



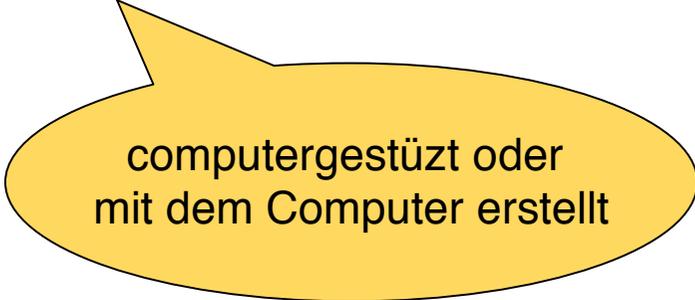
# Langzeit-Archivierung und Datensicherung

Max Riegel

2007-11-25

## ■ **Dokumente und Medien werden zunehmend 'digital'**

- e-mail anstelle von verschicktem Papier
- elektronische Steuererklärung
- Adressbuch und Kalender
- Notizen und Tagebücher
- digitale Fotografie
- MP3-/iPod Music Player anstelle CD
- digitale Camcorder

A yellow speech bubble with a black outline, pointing towards the list of digital media examples. It contains the text 'computergestützt oder mit dem Computer erstellt' in black font.

computergestützt oder  
mit dem Computer erstellt

## ■ **Die 'Digitalisierung' erfasst auch das Privatleben**

- Festplatten ersetzen die Papierberge

## ■ **Unsere Identität wird digital**

- Identität ergibt sich aus Erfahrung und Erinnerungen
- Erinnerungen sind nur noch digital dokumentiert

## ■ **Wie sichern und schützen wir in der Zukunft unsere Identität?**

- 'lebenslang' ist länger als der Planungshorizont der IT Industrie

## ■ Prolog

----

## ■ Grundlagen der Archivierung

- Revisionssichere elektronische Archivierung
- Elektronische Langzeitarchivierung
- Haltbarkeit von Datenträgern
- Sicherstellung der Verfügbarkeit

## ■ Langzeitarchivierung im privaten Bereich

- Lösungsansätze
  - Dateiformate
  - Systemwechsel
  - Dateiablage
- Datenvolumen und Wachstum

## ■ Integration von Backup und Langzeitarchivierung

# Grundlagen: Revisionssichere Archivierung



## ■ **Sichere und ordnungsgemäße Aufbewahrung von kaufmännischen Dokumenten entsprechend HGB §239 und §257 sowie der Abgabenordnung für eine Frist von 6-10 Jahren**

1. Jedes Dokument muss unveränderbar archiviert werden
2. Es darf kein Dokument auf dem Weg ins Archiv oder im Archiv selbst verloren gehen
3. Jedes Dokument muss mit geeigneten Retrievaltechniken wieder auffindbar sein
4. Es muss genau das Dokument wiedergefunden werden, das gesucht worden ist
5. Kein Dokument darf während seiner vorgesehenen Lebenszeit zerstört werden können
6. Jedes Dokument muss in genau der gleichen Form, wie es erfasst wurde, wieder angezeigt und gedruckt werden können
7. Jedes Dokument muss zeitnah wiedergefunden werden können
8. Alle Aktionen im Archiv, die Veränderungen in der Organisation und Struktur bewirken, sind derart zu protokollieren, dass die Wiederherstellung des ursprünglichen Zustandes möglich ist
9. Elektronische Archive sind so auszulegen, dass eine Migration auf neue Plattformen, Medien, Softwareversionen und Komponenten ohne Informationsverlust möglich ist
10. Das System muss dem Anwender die Möglichkeit bieten, die gesetzlichen Bestimmungen sowie die betrieblichen Bestimmungen des Anwenders hinsichtlich Datensicherheit und Datenschutz über die Lebensdauer des Archivs sicherzustellen

## ■ ***Private Anforderungen gehen über 6..10 Jahre hinaus...***

# Grundlagen: Elektronische Langzeitarchivierung



- Erfassung, langfristige Aufbewahrung und Erhaltung der dauerhaften Verfügbarkeit

*„Langzeit‘ bedeutet für die Bestandserhaltung digitaler Ressourcen nicht die Abgabe einer Garantieerklärung über fünf oder fünfzig Jahre, sondern die verantwortliche Entwicklung von Strategien, die den beständigen, vom Informationsmarkt verursachten Wandel bewältigen können.“*

- Zeithorizont >10 Jahre

- Im Gegensatz zum Backup, das die Daten nur kurzzeitig sichert

- Probleme:

- Schneller Medien- und Systemwandel
  - Veraltete Datenformate (Programme)
  - Veraltete Datenträger und Dateisysteme
- Auffinden von Informationen
- Datenkonsistenz
- Haltbarkeit von Trägermedien
  - siehe nächste Folie

# Grundlagen: Haltbarkeit von Trägermedien



## Lebensdauer einiger Datenträger

Medium	Erwartete Lebensdauer	Speicherdichte (kbit/kg)
Steintafeln und Steinmalereien	mehrere tausend Jahre (gesichert)	0,001...1
Nickelplatte	mehrere tausend Jahre (vermutet)	
Bücher und Handschriften aus säurefreiem Papier und mit säurefreier und nicht eisenhaltiger Tinte	mehrere hundert Jahre (gesichert)	3.000...30.000
Schwarzweißfilme aus Polyethylenterephthalat (PET)	bis zu 1000 Jahre (vermutet)	
Bücher und Handschriften aus säurehaltigem Papier	70 - 100 Jahre	3.000...30.000
Filme auf Zelluloid	mindestens 50 bis 70 Jahre, laut Hersteller mehrere hundert Jahre	
Herkömmliche Bücher	100 - 200 Jahre	
Optische Speichermedien CD-ROM / DVD	CD-R: 5-10 Jahre DVD-RAM: 30 Jahre	CD-ROM: $4 \times 10^8$ DVD: $2..4 \times 10^9$
Zeitungspapier	10 - 50 Jahre	
Disketten	5 - 10 Jahre	
Magnetbänder	bis zu 30 Jahre (gesichert)	
Iomega REV-Wechsellaufwerk	bis zu 30 Jahre (vermutet)	
USB-Stick	3 - 10 Jahre	

# Grundlagen: Sicherstellung der Verfügbarkeit



- **Standardisierung**
  - Standards für Dateiformate und Dateisystemformate, z.B.: MP3, TIFF, JPEG, PDF bzw. ISO 9660 (CD-ROM)
- **Migration**
  - Überführung der Informationen in eine neue Systemumgebung, z.B.: Mac OS 9 -> Mac OS X
- **Emulation**
  - Simulation eines älteren Systems in einer neuen Umgebung, z.B.: Classic in Mac OS X
- **Kapselung**
  - Archivierung der Laufzeitumgebung zusammen mit den Informationen, z.B.: Daten + Programm + Betriebssystem; virtuelle Rechner
- **Konvertierung zur Laufzeit**
  - Einsatz von Konvertern und Viewern zum Zugriff auf 'alte' Datenformate, z.B.: GraphicConverter auf Mac oder IrfanView auf Windows

# Langzeitarchivierung im privaten Bereich

## Lösungsansätze zu Dateiformaten



- **Standardisierte Datenformate verwenden**
  - viele Programme können mit den gleichen Daten umgehen
    - z.B.: TEXT, HTML, PDF, JPEG, PNG, MP3
  - hilft gegen beständiger Wechsel der Programme und Systeme
    - Größtes Problem bei der Langzeitarchivierung!
  
- **Standard Viewer-Format Kopie zusätzlich zur Originaldatei**
  - z.B.: PDF
  - wenn es keine 'brauchbaren' Standardformate gibt,
  - besondere Programme zum Einsatz kommen sollen
  - die Originaldateien Abhängigkeiten zur Systemumgebung haben
    - z.B. Druckereinstellungen bei Textverarbeitungsprogrammen

# Langzeitarchivierung im privaten Bereich

## Lösungsansätze zur Systemumstellung



- **Sorgfältige Planung bei Programmwechsel und Systemumstellung**
  - keine voreiligen Entscheidungen
    - möglicherweise ist die Weiterverwendung der alten Programme besser
  - automatisierte Erstellung von Viewer-Kopien durch Scripte
  - Wie kann die alte Systemumgebung wieder hergestellt werden?
    - Archivierung der Hardware ist keine Lösung!
  
- **Einsatz von Virtualisierungstechniken (Emulation, Kapselung)**
  - Virtualisierung erlaubt die Realisierung von virtuellen Rechnern
  - Als Gast-System kommt das 'alte' Betriebssystem mit dem 'alten' Programm zum Einsatz.
    - auch die Originaldateien können im Gast-System abgelegt werden (Kapselung)
    - Datenaustausch mit dem Host-System sicherstellen!
  - Zum schnellen Zugriff werden Viewer-Kopien erstellt und direkt im Host-System abgelegt

# Langzeitarchivierung im privaten Bereich

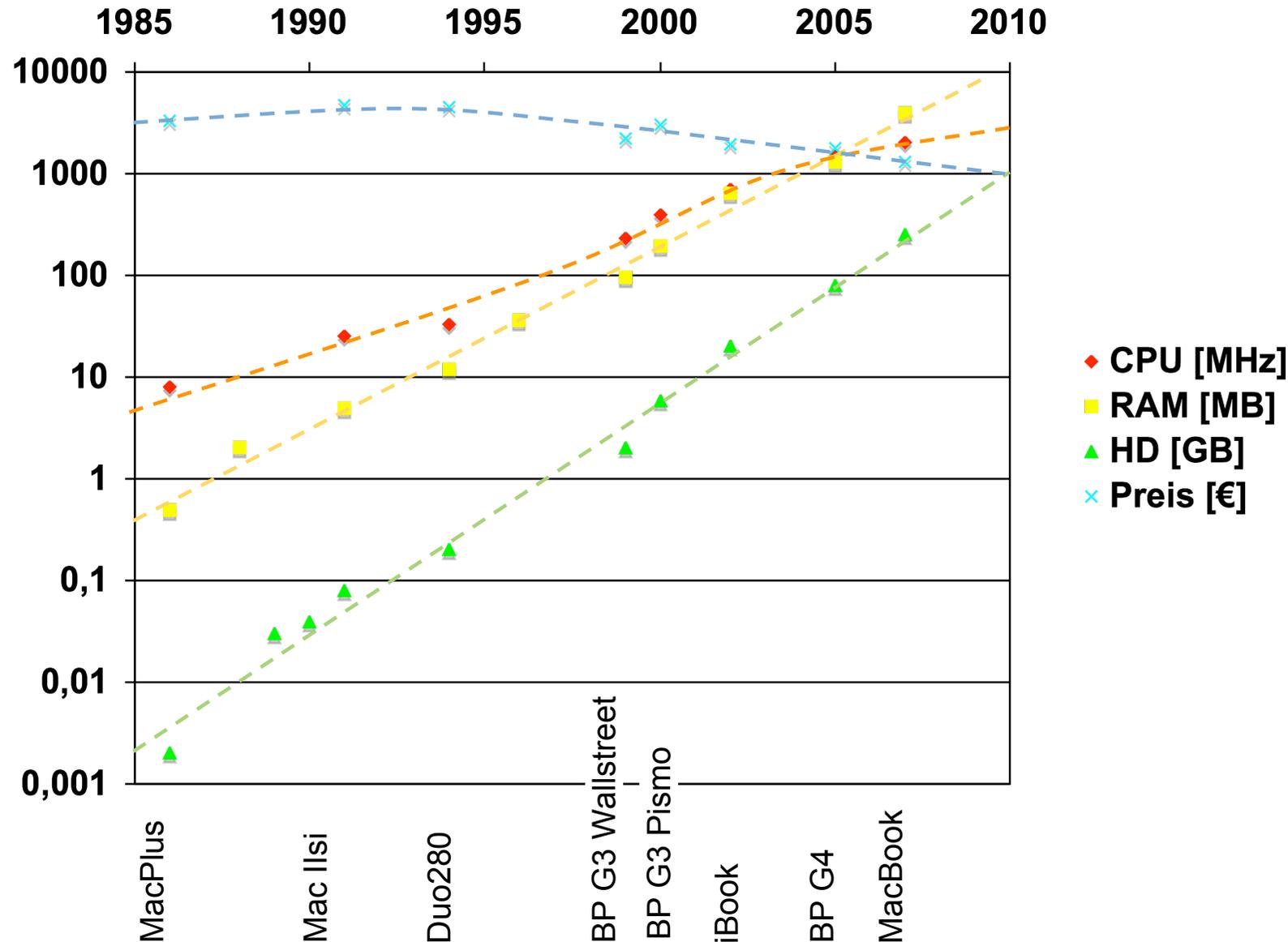
## Lösungsansätze zur Dateiablage



- **Dedizierte Ablage für Langzeitarchivierung**
  - Spezielle Anforderungen an das Speichermedium
    - langzeitstabil
    - betriebssystemunabhängig
    - katastrophensicher
    - kostengünstig
  - Lösungen für den Hausgebrauch
    - CD-ROM: Speicherkapazität für heutige Datenmengen unzureichend
    - DVD: Systemwirrwarr, Kapazität für heutige Verhältnisse knapp
    - Speicherung im Netz: im Prinzip sehr gut, aber wenig Erfahrung mit Stabilität der Anbieter und hohe Anforderungen an die Bandbreite  
z.B.: <http://www.jungledisk.com/>, Kosten ca. 2 US\$/GB/Jahr
- **Integration der Langzeitarchivierung in eine Backup Lösung**
  - Datenvolumen wachsen so schnell, dass die zusätzliche Datenmenge für die Langzeitarchivierung vernachlässigbar wird

# Datenvolumen und Wachstum

## Meine persönliche Geschichte



# Integration von Backup und Langzeitarchivierung



- **Mit rsback&rsync gibt es inzwischen hervorragende Tools zum automatisierten Backup auf externe Festplatten oder Netzlaufwerke**
  - siehe Vortrag von Hans-Jürgen Beie beim KNF-Kongress '07
  - 'TimeMachine' ist eine entsprechende kommerzielle Implementierung
- **Idealerweise wird das lokale Backup noch 'remote' gespiegelt, um die gewünschte Katastrophensicherheit zu erreichen**
  - zumindest entferntes Gebäude, besser noch über Netz weiter weg
    - Flatrate DSL macht es für Privatpersonen erschwinglich
    - Privatinitiative unter Freunden hilft Geld sparen
- **Backup Systeme müssen alle paar Jahre erneuert werden**
  - mehr Kapazität wegen gewachsenem Datenvolumen
  - Festplatten leben nicht ewig - aber normalerweise sind sie voll bevor sie ausfallen;-), wenn man kostenbewusst investiert
- **Mit dem Backup System wird implizit auch das Langzeitarchiv auf aktuelle Soft- und Hardware migriert**
  - hält Langzeitarchiv immer auf einem aktuellen Dateisystem

- **Private Langzeitarchivierung ist mehr als Backup.**
- **Größtes Problem ist der beständige Wechsel der Systeme und Programme.**
- **Mehr Standardisierung bei den Dateiformaten notwendig**
- **Virtualisierung erlaubt elektronische Archivierung alter Hardware.**
- **Ein automatisiertes Backup System auf Netzlaufwerke bietet auch eine sichere Dateiablage für die Langzeitarchivierung.**
- **Es ist hilfreich, dass beim derzeitigen Datenwachstum Festplatten schneller voll als defekt sind;-).**
- **Umkopieren hält Daten jung!**

***Vielen Dank für Eure Aufmerksamkeit!  
Fragen, Anmerkungen?***