

Emulation in the limelight

Review of the iPres workshop - use cases,
feedback and lessons learned

*OPF Hackathon, Freiburg, Germany November
13 – 15, 2012*

Dirk von Suchodoletz, Klaus Rechert & Isgandar Valizada



Oh, no, not emulation!¹

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- First whole day event featuring emulation at iPres conference
- Rising importance of emulation in DP and access
- Ca. 50 attendees from a wide range of backgrounds



¹ Chris Rushbridge on his blog "Unsustainable Ideas"

- Workshop structured in four blocks
 - Introduction, Keynote “Beyond the controversy: a quick overview of emulation as an approach to digital preservation” (Titia van der Werf, OCLC)
 - Use-Cases and Challenges / Research and Practitioners
 - Working Groups 1st session
 - Working Groups 2nd session
 - Report, summary and outlook



Use Cases and Challenges

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Challenges by content holders – Practitioners
 - Access to multimedia and other non-standard objects in library collections (German National Library, Dutch National Library)
 - Digital Art (Karlsruhe University of Arts and Design)
 - Challenges of access to web harvested content, outdated formats, games (Danish National Library, British Library)
 - Metadata for emulation and original environments (University of Portsmouth, University of Cologne)
 - Challenges of scientific/research data (bwFLA)
 - Software preservation challenges of NLA

Common Challenges

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

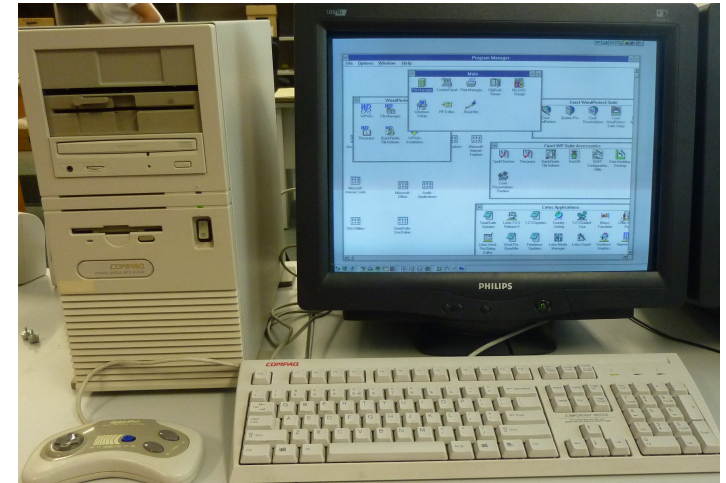
- Digital revolution changes material received by archives (spreadsheets, databases, email inboxes, end user programming, ...)
- New type of media because of legal deposit acts (computer games, apps, multimedia artefacts, ...)
- New forms of traditional published contents (e-books)
- Web archiving

- Ongoing research and projects
 - Use case: Digital Art
 - Use case: Development environments (software/industry)
 - Use case: Scientific/Business: Financial Data / Databases
 - Meta data for emulation workflows and solutions (TOTEM, ...)
 - Sustainable emulators and software archiving
 - Emulator and software archiving challenges (Brown, Indiana)
 - TIMBUS technical approaches to business processes (SBA Vienna)

Complete Systems

Albert-Ludwigs-Universität Freiburg

- Objects like
 - Famous persons machines
 - Scientific desktops (electronic lab-journal)
 - Databases
- Non-linear, user inter-action, multiple views





- The plenum agreed on four working groups
 - Digital Art / Computer Games
 - Metadata / Context
 - Webarchiving / Big Data
 - Sustainable Emulation / Community Involvement

- Museums spend millions on digital art
 - Often dynamic, interactive artefacts
 - Migration is seldom an option
 - Strategies required to preserve and display digital assets in the long run
- Many memory institutions receive digital legacy of famous authors, politicians, researchers, or wealthy donors (especially in the U.S. a much wider range of objects than usually found in Europe)

- Institutions see rising demand in new and versatile solutions
- Mismatch in institutional funding (gap between research and needs)
- What acceptable levels of digital art/game preservation (no guidance)
- What future users expect
- Legal aspects challenging: Game rights can be handed over from publisher, but those might not include rights on contained music

- Digital art and games more challenging than plain office material or traditional applications
- Artefacts usually try to make most of the hardware and/or use special effects which result from the specific hardware
 - Special screen, visual effects (achieved because of e.g. slow CRT refresh)
 - Non-standard use of floppy drives for e.g. sound or copy protection
 - Cutting edge features like 3D

Digital Art / Computer Games

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Solving emulation for art and games should solve most other emulation cases, or ...
- As some museum lady put it: “Digital art and games are the e-books of tomorrow”
- Actual e-book standards allow much more than plain PDF/A

Domain Specific Challenges

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

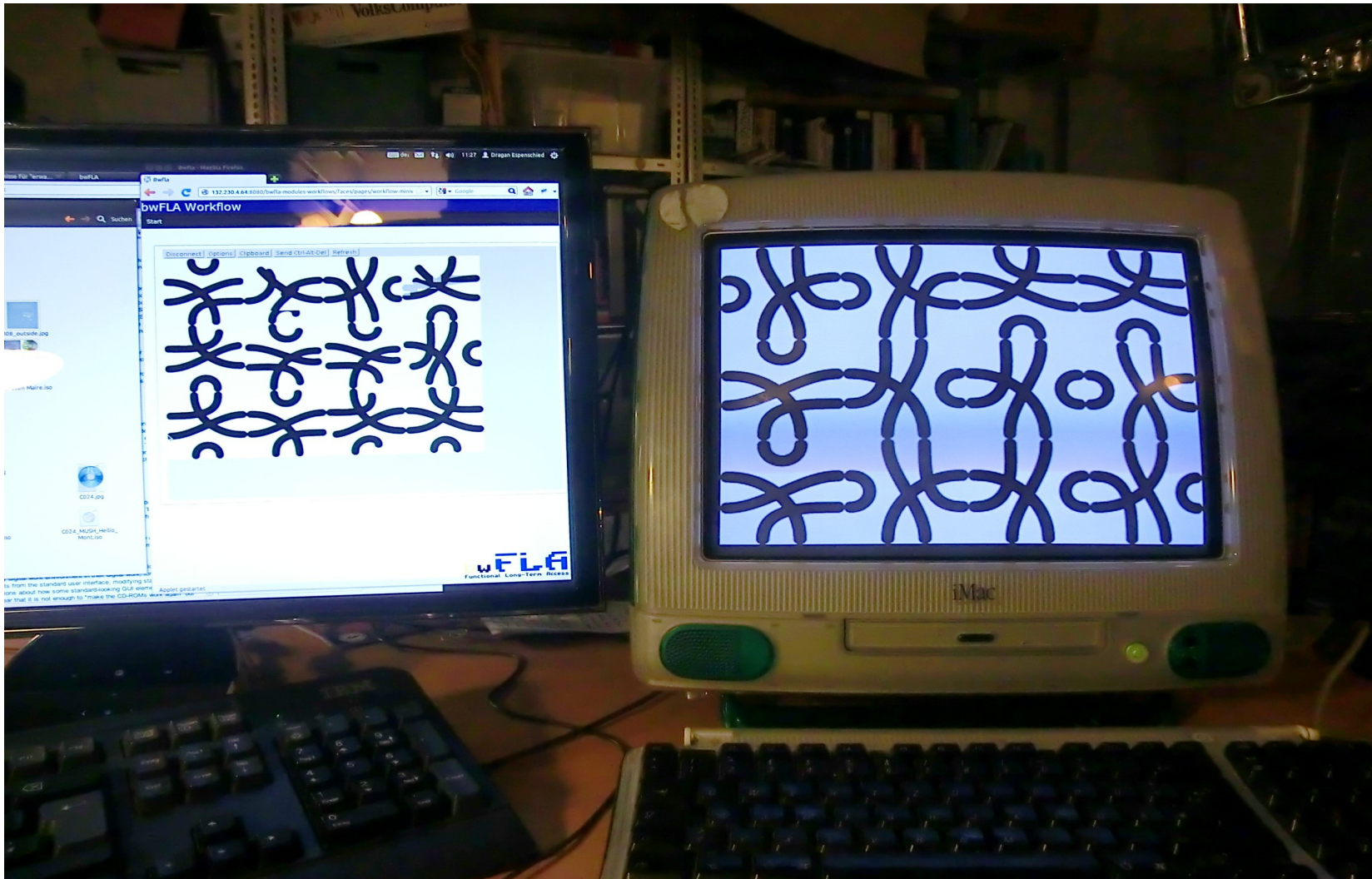
- Screen type, resolution and color depth oft crucial
 - A picture may look pretty different on an old (black&white) PAL TV set, 15" CRT in 1024x768 or on a modern 24" TFT offering 1920x1080
- Non standard types of input like audio or video for digital art or different types of joysticks for games
- Real time requirements in user interaction (time lags, execution speed of art works or games, ...)
- May require additional type of technical meta data

Screen / Environment Challenge

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG



Domain Specific Challenges

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Special challenges of preserving the “intent” of an artist
 - Does the artist intend his/her artwork to be “eternal”, may the artist “allow” emulation
 - Should the acquiring institution get a waiver of the artist consent to modifications for long term preservation
 - What kind of modification is acceptable, which not
 - What kind of additional meta data, documentation is required
 - Additional technical requirements

- When dealing with original environments descriptive meta data is required
 - Describe software and hardware components
 - Describe complete original environment in an unambiguous way
 - Preserve software licenses and documentation
 - Acquire tacit knowledge of operation
- Map digital artefacts to it's original environments

- Traditional preservation strategies often lose object's context
 - Single out objects
 - Miss linked in components
 - Lose context within original filesystem and working environment
- Business processes, research environments, (software) development systems usually consist of specially configured software/hardware/network environments

Meta Data and Registries

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Traditional registries incomplete, do not honor object and software components dependencies
 - Unusable to produce object view paths from it
 - Do not describe standard software items and machines

- Web archiving and big data (scientific data management) more and more relevant topics for memory institutions
- Legal deposit requires access for users to archived material over long periods
- Increasing complexity of object types like Flash or SilverLight, JavaScript, ...
- Difficult to “plan” or restrict acceptable object types
- Old formats refuse to fade away (PPT 4.0 in BL crawls)

- Modern sites often complex to archive, often crawls cannot be complete
- Institutions need to think about future use cases of web archives and be pragmatic how they invest their resources including access strategies
 - Emulation seems inevitable, what quality is good enough (for how long periods)
 - More meta data might be required to be gathered for future access (e.g. screenshots of pages in popular browsers)

- Browser environments and rendering characteristics change over time
 - Try to access a *wayback machine* archived page in a modern browser – does not necessarily reproduces original look&feel
- Instead of programming special purpose browsers use emulation to:
 - Preserve old rendering environments including virtual network access
 - Get access to deprecated formats like old documents or RealMedia audio streams

Big Data Challenge

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Big data – how defined, limited?
 - Datawarehousing, linked data
 - Scientific primary research data
- Some memory institutions started to look into large scale (research) data management (using storage strategies like Apache Hadoop technology)
- Proper data description / meta data strategic for preservation, access and re-use

Big Data Challenge

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Data often meaningless without context like creating application or associated processes
- Emulation can help to keep creation environments complete and re-run original tools
- Pretty early stage of discussions, generally concerns how migration or emulation strategies should look like

- Emulation based preservation and access strategies heavily depend on
 - Availability of suitable emulators (or VM's)
 - Archived versions of original environments consisting of a wide range of software components
 - Maintain matching hardware driver sets (significant challenge e.g. for preservation of MS operating systems)
 - Certain hardware toolsets, e.g. for disk imaging

- Emulation still approach no 100% replication of the original system/environment
- Emulators subjective because of programmers personal interest/passion
- Many emulators depend on a very small group of people
- Passions on platforms not evenly distributed (famous vs. rare, unpopular systems)
- Lack of institutional interest, no big investments (beside the Dutch) by memory institutions in emulations

- Two strands of emulation toolsets
 - Broad range of open source hardware emulators in different states of quality and usability
 - Often driven by games communities
 - Hobbyist approach – seen as a “hack” only by some memory institutions
 - Commercial approach (mostly virtual machines, preferred by companies, no full emulation) with limited scope
 - Memory institutions barely visible yet
- Pretty much three fractions without much interaction and interlinkage

Emulation Communities

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Need a clear set of characteristics and designated communities for quality ensurance
 - Much better interaction between communities required

Legal Challenges of Emulation

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

- Legal challenges of emulation
 - Run old environments in emulation
 - Required proprietary firmwares/bios/... for emulators
 - Software licenses not mentioning emulation
 - Break copy protection (to e.g. preserve software)
- Discussion/definition of what is (il)legal
 - Lead function of National Libraries & Archives
 - Library law may permit more action than general regulation
- Owning data in proprietary formats (what does this mean for memory institutions)

- Wrap-up and conclusions by Bram van der Werf of OPF
- Challenges are interdisciplinary
 - Curational
 - Legal
 - Technical
- Need to have an emulation agenda

- Migration as a strategies pushed it 's limits, emulation extends them significantly
 - Big institutions may have not hit their crisis point yet
 - Database preservation is poorly done
 - Discussion of authenticity via significant properties
 - Produce good use cases to demonstrate need and usefulness of emulation
 - Fall back solutions of objects failed in migration

- New use cases expected from
 - Digital art
 - 3D modeling
 - Idiosyncratic databases
 - Range of custom made/special applications
 - Long term accountability for auditing
- Available tools should be put out for demonstration to see take up/denial by (institutional) users to build upon this

- Data management plans now central for many (research) funding agencies
- Quality of available tools and workflows in emulation still a challenge
- Sustainability of tools and services a major requirement
- Couple of research projects up to now, but production stable, usable tools and services still missing

- Legal issues pretty much a mystery
 - To be addressed via lobbying
 - Technical issues should be split from legal ones
- Create consortium of relevant libraries and archives
 - Distributed approach for archiving standard software
 - Use similar established technology for different tools and services

- New challenges of platforms like the iPad
 - Heavily secured, protected, cannot be easily emulated
- Preservation of (networked), dynamic, permanently changing data/objects like business processes
 - Distributed systems (network P2P, client-server applications)
- Dependence on external data sources
- And: Tools for emulation are software and need to be updated over time

Thank you for your Attention!

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

Questions / Comments?

Dirk von Suchodoletz

Klaus Recher

Isgandar Valizada

Chair on Communication Systems

Faculty of Engineering

University of Freiburg / Germany

{givenname.surname}@rz.uni-freiburg.de

