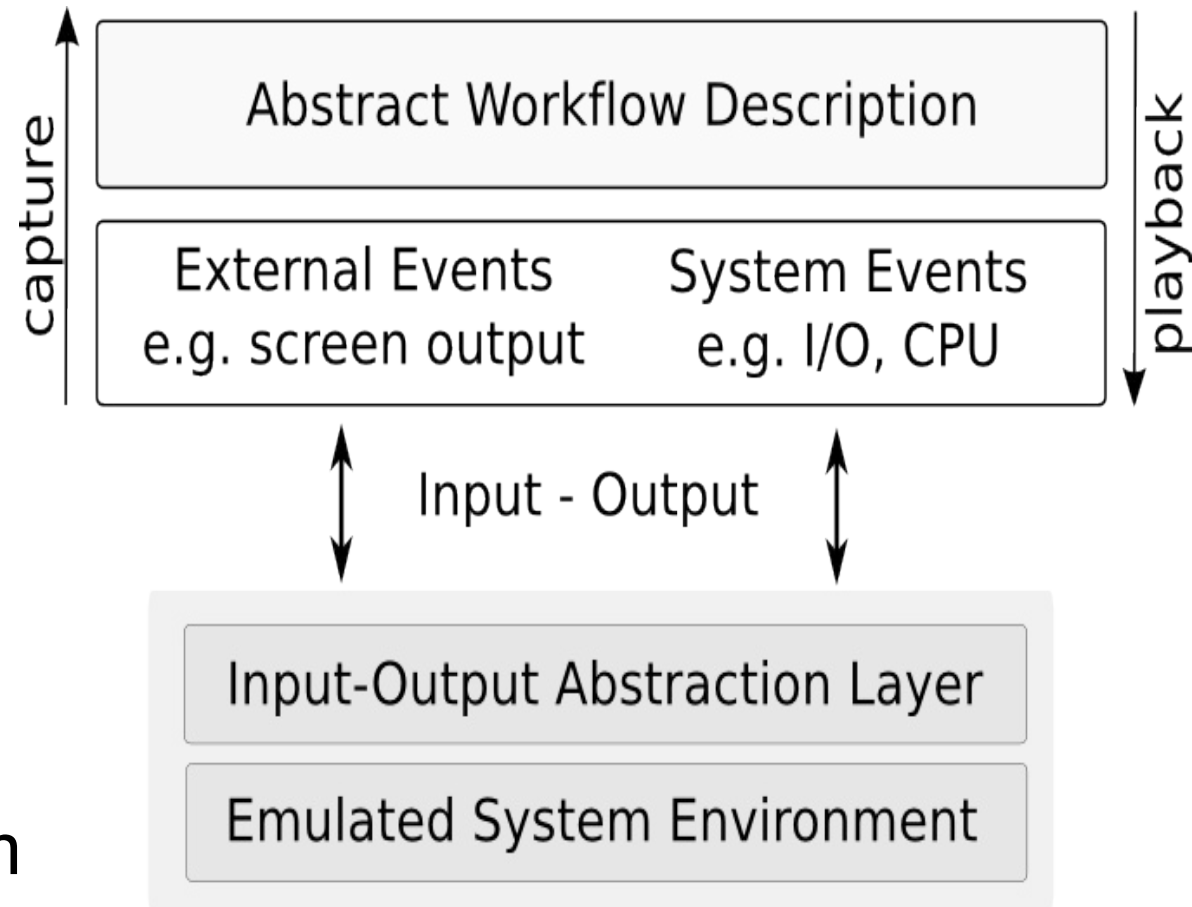
Specific for certain operation systems and might be unavailable on certain systems.

Difficult to get a generalized abstract description of recorded workflows.

No control over the host environment.

## Using emulation

- Technical and organizational separation
- Abstract workflow description
- Single format for all possible OS
- Control and monitor the emulated system



# Prerequisite: Data Exchange

Albert-Ludwigs-Universität Freiburg



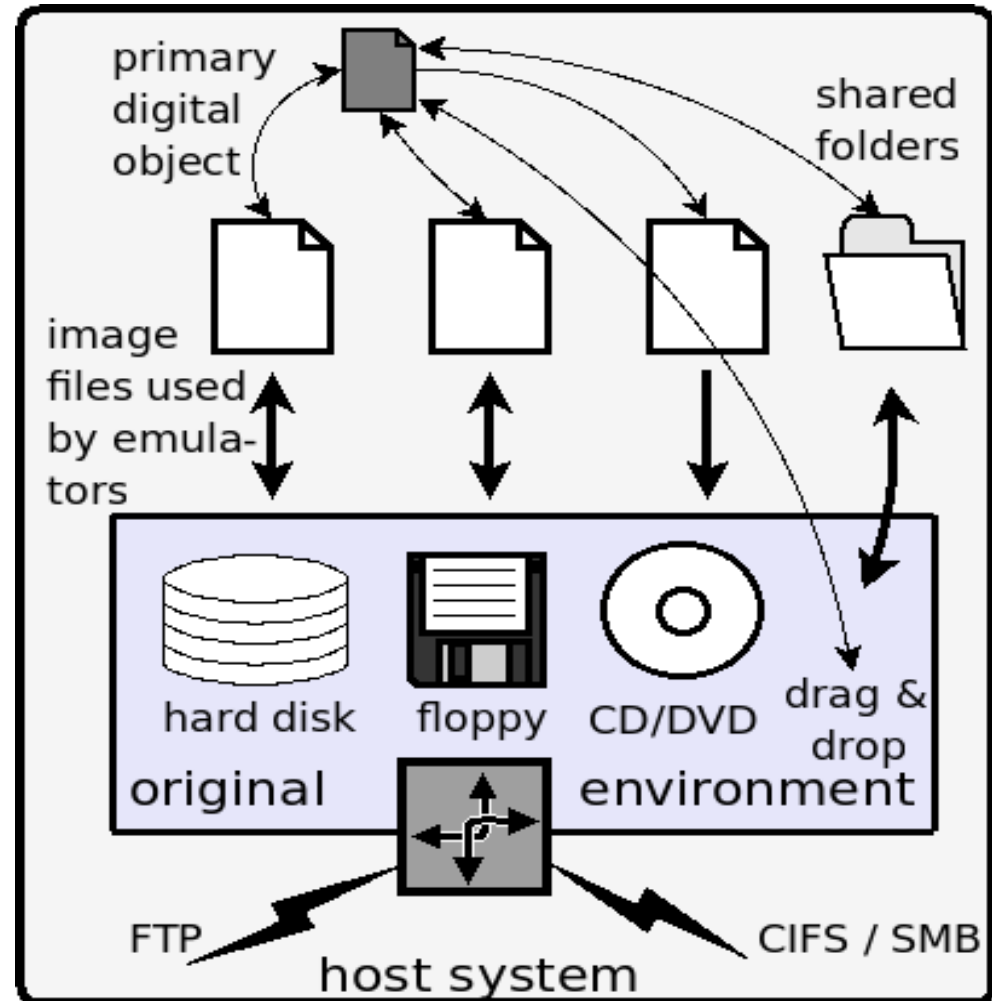
Virtual optical (ISO) or floppy disks as images

Disk container files

Network connections like FTP, SMB/CIFS

“Shared Folders” (as e.g. found in VMware or VirtualBox)

Copy&Paste (e.g. text areas in Dioscuri)



# Prerequisite: View-Path and Software Archive

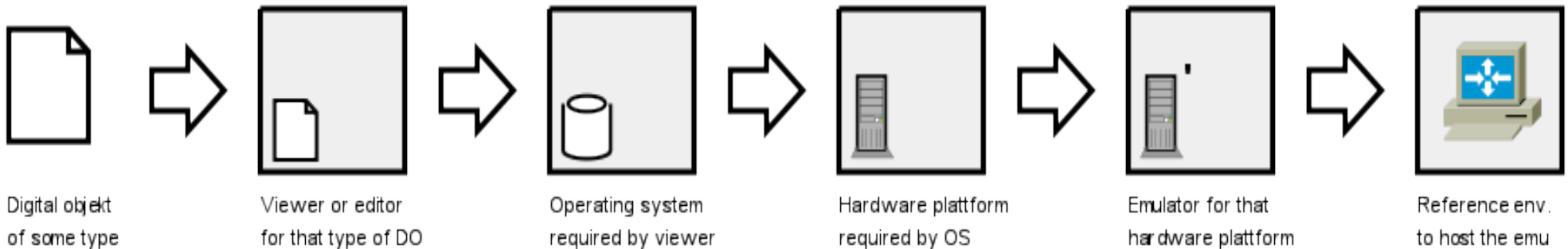
Albert-Ludwigs-Universität Freiburg



Digital objects (*primary objects*) can not be used by themselves.

Suitable context (hardware and software) is required (*secondary objects*).

A Software Archive containing all necessary secondary objects or prefabricated *View-Paths*.





An *Interactive Workflow* is ordered list of interactive events passed on to the emulated environment (e.g. mouse and keyboard events).

Each event is linked with a *precondition* and an *expected outcome*.

Built our solution on top of the *VNC-Play* tool, which offers visual synchronization points.



# Example Workflow

Albert-Ludwigs-Universität Freiburg



UNI  
FREIBURG

## Goal

Opening Lotus AMI Pro documents (SAM)  
and rendering them as PostScript documents.

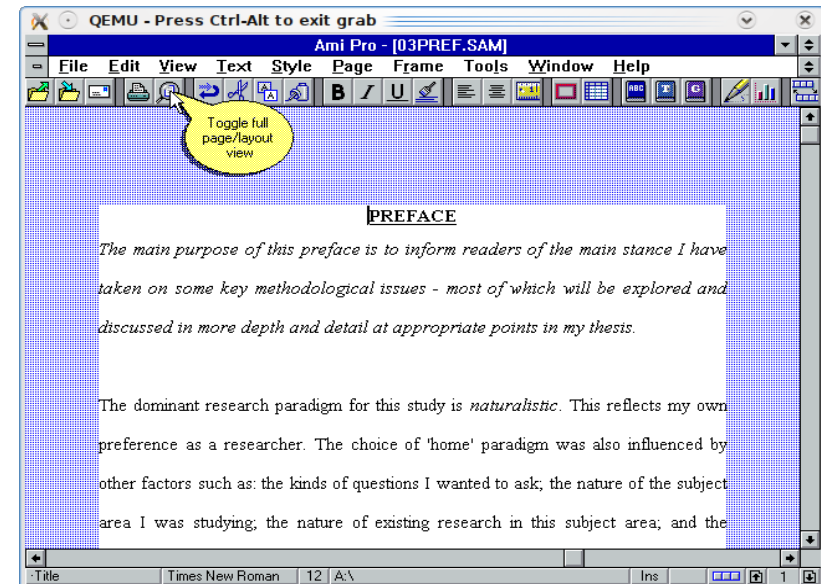
Requires View-Path consisting of:

DOS 6.20 with CD-ROM support

Win 3.11

Lotus AMI Pro

HP PostScript Printer Driver





## Aggregating the necessary View-Path

Boot a clean Win 3.11 installation.

Attach ISO-image with AMI Pro installation files and replay installation process.

Reboot the system.

Attach ISO-image with printer driver and replay installation.

Reboot the system.

Configure printer driver to print to A:\out.ps.



## Migrating SAM documents as batch-job

Create a floppy image with a single SAM file, attach it to the emulator.

Boot and wait for the system to be ready.

Open drive A: and open the first file.

Wait until the application is ready.

Print the file and shutdown.

Collect the PS file for further processing.

Repeat with next SAM file.



Reliable playback under different environments and conditions.

E.g. heavy I/O and/or CPU load.

Creating batch jobs migrating several documents in an unattended manner.

Aggregating View-Paths on demand and without human interaction.

**Note:** Creating reliable playbacks requires some training and carefulness.



## Recording-Toolbox

- Manual synchronization points
- Alter synchronization area and method
- Checkpoints, branches and failure strategies

## Preserve Knowledge

- Adding description of recorded steps

# Thank you!

Albert-Ludwigs-Universität Freiburg



UNI  
FREIBURG

## Questions?

