

Langzeitarchivierung digitaler Sammlungsdaten in Kultureinrichtungen und Museen

Werner Schweibenz



Übersicht

- Einführung
- Vergleich Bibliothek – Museum
- Konsequenzen für Museen
- Archivierungstechniken
- Langzeitarchivierung im KHI
- Zusammenfassung
- Diskussionsvorschläge



Einführung

“Die Langzeitarchivierung digitaler Dokumente ist zum Problem geworden, das noch dadurch gesteigert wird, daß dieses Faktum im Allgemeinen gleichmütig hingenommen wird.”

(Hermann Leskien,
Generaldirektor der Bayerischen Staatsbibliothek)



Einführung

Szenarien von Rothenberg (RAND):

- Brief mit Schatzkarte auf Papier
- Brief mit Schatzkarte auf CD-ROM

Schlussfolgerungen:

- Langzeitarchivierung als unterschätztes Problem
- Gefahr des heraufziehenden „digital dark age“ wegen Datenverlusten?



Einführung

- Handschriften und Bücher können ohne Hilfsmittel gelesen werden.
- Für den Zugriff auf digitale Objekte benötigt man technische Hilfsmittel.
- Technische Hilfsmittel sind problematisch, wegen rapidem Wandel.
- Beispiele:
 - 5 1/4-Zoll-Disketten
 - Texte, die mit IBM-Text4 geschrieben wurden
 - Daten, die in Geoworks erfasst wurden



Einführung

- Beispiele für Datenverluste in den USA:
 - NASA- Daten über die Entwicklung des brasilianischen Regenwaldes
 - etwa 10.000 Datensätze aus der Volkszählung von 1960
- Beispiele für Probleme ohne Metadaten:
 - Bild-, Audio- oder Bewegtbilddateien per se sind nicht „sprechend“ und identifizierbar
 - Nutzen ist eingeschränkt

Vergleich

Bibliothek - Museum

- Objekte sind vielfach vorhanden
 - Metadaten werden zentral erstellt (DDB-Service)
 - Daten werden in Verbunddatenbanken gehalten
 - Daten können zentral archiviert werden
- Objekt ist i.d.R. einmalig oder selten
 - Metadaten müssen im Museum erstellt werden
 - Daten werden häufig nur vor Ort gehalten
 - Daten werden vor Ort archiviert (wenn überhaupt)

Vergleich

Bibliothek - Museum

- Musterdatensatz:
Die Deutsche
Bibliothek verzeichnet
diese Publikation in
der Deutschen
Nationalbibliographie;
detaillierte
bibliographische Daten
sind im Internet über
<http://dnb.ddb.de>
abrufbar.
- Musterdatensatz:
???



Konsequenzen für Museen

- Dokumentation im Haus verbessern
- Anschluss an Datenverbünde suchen
- Beachtung von Metadatenstandards für Datenaustausch und Archivierung
- Suche nach (gemeinsamen) Archivierungsverfahren und -techniken



Archivierungstechniken

- Erstellung von Hardcopies
- Erhaltung von Hard- und Software
- Datenmigration
- Emulation
- Hybridlösungen:
 - Kurzzeitarchivierung in digitaler Form
 - Langzeitarchivierung auf Mikrofilm



Langzeitarchivierung im KHI

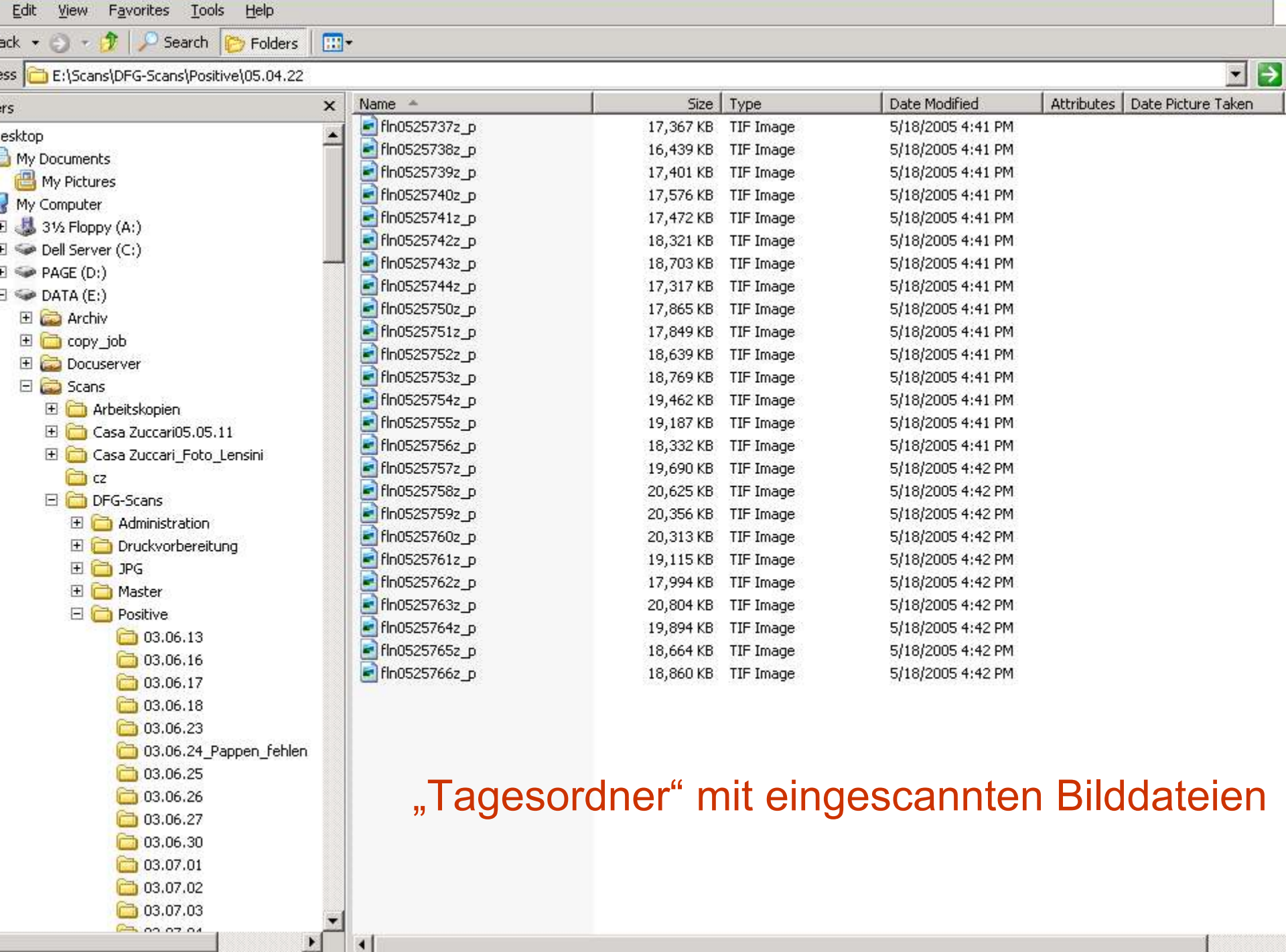
- Das KHI, gegründet 1897, ist seit 2002 Teil der Max-Planck-Gesellschaft (MPG)
- Sammlung der Photothek umfasst ca. 580.000 Fotos, ca. 20.000 sind digitalisiert
- Frage nach der Langzeitarchivierung



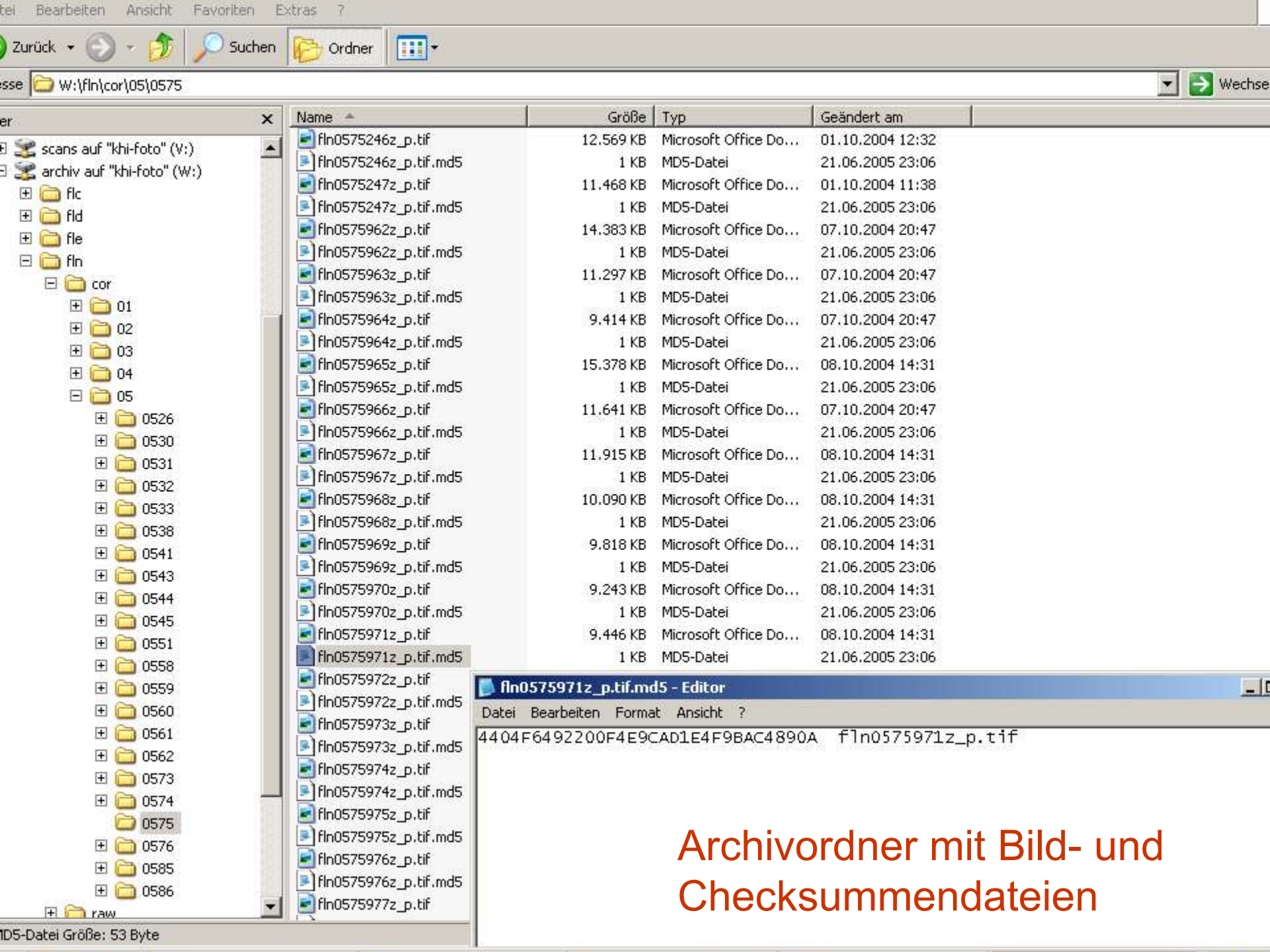
Langzeitarchivierung im KHI

Datenfluss bei der Langzeitarchivierung:

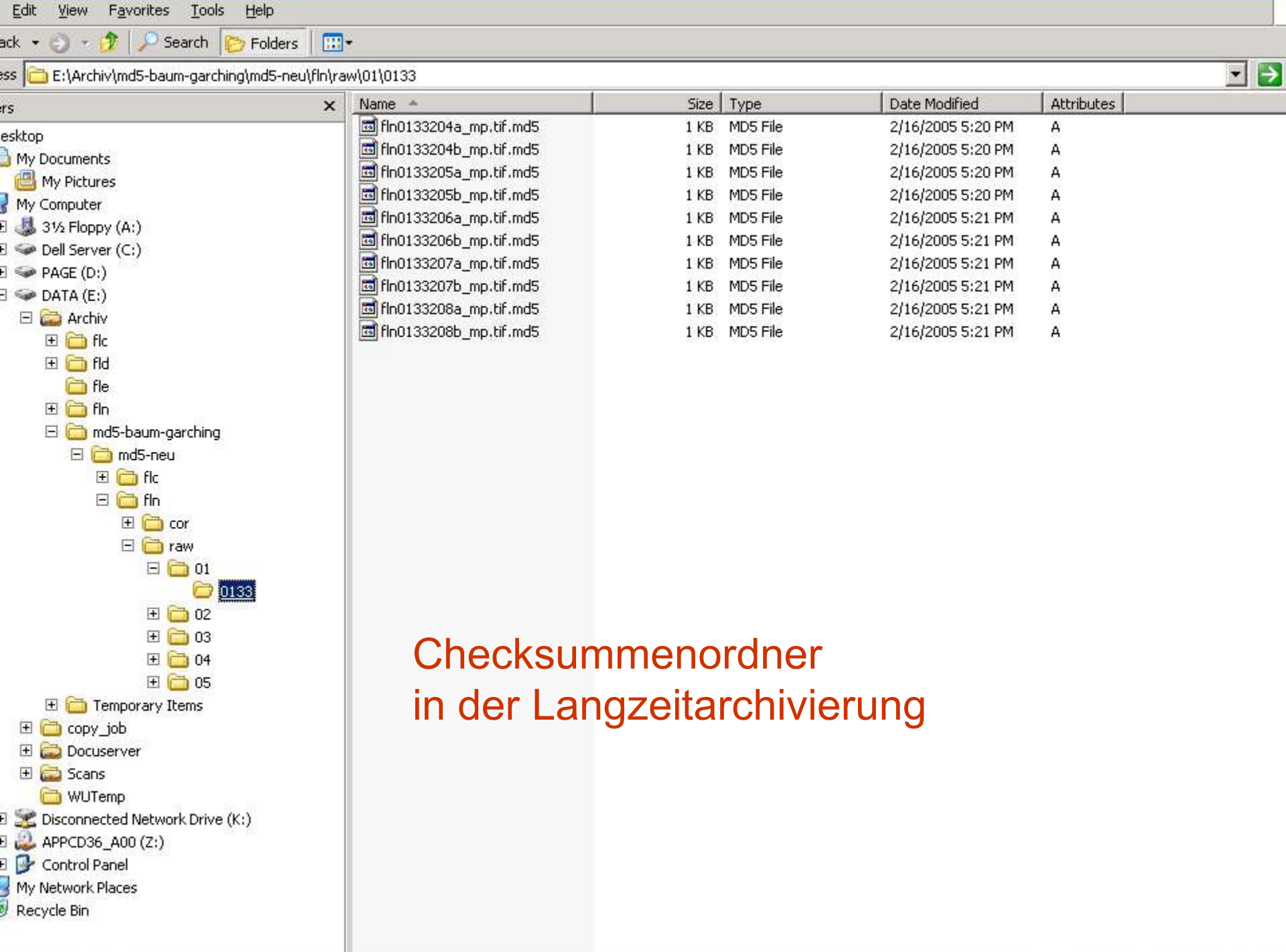
- Masterscans und bearbeitete Scans in „Tagesordnern“
- Übertragung in Archivordner
- Erzeugung von Checksummendateien
- Erzeugung von Metadatendateien aus HiDA (geplant)
- Übertragung in Langzeitarchivierung per Platten- und Onlinetransfer



„Tagesordner“ mit eingescannten Bilddateien



Archivordner mit Bild- und Checksummendateien



Checksummenordner in der Langzeitarchivierung

Bearbeiten Ansicht Einfügen Format Extras Tabelle Fenster ?

Tabellengitternetz Times New Roman 12 F K U

130%

gültige Version enthält Markups Anzeigen

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

| Feldnr. in HiDA | Feldname in HiDA | Alternativ falls nicht vorhanden | Anmerkung |
|-----------------|--------------------|----------------------------------|--|
| 5000 | Dokument-Nr. | <input type="checkbox"/> | Stellt ggf. Bezug zu HiDA und damit zu weiterführenden Informationen |
| 5100 | Künstlername | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 5130 | Entstehungsort | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 5064 | Datierung | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 5108 | Ort | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 52700 | Bauwerk | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 52664 | Standort | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 52864 | Verwalterort | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 52900 | Sozietätname | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 5064 | Datierung | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 5108 | Ort | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 52200 | Objekttitel | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 52202 | Bauwerkname | <input type="checkbox"/> | Identifizierungshilfe für Bildinhalt |
| 534fl | Florenz-Nr. | <input type="checkbox"/> | Stellt Bezug zur Pappelerbe her |
| 53470 | Neg.-Nr. | <input type="checkbox"/> | Stellt Bezug zum Negativ her |
| 53494 | Aufnahmedatum | <input type="checkbox"/> | Historisch interessant |
| 53540 | digitale Repro-Nr. | <input type="checkbox"/> | Dient der Kontrolle des Dateinames (identisch mit digitaler Repro-Nr.) |

Entwurf
Meta-
daten-
extraktion
aus HiDA



Langzeitarchivierung im KHI

- Rechenzentren der MPG als Archivierungspartner:
 - Institut für Plasmaphysik in Garching
 - Gesellschaft für Wissenschaftliche Datenverarbeitung in Göttingen



Langzeitarchivierung im KHI

- Migration als Archivierungstechnik
- Wird organisiert von den Archivierungspartnern
 - Garching: Bandsicherung, Bilddaten getrennt von Checksummen und Metadaten
 - Göttingen: Online-Haltung aller Dateien an zwei Standorten



Zusammenfassung

- Langzeitarchivierung als Herausforderung für Museen und Kultureinrichtungen
- Bewusstsein muss geweckt werden
- Anzustreben sind Erfahrungsaustausch und gemeinsames Vorgehen



Diskussionsvorschläge

- Notwendigkeit einer zentralen Archivierungseinrichtung?
- Möglichkeit der Archivierung durch Verteilen im Internet (Ansatz Rosenthaler, Basel)?