



Stiftung  
Preußischer Kulturbesitz



Staatsbibliothek  
zu Berlin  
Preußischer Kulturbesitz

# TIFF oder JPEG2000?!

Auf der Suche nach einem Format für die Archivierung von digitalisiertem Kulturgut

Felix Schäfer (Stiftung Preußischer Kulturbesitz), Doris Škarić (Staatsbibliothek zu Berlin)



Angesichts **steigender Speicherkosten** und **wachsendem Energieverbrauch** stellt sich die Frage nach einer Neubewertung der bisher weit verbreiteten Praxis,

**Masterdigitalisate im unkomprimierten TIFF (Baseline 6.0)**

zu speichern und zu archivieren. Es lohnt sich (erneut), einen Blick auf **alternative, effizientere, kostengünstigere Archivformate** zu werfen und eine neue Kosten-Nutzen-Abwägung zu versuchen.

1. Worin bestehen die **technischen und qualitativen Vor- und Nachteile** von TIFF und JPEG2000 für die Langzeitarchivierung von Digitalisaten aus dem Kulturerbe-Bereich?
2. **Wie verbreitet** sind von TIFF und JPEG2000 in GLAM-Einrichtungen in Deutschland, Europa und der Welt tatsächlich, sowohl in Einrichtungen, die Digitalisate erzeugen bzw. verwalten, also auch in Einrichtungen, die diese Digitalisate archivieren?
3. Wie ist **der aktuelle Diskussionsstand** darüber, wie gut oder schlecht TIFF und JPEG2000 als Dateiformate für Masterfiles der Digitalisierung und für die Archivierung geeignet sind?

- Das Master-Digitalisat
- Empfehlungen und Einschätzungen
- JPEG2000 im internationalen Kontext
- Umfrage im deutschen Kontext:
  - Motivation und Durchführung
  - Verbreitungsgrad Formate
  - Entscheidungsfaktoren
  - Gründe für und gegen JPEG2000
- Exkurse: Bit Rot und Softwarelandschaft JPEG2000
- Resümee

Zweck Master-Digitalisat: Authentizität und Vollständigkeit von abglichteten Informationen

Archivarisches Ziel: Erhaltung und Bewahrung vor Beschädigungen und Verlust

Anforderungen an Dateiformat:

- **Verlustfreiheit und Qualität**,  
d.h. vollständiger Erhalt aller signifikanten Eigenschaften (u.a. Auflösung, Farbtiefe, Bildinformationen)
- **nicht proprietäres Dateiformat** mit offenen Lizenzen, frei von Patenteinschränkungen
- ISO-normiert oder de-facto **Standard**
- **Zuverlässige Verarbeitbarkeit** durch etablierte Tools (idealerweise Open Source) zur Darstellung, Konvertierung und Validierung
- **Effiziente, hohe Speicherdichte**, ggf. mit Optionen für verlustfreie Kompression
- Langzeitstabilität und **gute Konvertierbarkeit** in neue Dateiformate
- **hohe Robustheit** gegenüber Speicherfehlern, d. h. Fehler im Bitstream hat nur minimale Auswirkungen

Zweck Master-Digitalisat:

Authentizität und Vollständigkeit von abglichteten Informationen

Archivarisches Ziel:

Erhaltung und Bewahrung vor Beschädigungen und Verlust

Anforderungen an Dateiformat:

- **Verlustfreiheit und Qualität**, d.h. vollständiger Erhalt aller signifikanten Eigenschaften (u.a. Auflösung, Farbtiefe, Bildinformationen)
- **nicht proprietäres Dateiformat** mit offenen Lizenzen, frei von Patenteinschränkungen
- ISO-normiert oder de-facto
- **Zuverlässige Verarbeitung** (Darstellung, Konvertierung)
- **Effiziente, hohe Speichereffizienz**
- Langzeitstabilität und **gute Konvertierbarkeit** in neue Dateiformate
- **hohe Robustheit** gegenüber Speicherfehlern, d. h. Fehler im Bitstream hat nur minimale Auswirkungen

Ökonomische Auswirkung und ökologische Nachhaltigkeit eines Dateiformates werden in jüngster Zeit zunehmend problematisiert und als Kriterium angemahnt

## Fazit | Juli 2021:

*“There is no single perfect format for the preservation and future use of raster images. Decisions made on file formats should be dependent on the features and functionality to be preserved and the future use cases to be supported. [...] It is recommended that careful research and analysis is carried out before migrating files to a new format.”*

<b>TIFF</b>	<ul style="list-style-type: none"><li>▪ most commonly used preservation format</li></ul>
<b>JPEG2000</b>	<ul style="list-style-type: none"><li>▪ supports both lossy &amp; lossless compression, but some compression standards are proprietary and not widely available.</li><li>▪ considered a challenger to TIFF as image preservation format when it was introduced</li><li>▪ impressive lossless compression and small file sizes, thus provides a cost saving option over TIFF</li><li>▪ IIF and mass digitization have led to JP2 being selected as an access format by several heritage organizations</li><li>▪ Format is more complex than TIFF, uses a number of proprietary compressions standards, and is supported by few affordable software options.</li><li>▪ These factors, as well as issues that can arise when converting JP2 files to other formats have slowed uptake of JP2 as a preservation format</li></ul>

Kriterium	TIFF	JPEG2000	JPG
Offenheit	3	4	4
Lizenzfreiheit	3	3	4
Verbreitung	4	3	4
Funktionalität	4	4	3
Implementierung	4	4	4
Speicherdichte	2	3	4
Verifizierbarkeit	4	4	4
Best Practice	3	3	2
Perspektive	2	4	1
Formatklasse	A	B	A
<b>SUMME</b>	<b>29</b>	<b>32</b>	<b>30</b>

## Fazit | Juni 2024:

- Obwohl nicht ursprünglich für Archivierungszwecke entwickelt, eignet sich JPEG2000 gut für die Archivierung.
- Insbesondere verlustfreie Komprimierung behebt entscheidenden Nachteil des ursprünglichen JPEG-Formats.
- JPEG2000 hat sich bereits als Bildformat bei den integrierten Bildern in PDF-Dateien durchgesetzt.
- JPEG2000 als reine Datei bietet sich als Kandidat an für ein zukünftiges Archivformat, dessen Gebrauch den abliefernden Stellen empfohlen werden kann.
- Im Unterschied zu PNG [...] eignet sich JPEG2000 speziell für Fotografien.
- **Angesichts der Diskussionen um Speicherkosten und mit wachsender Verbreitung anderer Bildformate wird TIFF [...] seinen Quasi-Monopolstatus im Archivbereich verlieren.**

## nestor Handbuch | August 2010:

*„Aus Perspektive der Langzeitarchivierung kommen also generell die Datenformate TIFF und JPEG2000 als Datenformat für das „Digital Master“ in Frage. Allerdings sind beide Formate so flexibel, dass diese Aussage spezifiziert werden muss.“*

<b>Allgemein</b>	<ul style="list-style-type: none"><li>▪ Offene, standardisierte Formate</li><li>▪ Weit verbreitet &amp; gut dokumentiert</li><li>▪ Auswahl abhängig vom Anwendungskontext</li></ul> <p>➔ TIFF = bewährter Standard   JPEG2000 = Alternative mit spezifischen Vor- und Nachteilen (nur verlustfrei)</p>
<b>JPEG2000</b>	<ul style="list-style-type: none"><li>▪ Grundsätzlich archivtauglich</li><li>▪ Vorteile bei Fehlertoleranz (Bitfehler)</li><li>▪ Nur verlustfrei empfohlen - verlustbehaftete Kompression ungeeignet</li></ul>
<b>TIFF</b>	<ul style="list-style-type: none"><li>▪ Etabliertes Archivformat (häufig verwendet)</li><li>▪ Geeignet als Digital Master</li><li>▪ Bevorzugt: unkomprimiert/verlustfrei - Achtung bei bestimmten Kompressionen (z.B. LZW)</li></ul>

## Praxisregeln Digitalisierung | Dezember 2022:

Die Speicherung der Master erfolgt in Form unkomprimierter Baseline-TIFF-Dateien. Die weitergehenden Optionen von extended TIFFs sollten für den digitalen Master nicht genutzt werden. Neben TIFF kann auch TIFF-LZW oder JPEG2000 in seiner verlustfreien Form als Format für den Bildmaster verwendet werden. Generell sind jedoch Formate ohne Komprimierung denen mit verlustfreier Komprimierung vorzuziehen. Für die Speicherung von Mastern im JPEG2000-Format ist darauf zu achten, dass nur die lizenzfreien Bereiche Verwendung finden. Zu den mit JPEG2000 und TIFF-LZW einhergehenden Risiken bezüglich der Langzeitarchivierung vergleiche Kapitel → [3.2.1.4](#).

**JPEG2000 zwar erlaubt, Formate ohne Komprimierung sind vorzuziehen**



# JPEG2000 im internationalen Kontext: Masterfiles in ausländischen Bibliotheken



Library of Congress	kurz nach 2000	JP2	<a href="https://digitalpreservation.gov/series/challenge/formats2_challenge.html">https://digitalpreservation.gov/series/challenge/formats2_challenge.html</a>
US National Archives	seit 2002	TIFF	<a href="https://www.archives.gov/records-mgmt/initiatives/scanned-textual.html">https://www.archives.gov/records-mgmt/initiatives/scanned-textual.html</a>
Königliche Bibliothek der Niederlande	seit 2007	JP2	<a href="#">Generating lossy access JP2s from lossless preservation masters</a>
Österreichische Nationalbibliothek	seit 2012	JP2	<a href="https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Digitising_Manuscripts_to_Europeana/Deliverables/D1.1%20-%20Requirements%20Report.pdf">https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Digitising_Manuscripts_to_Europeana/Deliverables/D1.1%20-%20Requirements%20Report.pdf</a>
British Library	seit ca. 2014	JP2	Newspaper, <a href="https://www.britishnewspaperarchive.co.uk/help/about">https://www.britishnewspaperarchive.co.uk/help/about</a>
Bibliothèque nationale de France	seit 2015	JP2	<a href="https://github.com/hackathonBnF/FichesFormat/wiki/JPEG2000">https://github.com/hackathonBnF/FichesFormat/wiki/JPEG2000</a>
Ghent University Library	vor 2025	JP2	<a href="#">Matthias Vandermaesen - Converting 430.000 images from JPG2000 to JPEG</a>

- **Motivation:** Einbeziehung von Erfahrungen aus der Praxis

- **Ziel der Umfrage:**



für einen Vortrag mit dem Titel „**TIFF oder JPEG2000?! – Auf der Suche nach einem Format für die Archivierung von digitalisiertem Kulturgut**“, der im März auf dem 8. Workshop zur Retrodigitalisierung „*Digitalisierung für die Ewigkeit? Datenqualität in der Praxis*“ in Berlin gehalten wird, möchten wir gerne Erfahrungen aus der Praxis einbeziehen.

Wir würden uns daher sehr freuen, wenn Sie uns Ihre Strategien, Entscheidungen und Erfahrungen zur Verwendung von **TIFF** oder **JPEG2000** als Archivierungsformat für die Masterfiles der Retrodigitalisierung mitteilen könnten, insbesondere dann, wenn Sie sich mit JPEG2000 auseinandergesetzt haben. Interessant sind für uns beispielsweise:

- Gründe für oder gegen die Wahl eines Formats und dessen Spezifikation
- institutionelle Policies, Richtlinien oder Empfehlungen
- praktische Erfahrungen in der Kuratierung und in der Langzeitarchivierung
- eventuelle Wechsel oder Neubewertungen einer Strategie

- **Zielgruppe:** Kulturerbeeinrichtungen, die digitalisieren oder archivieren
- **Methodik:** Anfrage über zwei Mailinglisten mit archivischer Expertise:  
Archivliste (Archivschule Marburg) & Langzeitarchivierungsliste (nestor)

- **einundzwanzig Rückmeldungen**, mitunter sehr ausführlich
- **nicht repräsentativ**, aber ein Schlaglicht
- von kleinen Archiven (Kreisarchiv, Stiftungsarchiv) bis zu großen Bibliotheken und Landesarchiven
- von Bibliotheken geringerer Rücklauf
  - eventuell falscher Verteiler für diese Zielgruppe
- teilweise **individuelle Einschätzungen**, teilweise **etablierte Praxis / Policy einer Institution**
- **Relevanz des Themas** wird betont, Interesse an Vortrag und Ergebnissen



# Umfrage: Ergebnisse - Formate

		JPEG	JPEG2000	TIFF
Archiv	DANRW & DiPS.kommunal c/o Landschaftsverband Westfalen-Lippe Münster		JP2 (verlustfrei)	TIFF (Baseline 6.0), inkl. Exif, XMP oder LZW-Kompression
Archiv	Landesarchiv NRW	JPEG (Qualität 100%) für „Schutzdigitalisierung“ bei Vorlagentyp „Aktenarchivgut“)	JP2 (verlustfreie) für Vorlagentyp „Urkunden, Karten, Pläne, Fotos“ [seit 2005]	
Archiv (Institutionenarchiv)	Archiv der Max-Planck-Gesellschaft			TIFF (mit LZW-Kompression)
Archiv (Institutionenarchiv)	Bundesanstalt für Geowissenschaften und Rohstoffe (BGR)			TIFF (Masterimages)
Archiv (Institutionenarchiv)	Politisches Archiv, Auswärtiges Amt, Berlin			TIFF
Archiv (Institutionenarchiv)	Universitätsarchiv UB Bonn		langfristig JP2	aktuell TIFF
Archiv (Kirchenarchiv)	Landeskirchliches Archiv, Evangelische Kirche von Westfalen, Bielefeld		JP2 (verlustfrei)	
Archiv (Kreisarchiv)	Kreisarchiv Reutlingen	JPEG (Qualität 100%) aus TIFF-Digitalisat erzeugt [seit 2017]		TIFF (bis 2016)
Archiv (Kreisarchiv)	Kreisarchiv Verden			TIFF
Archiv (Landesarchiv)	Landesarchiv Baden-Württemberg	?	?	?
Archiv (Landesarchiv)	Landesarchiv Sachsen-Anhalt		JP2 (verlustfrei) TIFF werden migriert	
Archiv (Landesarchiv)	Staatliche Archive Bayern		JP2	
Archiv (Spezialarchiv)	Archiv der sozialen Demokratie			TIFF (Baseline 6.0)
Archiv (Stadtarchiv)	Stadtarchiv Münster	JPEG (Qualität 95%) als Master für alle anderen Fälle	JP2 für Ersatzdigitalisate bei nicht erhaltungsfähigen Originalen	
Bibliothek	Bayrische Staatsbibliothek			TIFF (Baseline 6.0) mit ICC-Farbprofil
Bibliothek	Staatsbibliothek zu Berlin			TIFF (Baseline 6.0) mit ICC-Farbprofil
Bibliothek	SLUB Dresden			TIFF (Baseline 6.0) mit ICC-Farbprofil
Museum	Deutsches Museum München		JP2	
Museum	Germanisches Nationalmuseum, Historisches Archiv			TIFF (unkomprimiert)





- Richtlinien und Empfehlungen
  - kleinere Archive lehnen sich an Praxis des Bundesarchives an
  - DFG-Praxisregeln Digitalisierung relevant für Bibliotheken und DFG-geförderte Projekte
  - KOST
- Dateigröße (Speicherbedarf, Datenmanagement)
- Verbreitung der Formate
- Komplexität der Formate
- Know-How / Kapazitäten
- Abhängigkeiten von Systemen / Verbundlösungen
- Robustheit / Sicherheit
- Nicht erwähnt wurde bessere / schlechtere Einbindung von Metadaten

- **ISO-Standard** (ISO/IEC 15444), seit 2000 produktiv
- benötigt deutlich **weniger Speicherplatz** als TIFF (bis zu 50-60% weniger)
- bietet auch **verlustfreie Kompression**
- **einfacheres Datenhandling** wg. verringerte Übertragungsdauer zwischen den Systemen
- **bessere Robustheit** gegen bitrot wg. eingebauter Fehlertoleranz
- systembedingt (dips.kommunal)
- **Software-Ökosystem existiert**, proprietär und Open Source
- **qualitativ identisch** zu TIFF (Auflösung, Abmessungen, Farbtiefe, Farbraum, Farbprofil)





- **mangelnde praktische Erfahrungswerte** und auch Beispiele aus anderen Häusern mit ähnlichen Problemstellungen, an denen man sich orientieren könnte
- **geringe Verbreitung**, Nischenlösung, nicht wirklich durchgesetzt im deutschsprachigen GLAM Bereich
- **Lizenzfragen** für Part 1 nur de facto, aber nicht de jure geklärt
- **höhere Rechenzeit** zum Erzeugen von JP2 als TIFF und JPEG
- **geringe Anzahl** an unterstützender Software



*„Solide Digitale Archivsysteme überwachen permanent mit Prüfsummen die Integrität.  
Wieso sollte also Robustheit gegen Bit Rot ein Kriterium sein?“*

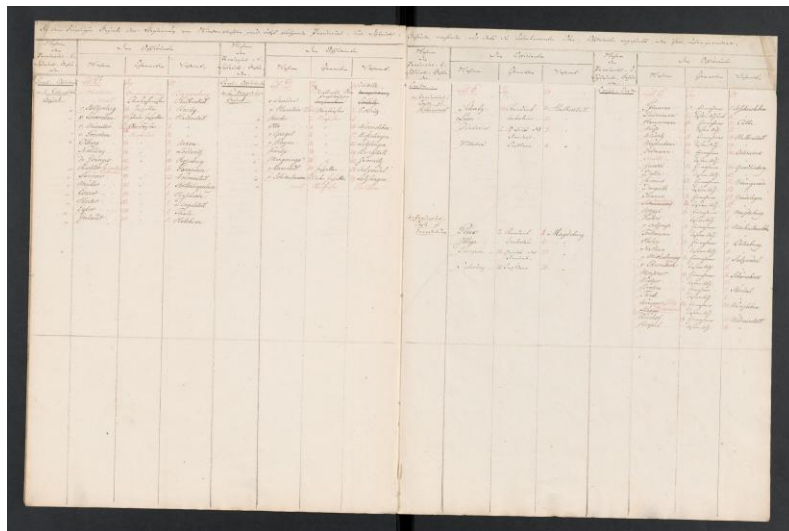
*„Bei geeigneten Speichern ist das häufig in Bezug auf JPEG beschworene Phänomen  
des Bit Rot bei uns noch nie aufgetreten“*

*„Und in dem Zusammenhang sollten auch die (geringe) Eintrittshäufigkeit möglicher und meiner Ansicht überschätzter  
Risiken wie „Bit-Rotting“ oder „Silent Corruption“, deren (überaus geringe) Eintrittswahrscheinlichkeit und deren (Nicht-  
)Auswirkungen beim Datentyp „Schutz-Digitalisat“ berücksichtigt werden.“*

*„Im Rahmen eines OAIS-konformen digitalen Archivs spielt die physische Robustheit einzelner Dateien (im Sinne von  
Fehleranfälligkeit durch Bitverlust) eine untergeordnete Rolle, da Redundanz, Checksummenverfahren und systematische  
Integritätsprüfungen vorgesehen sind.“*

*„Für die Praxis bedeutet das auch, dass JPEG 2000 im Hinblick auf die Langzeitspeicherung bzw. Erhaltung elektronischen  
Schriftguts nach OAIS ein robustes Format darstellt. [...] als auch die bessere Robustheit gegen bit rot waren für uns  
ausreichende Argumente, um seitdem JP2 als Masterformat einzusetzen.“*

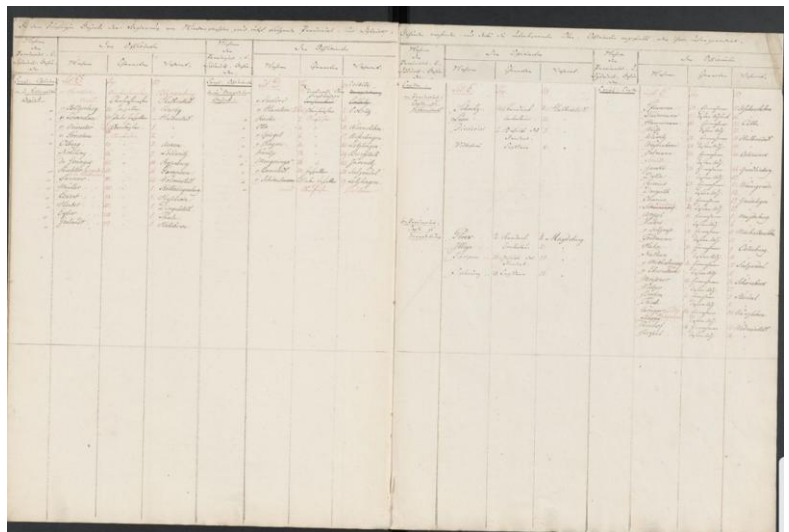
TIFF-LZW  
(Original)



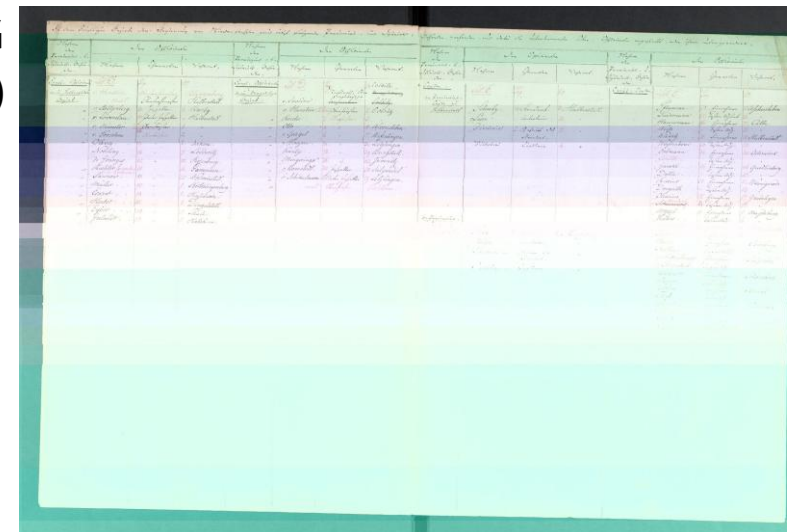
TIFF-LZW  
(Bit Rot)



JPEG2000  
(Bit Rot)



JPEG  
(Bit Rot)





- Erstellung / Konvertierung: Kakadu, OpenJPEG (mit Grok), FileStar, XnConvert, ImageMagick, libvips
- Metadaten-Extraktion: exiftool
- Validierung: jpylyzer, JHOVE/JHOVE2, KOST-Val
- Viewer: XnView MP, IrfanView (inkl. Plugins), KOST-Simy

*„Es gibt hinreichend viele und aktiv entwickelte Werkzeuge und Libraries, die insbesondere für den internen Zugriff und die Prozessierung durch Mitarbeitende und technische Systeme ausreichend sein können“*

- Pro:

- Zunehmende Verbreitung, zuletzt in Firefox und Chrome als unterstütztes Dateiformat
- Bitgenaue Rekonstruktion von klassischen, eingebetteten JPEG-Datenströmen, dies erlaubt die Nutzung auch für born-digital Objekte
- Wegfall ICC-Farbprofile durch Kodierung in „allumfassenden“ Farbraum
- Hohe Datenkompression in lossless-Variante von bis zu 30% (im Vergleich PNG: 50%)



- Contra:

- höhere Komplexität
- niedrigere Robustheit
- Entwicklung seit 2016 ohne essenziellen Durchbruch
- Prüfungen auf Tauglichkeit und Unterstützung durch breite Toollandschaft fehlen noch





- Gute **Softwareunterstützung**:
  - Genutzte Archivsoftware erlaubt und unterstützt Nutzung des Formats
  - Abhängigkeiten vermeiden: Es gibt ausreichend Viewer und andere benötigte Tools
  - Weite Akzeptanz auch in der Industrie und Software-Domäne
- **Offenheit** von Richtlinien und Quasi-Standards (z.B. DFG-Praxisregeln Digitalisierung)
- „Vorreiter“
- **Wissenstransfer und Erfahrungsaustausch**

- Relevanz von Komprimierung, Datensicherheit, Authentizität, technischer Qualität neu bewerten
  - **Wie wahrscheinlich und wie „dramatisch“ ist ein Datenverlust?**  
Was verlieren wir wirklich im Falle eines Dateifehlers oder Bit rots – insbesondere vor dem Hintergrund, wenn auch die physische Vorlagen dauerhaft erhalten wird und im Bedarfsfall erneut digitalisiert werden kann?
- Was ist **das archivarische Erhaltungsziel**: die Bilddatei als digitales Objekt oder die visuelle Bildinformation?
  - Maximaler Qualitätsanspruch an Unveränderlichkeit von Masterfiles muss man sich auch „leisten“ können
  - Es gibt keine **„absolute, kontrollierte Authentizität“** im Bezug auf die analoge Vorlage

- Aufwand der Berechnung bzw. (nachträglichen) Konvertierung von TIFF zu JPEG2000 wird als zu **ressourcenintensiv** erachtet
  - aber die langfristig höheren Speicherbedarfe und Kosten werden in Kauf genommen
  - lieber eine Archivierung mit komprimierten Dateiformaten anstatt einer Archivierung, die man sich (technisch und finanziell) nicht mehr leisten kann
  - Verlustbehaftete Bildkompression kann in Einrichtungen mit begrenzten Ressourcen vor dem Totalverlust durch «kalte Kassation» schützen
- Formatentscheidungen auch danach abwägen, ob die Konsequenzen **ökologische nachhaltig und langfristig finanzierbar** sind
  - reale Kosten eines „kontrollierten“ Qualitätsverlustes (ohne Verluste der signifikanten Eigenschaften!) und des „überschaubaren“ Rest-Risikos von Datenverfall berücksichtigen

- Es gibt international etablierte Akteure, die JP2 seit vielen Jahren erfolgreich produktiv einsetzen
  - **JP2 ist keine Nische mehr**, es gibt hinreichend langjährige Erfahrungen und Software
  - JP2 kann als archivischer Standard für Masterformate gelten, JPEG als Masterformat ist nicht zu empfehlen
  - Anpassung der DFG-Richtlinien sinnvoll
- Differenzierterer Einsatz der unterschiedlichen Datenformate anstreben
  - Format-Optionen und Varianten - ohne Kompression, mit Kompression (verlustfrei), mit Kompression (verlustbehaftet) - **ausprobieren, Erfahrungen sammeln, darüber berichten**
  - JPEG XL als mögliches (neues) Archivmasterformat im Blick haben
- Mehrwert von TIFF, der gegenüber JP2 höhere Speicher- und Energiekosten rechtfertigt, sollte aktiv begründet werden – **und wenn es keine guten Gründe gibt JPEG2000 nutzen!**

# VIELEN DANK!

- Fürs Zuhören!
- Und ein **herzliches Dankeschön** an alle, die an unserer Umfrage teilgenommen haben!

# FRAGEN?

## KONTAKT:

- Doris Škarić: [doris.skaric@sbb.spk-berlin.de](mailto:doris.skaric@sbb.spk-berlin.de)
- Felix Schäfer: [F.Schaefer@spk-berlin.de](mailto:F.Schaefer@spk-berlin.de)

