

**Analyse PDF/A-1b mit eingebetteter TrueType Schrift**  
***Fehlerhafte Einbettung der TrueType Schrift***

1	Management Summary	2
1.1	Deutsch	2
1.2	English	3
1.3	Français	4
2	Analysen	5
2.1	Erstellende Applikation	5
2.2	Druckergebnisse	5
2.3	PDF/A-1B	7
3	Fazit der Analysen	8
4	Preservation-Planning-Massnahmen	9

# 1 Management Summary

## 1.1 Deutsch

Am 20. August 2014 hat die KOST einen beunruhigenden Fehler beim Drucken eines PDF-Dokuments<sup>1</sup> mit einem PCL-Drucker festgestellt. Der Fehler besteht hauptsächlich darin, dass mehrfach ganze Passagen nicht gedruckt werden und die visuelle Reproduzierbarkeit des Dokuments somit nicht gegeben ist (siehe Abbildung 1, entsprechende Passagen mit rosa Marker hervorgehoben).

Anzeige	Druckergebnis
<p>Reinventing Archival Methods Presentation for Roundtable event in honour of Hans Hofman, National Archives of the Netherlands, The Hague, January 27 2014</p> <p>Cassie Findlay</p> <p><i>This paper has been based on one of the same name prepared and delivered at the Australian Society of Archivists' conference in 2013 with Kate Cumming, a fellow founder of the Recordkeeping Roundtable.</i></p> <p>In 1986 David Bearman first argued that the core methods of the archival profession – appraisal, description, preservation and access – were fundamentally unable to cope with the volumes of information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></p> <p>While much has been done in the intervening 25 years, as a profession, our archival methods are still today ill-equipped to deal with the volume, fragility and complexity of contemporary archival records.</p> <p>Inspired by Bearman, in November 2012 the Sydney-based discussion group, the Recordkeeping Roundtable, hosted a workshop called "Reinventing Archival Methods". At the workshop participants shared concerns that that archival professional methods are not coping with the scale and complexity of contemporary recordkeeping challenges and that they are falling at a time of critical risk.</p> <p>Participants explored how as a profession we can fundamentally reassess our methods and create a stable archival record of the 21st century. Many of the ideas discussed at the workshop have been distilled into two issues papers developed by the Recordkeeping Roundtable ('Appraisal', by Kate Cumming and Anne Picot, and 'Access', by Barbara Reed) that examine the archival methods of access and appraisal.<sup>2</sup></p> <p>Following on from that work and discussions flowing from it, today I would like to talk about some of the things that I think we as a profession should stop doing, and also what I believe we should be doing more of, to explore some strategies for responding to the extensive challenges posed by contemporary digital information and for ensuring the creation of an robust and useful archival record.</p> <p>But first – setting the scene. What is the contemporary business landscape and how is information being managed in it? How are records being made, kept and used, and are these methods compatible with the real world?</p> <p>A world characterised by:</p> <p><sup>1</sup> David Bearman, 'Archival Methods', Archives and Museum Informatics Technical Report no. 9, Pittsburgh, Archives and Museum Informatics, 1989, accessible via <a href="http://www.archimuse.com/publishing/archival_methods/">http://www.archimuse.com/publishing/archival_methods/</a></p> <p><sup>2</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: Report and what's next', December 2012, accessible via <a href="http://rroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/">http://rroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/</a></p> <p><sup>3</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: issues papers - Access and Appraisal', September 2013, accessible via <a href="http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/">http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/</a></p>	<p>Reinventing Archival Methods Presentation for Roundtable event in honour of Hans Hofman, National Archives of the Netherlands, The Hague, January 27 2014</p> <p>Cassie Findlay</p> <p><i>This paper has been based on one of the same name prepared and delivered at the Australian Society of Archivists' conference in 2013 with Kate Cumming, a fellow founder of the Recordkeeping Roundtable.</i></p> <p>In 1986 David Bearman first argued that the core methods of the archival information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></p> <p>While much has been done in the intervening 25 years, as a profession, our archival methods are still today ill-equipped to deal with the volume, fragility and complexity of contemporary archival records.</p> <p>Inspired by Bearman, in November 2012 the Sydney-based discussion group, the Recordkeeping Roundtable, hosted a workshop called "Reinventing Archival Methods". At the workshop participants shared concerns that that archival professional methods are not coping with the scale and complexity of contemporary recordkeeping challenges and that they are falling at a time of critical risk.</p> <p>Participants explored how as a profession we can fundamentally reassess our methods and create a stable archival record of the 21st century. Many of the ideas discussed at the workshop have been distilled into two issues papers developed by Cumming and Anne Picot, access and appraisal.<sup>2</sup></p> <p>Following on from that work and discussions flowing from it, today I would like to talk about some of the things that I think we as a profession should stop doing, and also what I believe we should be doing more of, to explore some strategies for responding to the extensive challenges posed by contemporary digital information and for ensuring the creation of an robust and useful archival record.</p> <p>But first – setting the scene. What is the contemporary business landscape and how is information being managed in it? How are records being made, kept and used, and are these methods compatible with the real world?</p> <p>A world characterised by:</p> <p><sup>1</sup> Archives and Museum Informatics, 1989, accessible via <a href="http://www.archimuse.com/publishing/archival_methods/">http://www.archimuse.com/publishing/archival_methods/</a></p> <p><sup>2</sup> accessible via <a href="http://rroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/">http://rroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/</a></p> <p><sup>3</sup> 2013, accessible via <a href="http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/">http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/</a></p>

Abbildung 1: links PrintScreen der ersten Seite, rechts das Druckergebnis mit den fehlenden Passagen in rosa markiert

Die Tests der KOST haben ergeben, dass einige Zeichen der TrueType Schrift "Calibri" offensichtlich fehlerhaft eingebettet wurden. Dennoch wird das Dokument von den führenden PDF/A-Validatoren als valid identifiziert.

Die vorliegende Analyse definiert Massnahmen auf verschiedenen Ebenen: Präzisierung der ISO 19005-Standards, Korrektur des Fehlers in den entsprechenden PDF/A-Konvertern sowie Erkennung des Fehlers durch PDF/A-Validatoren.

Erst wenn die PDF/A-Validatoren den Fehler erkennen, können die Archive die betroffenen PDF-Dokumente ermitteln und korrigieren.

<sup>1</sup> Das PDF-Dokument (MD5-Summe: 05ff9afaf7ded808c3200eb1beff69fc) wurde von [http://www.nationaalarchief.nl/sites/default/files/docs/nieuws/cassie\\_findlay\\_reinventing\\_archival\\_methods\\_the\\_hague\\_27jan\\_2014a.pdf](http://www.nationaalarchief.nl/sites/default/files/docs/nieuws/cassie_findlay_reinventing_archival_methods_the_hague_27jan_2014a.pdf) heruntergeladen.

## 1.2 English

On 20 August 2014, KOST identified a worrying error when printing a PDF document<sup>2</sup> using a PCL printer. The main problem is that in a number of instances, entire passages are not printed and the document's visual appearance is not preserved (see Fig. 1; the relevant passages are highlighted with pink marker).

Document as displayed	Printed document
<p>Reinventing Archival Methods Presentation for Roundtable event in honour of Hans Hofman, National Archives of the Netherlands, The Hague, January 27 2014</p> <p>Cassie Findlay</p> <p>This paper has been based on one of the same name prepared and delivered at the Australian Society of Archivists' conference in 2013 with Kate Cumming, a fellow founder of the Recordkeeping Roundtable.</p> <p>In 1986 David Bearman first argued that the core methods of the archival profession – appraisal, description, preservation and access – were fundamentally unable to cope with the volumes of information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></p> <p>While much has been done in the intervening 25 years, as a profession, our archival methods are still today ill-equipped to deal with the volume, fragility and complexity of contemporary archival records.</p> <p>Inspired by Bearman, in November 2012 the Sydney-based discussion group, the Recordkeeping Roundtable, hosted a workshop called 'Reinventing Archival Methods'.<sup>2</sup> At the workshop participants shared concerns that that archival professional methods are not coping with the scale and complexity of contemporary recordkeeping challenges and that they are falling at a time of critical risk.</p> <p>Participants explored how as a profession we can fundamentally reassess our methods and create a stable archival record of the 21st century. Many of the ideas discussed at the workshop have been distilled into two issues papers developed by the Recordkeeping Roundtable ('Appraisal', by Kate Cumming and Anne Picot, and 'Access', by Barbara Reed) that examine the archival methods of access and appraisal.<sup>3</sup></p> <p>Following on from that work and discussions flowing from it, today I would like to talk about some of the things that I think we as a profession should stop doing, and also what I believe we should be doing more of, to explore some strategies for responding to the extensive challenges posed by contemporary digital information and for ensuring the creation of an robust and useful archival record.</p> <p>But first – setting the scene. What is the contemporary business landscape and how is information being managed in it? How are records being made, kept and used, and are these methods compatible with the real world?</p> <p>A world characterised by:</p> <p>_____</p> <p><sup>1</sup> David Bearman, 'Archival Methods', Archives and Museum Informatics Technical Report no. 9, Pittsburgh, Archives and Museum Informatics, 1989, accessible via <a href="http://www.archimuse.com/publishing/archival_methods/">http://www.archimuse.com/publishing/archival_methods/</a></p> <p><sup>2</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: Report and what's next', December 2012, accessible via <a href="http://rroundtable.org/2012/12/14/reinventing-archivalmethods-report-whats-next/">http://rroundtable.org/2012/12/14/reinventing-archivalmethods-report-whats-next/</a></p> <p><sup>3</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: Issues papers - Access and Appraisal', September 2013, accessible via <a href="http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/">http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/</a></p>	<p>Reinventing Archival Methods Presentation for Roundtable event in honour of Hans Hofman, National Archives of the Netherlands, The Hague, January 27 2014</p> <p>Cassie Findlay</p> <p>This paper has been based on one of the same name prepared and delivered at the Australian Society of Archivists' conference in 2013 with Kate Cumming, a fellow founder of the Recordkeeping Roundtable.</p> <p>In 1986 David Bearman first argued that the core methods of the archival _____</p> <p>information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></p> <p>While much has been done in the intervening 25 years, as a profession, our archival methods are still today ill-equipped to deal with the volume, fragility and complexity of contemporary archival records.</p> <p>Inspired by Bearman, in November 2012 the Sydney-based discussion group, the Recordkeeping Roundtable, hosted a workshop called 'Reinventing Archival Methods'.<sup>2</sup> At the workshop participants shared concerns that that archival professional methods are not coping with the scale and complexity of contemporary recordkeeping challenges and that they are falling at a time of critical risk.</p> <p>Participants explored how as a profession we can fundamentally reassess our methods and create a stable archival record of the 21st century. Many of the ideas discussed at the workshop have been distilled into two issues papers developed by Cumming and Anne Picot, _____ access and appraisal.<sup>3</sup></p> <p>Following on from that work and discussions flowing from it, today I would like to talk about some of the things that I think we as a profession should stop doing, and also what I believe we should be doing more of, to explore some strategies for responding to the extensive challenges posed by contemporary digital information and for ensuring the creation of an robust and useful archival record.</p> <p>_____ porary business landscape and how is information being managed in it? How are records being made, kept and used, and are these methods compatible with the real world?</p> <p>A world characterised by:</p> <p>_____</p> <p><sup>1</sup> _____ Informatics Technical Report no. 9, Pittsburgh, Archives and Museum Informatics, 1989, accessible via <a href="http://www.archimuse.com/publishing/archival_methods/">http://www.archimuse.com/publishing/archival_methods/</a></p> <p><sup>2</sup> _____</p> <p><sup>3</sup> _____ accessible via <a href="http://rroundtable.org/2012/12/14/reinventing-archivalmethods-report-whats-next/">http://rroundtable.org/2012/12/14/reinventing-archivalmethods-report-whats-next/</a></p> <p><sup>4</sup> _____ 2013, accessible via <a href="http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/">http://rroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/</a></p>

Fig. 1: left: screenshot of the first page; right: the document as printed, with the missing passages highlighted in pink

Tests conducted by KOST have revealed that some characters in the TrueType font "Calibri" appear to have been incorrectly embedded. However, the leading PDF/A validators identify the document as valid.

This analysis defines measures on various levels: further clarification of the ISO 19005 standard, correction of the error in the corresponding PDF/A converters, and recognition of the error by PDF/A validators.

Archives will not be able to identify and correct the PDF documents affected until the PDF/A validators recognise the error.

<sup>2</sup> The PDF document (MD5 hash: 05ff9afaf7ded808c3200eb1beff69fc) was downloaded from [http://www.nationaalarchief.nl/sites/default/files/docs/nieuws/cassie\\_findlay\\_reinventing\\_archival\\_methods\\_the\\_hague\\_27jan\\_2014a.pdf](http://www.nationaalarchief.nl/sites/default/files/docs/nieuws/cassie_findlay_reinventing_archival_methods_the_hague_27jan_2014a.pdf).

### 1.3 Français

Le 20 août 2014, le CECO a constaté une erreur préoccupante survenue lors de l'impression d'un document PDF<sup>3</sup> avec une imprimante PCL. L'erreur réside principalement dans le fait qu'à plusieurs reprises des passages entiers n'ont pas été imprimés et que la reproductibilité visuelle du document n'a donc pas été réalisée (voir figure 1 : les passages concernés sont marqués en rose).

Affichage	Impression
<p>Reinventing Archival Methods Presentation for Roundtable event in honour of Hans Hofman, National Archives of the Netherlands, The Hague, January 27 2014</p> <p>Cassie Findlay</p> <p><i>This paper has been based on one of the same name prepared and delivered at the Australian Society of Archivists' conference in 2013 with Kate Cumming, a fellow founder of the Recordkeeping Roundtable.</i></p> <p>In 1986 David Bearman first argued that the core methods of the archival profession – appraisal, description, preservation and access – were fundamentally unable to cope with the volumes of information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></p> <p>While much has been done in the intervening 25 years, as a profession, our archival methods are still today ill-equipped to deal with the volume, fragility and complexity of contemporary archival records.</p> <p>Inspired by Bearman, in November 2012 the Sydney-based discussion group, the Recordkeeping Roundtable, hosted a workshop called 'Reinventing Archival Methods'.<sup>2</sup> At the workshop participants shared concerns that that archival professional methods are not coping with the scale and complexity of contemporary recordkeeping challenges and that they are failing at a time of critical risk.</p> <p>Participants explored how as a profession we can fundamentally reassess our methods and create a stable archival record of the 21st century. Many of the ideas discussed at the workshop have been distilled into two issues papers developed by the Recordkeeping Roundtable ('Appraisal', by Kate Cumming and Anne Picot, and 'Access', by Barbara Reed) that examine the archival methods of access and appraisal.<sup>3</sup></p> <p>Following on from that work and discussions flowing from it, today I would like to talk about some of the things that I think we as a profession should stop doing, and also what I believe we should be doing more of, to explore some strategies for responding to the extensive challenges posed by contemporary digital information and for ensuring the creation of an robust and useful archival record.</p> <p>But first – setting the scene. What is the contemporary business landscape and how is information being managed in it? How are records being made, kept and used, and are these methods compatible with the real world?</p> <p>A world characterised by:</p> <p><sup>1</sup> David Bearman, 'Archival Methods', Archives and Museum Informatics Technical Report no. 9, Pittsburgh, Archives and Museum Informatics, 1989, accessible via <a href="http://www.archimuse.com/publishing/archival_methods/">http://www.archimuse.com/publishing/archival_methods/</a></p> <p><sup>2</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: Report and what's next', December 2012, accessible via <a href="http://rkrroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/">http://rkrroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/</a></p> <p><sup>3</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: issues papers - Access and Appraisal', September 2013, accessible via <a href="http://rkrroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/">http://rkrroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/</a></p>	<p>Reinventing Archival Methods Presentation for Roundtable event in honour of Hans Hofman, National Archives of the Netherlands, The Hague, January 27 2014</p> <p>Cassie Findlay</p> <p><i>This paper has been based on one of the same name prepared and delivered at the Australian Society of Archivists' conference in 2013 with Kate Cumming, a fellow founder of the Recordkeeping Roundtable.</i></p> <p>In 1986 David Bearman first argued that the core methods of the archival information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></p> <p>While much has been done in the intervening 25 years, as a profession, our archival methods are still today ill-equipped to deal with the volume, fragility and complexity of contemporary archival records.</p> <p>Inspired by Bearman, in November 2012 the Sydney-based discussion group, the Recordkeeping Roundtable, hosted a workshop called 'Reinventing Archival Methods'.<sup>2</sup> At the workshop participants shared concerns that that archival professional methods are not coping with the scale and complexity of contemporary recordkeeping challenges and that they are failing at a time of critical risk.</p> <p>Participants explored how as a profession we can fundamentally reassess our methods and create a stable archival record of the 21st century. Many of the ideas discussed at the workshop have been distilled into two issues papers developed by Cumming and Anne Picot, and 'Access', by Barbara Reed) that examine the archival methods of access and appraisal.<sup>3</sup></p> <p>Following on from that work and discussions flowing from it, today I would like to talk about some of the things that I think we as a profession should stop doing, and also what I believe we should be doing more of, to explore some strategies for responding to the extensive challenges posed by contemporary digital information and for ensuring the creation of an robust and useful archival record.</p> <p>But first – setting the scene. What is the contemporary business landscape and how is information being managed in it? How are records being made, kept and used, and are these methods compatible with the real world?</p> <p>A world characterised by:</p> <p><sup>1</sup> David Bearman, 'Archival Methods', Archives and Museum Informatics Technical Report no. 9, Pittsburgh, Archives and Museum Informatics, 1989, accessible via <a href="http://www.archimuse.com/publishing/archival_methods/">http://www.archimuse.com/publishing/archival_methods/</a></p> <p><sup>2</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: Report and what's next', December 2012, accessible via <a href="http://rkrroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/">http://rkrroundtable.org/2012/12/14/reinventing-archival-methods-report-whats-next/</a></p> <p><sup>3</sup> Recordkeeping Roundtable, 'Reinventing Archival Methods: issues papers - Access and Appraisal', September 2013, accessible via <a href="http://rkrroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/">http://rkrroundtable.org/2013/09/28/reinventingarchival-methods-issues-papers/</a></p>

Figure 1 : à gauche capture d'écran de la première page, à droite le résultat de l'impression avec les passages manquants marqués en rose

Les tests du CECO ont révélé que quelques signes de la police d'écriture *true type* « Calibri » ont apparemment été intégrés de manière erronée. Le document a cependant été reconnu valide par les principaux validateurs PDF/A.

La présente analyse définit des mesures à différents niveaux : précision des normes ISO 19005, correction de l'erreur dans les convertisseurs PDF/A concernés ainsi que reconnaissance de l'erreur par les validateurs PDF/A.

Une fois que les validateurs PDF/A reconnaîtront l'erreur, les archives pourront identifier et corriger les documents PDF concernés.

<sup>3</sup> Le document PDF (MD5-Somme: 05ff9afaf7ded808c3200eb1beff69fc) a été téléchargé ici: [http://www.nationaalarchief.nl/sites/default/files/docs/nieuws/cassie\\_findlay\\_reinventing\\_archival\\_methods\\_the\\_hague\\_27jan\\_2014a.pdf](http://www.nationaalarchief.nl/sites/default/files/docs/nieuws/cassie_findlay_reinventing_archival_methods_the_hague_27jan_2014a.pdf)

## 2 Analysen

### 2.1 Erstellende Applikation

Die Dateieigenschaften des in Frage stehenden PDF-Dokuments zeigen auf, dass es mit "Microsoft® Word 2010" erstellt und mit "Acrobat Distiller 11.0 (Windows)" erzeugt wurde.

Beides sind aktuelle Systeme und entsprechend verbreitet im Einsatz.

### 2.2 Druckergebnisse

Wenn das PDF-Dokument mit einem Adobe-Produkt geöffnet und auf einem Drucker ohne PostScript (und ohne die erweiterte Druckoption "als Bild drucken") gedruckt wird, treten bedeutende Fehler im Erscheinungsbild auf. Dabei spielt es keine Rolle, welcher PCL-Drucker verwendet wird; der Fehler erscheint zudem bei der Konvertierung in XPS.

Zur Illustration mag der erste Absatz der ersten Seite dienen.

Ergebnis	Druckergebnis des ersten Absatzes
korrekt	In 1986 David Bearman first argued that the core methods of the archival profession – appraisal, description, preservation and access – were fundamentally unable to cope with the volumes of information that they were required to process. He called on the profession to completely reinvent its core methods. <sup>1</sup>
falsch	In 1986 David Bearman first argued that the core methods of the archival  information that they were required to process. He called on the profession to completely reinvent its core methods. <sup>1</sup>

Die folgende Tabelle fasst die Tests der KOST und die Rückschlüsse daraus zusammen. Zur Klassifizierung der Testergebnisse wird ein Farbcode verwendet (grün = korrekt, rot = falsch).

Nr	Viewer, Drucker, Ergebnis und Fazit
1	<p>Viewer: Adobe Acrobat Pro Version 10.1.10</p> <p>Drucker:           <ul style="list-style-type: none"> <li>HP Color LaserJet 4700 - HP Universal Printing PCL 6 (v5.1)</li> <li>HP Color LaserJet 4700 - HP Universal Printing PS (v5.1)</li> <li>Xerox WorkCentre 5645 - Xerox Global Print Driver PCL6</li> <li>Xerox WorkCentre 5645 - Xerox Global Print Driver PS</li> </ul> </p> <p>Fazit 1: Der PostScript-Treiber liefert korrekte, der PCL-Treiber fehlerhafte Resultate. Der Druckerhersteller kann als Ursache ausgeschlossen werden.</p>
2	<p>Viewer: Adobe Acrobat Pro Version 8.3.1</p> <p>Drucker: HP LaserJet CP1025nw color - HP LaserJet Pro CP1020 Series</p> <p>Fazit 2: Das Resultat aus Test 1 wird auf der Privat-Infrastruktur bestätigt. Die Büroinfrastruktur und die Version von Adobe Acrobat Pro können als Ursache ausgeschlossen werden.</p>

Nr	Viewer, Drucker, Ergebnis und Fazit	
3	Drucker: Viewer:	HP Color LaserJet 4700 - HP Universal Printing PCL 6 (v5.1) Adobe Acrobat Pro Version 10.1.10 Adobe Reader Version 10.1.10 Foxit Reader Version 5.4.3.0920 PDF-X-Change Viewer Version 2.5.204.0 3-Heights™ PDF Viewer Version 4.00 Fazit 3: Der Fehler taucht nur beim Gebrauch von Adobe Reader und Adobe Acrobat Pro auf.
4	Original: Gespeichert mit Viewer:	HP Color LaserJet 4700 - HP Universal Printing PCL 6 (v5.1) MD5-Summe: 05ff9afaf7ded808c3200eb1beff69fc Foxit Reader Version 5.4.3.0920: MD5-Summe: 05ff9afaf7ded808c3200eb1beff69fc PDF-X-Change Viewer Version 2.5.204.0: MD5-Summe: 8cab6cb6059eef3abd265963c7ae2d08 - Viewer: Adobe Acrobat Pro Version 10.1.10 - Drucker: HP Color LaserJet 4700 - HP Universal Printing PCL 6 3-Heights™ PDF Viewer Version 4.00: MD5-Summe: bd9ab58c7d602d50c7377b712cf8f69b - Viewer: Adobe Acrobat Pro Version 10.1.10 - Drucker: HP Color LaserJet 4700 - HP Universal Printing PCL 6 Fazit 4: PDF-X-Change Viewer und 3-Heights™ PDF Viewer verändern das PDF-Dokument beim Speichern. Dies hat jedoch kein Einfluss auf das Druckergebnis mit Adobe und PCL-Drucker.
5	Viewer: Drucker: Ausgabe: Drucker: Ausgabe:	Adobe Acrobat Pro Version 10.1.10 HP Color LaserJet 4700 - HP Universal Printing PCL 6 (v5.1) Standard Mit der erweiterten Adobe-Option "als Bild drucken" Microsoft XPS Document Writer (Office 2007) Standard Mit der erweiterten Adobe-Option "als Bild drucken" Fazit 5: Wird die erweiterte Adobe-Option "als Bild drucken" aktiviert, ist das Ergebnis korrekt. Es sind nicht nur PCL-Drucker betroffen, sondern auch die XPS-Konvertierung via XPS-Drucker.
6	Viewer: Drucker: Modifikation: Fazit 6:	Adobe Acrobat Pro Version 10.1.10 HP Color LaserJet 4700 - HP Universal Printing PCL 6 (v5.1) Eingebettete Schrift "Calibri" entfernt Das Problem liegt in der eingebetteten "Calibri" Schrift.

Nr	Viewer, Drucker, Ergebnis und Fazit
7	<p>Viewer: Adobe Acrobat Pro Version 10.1.10</p> <p>Drucker: HP Color LaserJet 4700 - HP Universal Printing PCL 6 (v5.1)</p> <p>Modifikation: In den unterdrückten Passagen der ersten Seite ist mindestens eines der folgenden Unicode-Zeichen enthalten: – [U+2013] oder ‘ [U+02BB] respektive ’ [U+02BB]. Im ersten Absatz wurde das erste – entfernt:</p> <p style="padding-left: 40px;">In 1986 David Bearman first argued that the core methods of the archival profession appraisal, description, preservation and access – were fundamentally unable to cope with the volumes of information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></p> <p>Ergebnis: <span style="border: 1px solid red; padding: 2px;">In 1986 David Bearman first argued that the core methods of the archival profession appraisal, desc information that they were required to process. He called on the profession to completely reinvent its core methods.<sup>1</sup></span></p> <p>Fazit 7: Die Unicode-Zeichen – [U+2013], ‘ [U+02BB] und ’ [U+02BC] der Schriftart Calibri sind nicht korrekt eingebettet.</p>

## 2.3 PDF/A-1B

### 2.3.1 PDF/A-1B – Ziel und Zweck

Der grundsätzliche Anspruch von PDF/A ist es, die visuelle Reproduzierbarkeit korrekt zu gewährleisten. Dieser Anspruch ist auch im dritten Absatz der *Introduction* explizit festgehalten:

The primary purpose of this part of ISO 19005 is to define a file format based on PDF, known as PDF/A, which provides a mechanism for representing electronic documents in a manner that preserves their visual appearance over time, independent of the tools and systems used for creating, storing or rendering the files.

Das vorliegende PDF/A-1b-Dokument verletzt mit der fehlerhaften Einbettung von Schriften zwar nicht ein konkretes Requirement<sup>4</sup>, aber sehr wohl dieses generelle Statement. Es wäre also zu erwarten, dass ein PDF/A-Validator darauf reagiert.

<sup>4</sup> In ISO 19005-1 steht nicht explizit, dass diese Schriften korrekt eingebettet werden müssen (6.3.2 Font types: All fonts used in a conforming file shall conform to the font specifications defined in PDF Reference 5.5.), sondern nur, dass alle verwendeten Zeichen eingebettet sein müssen (6.3.3 ff). Dass sie korrekt sein müssen, wird implizit angenommen, aber nicht ausdrücklich gesagt. In ISO 19005-2 wurde der Satz noch mit der Aussage ergänzt, dass die zitierten Spezifikationen zur PDF-Referenz konform sein müssen (6.2.11.2 Font types: All fonts and font programs used in a conforming file, regardless of rendering mode usage, shall conform to the provisions in ISO 32000-1:2008, 9.6 and 9.7, as well as to the font specifications referenced by these provisions.).

### **2.3.2 PDF/A-1B-Validierung**

Das in Frage stehende PDF/A-1b-Dokument wurde mit folgenden PDF/A-Validatoren getestet:

- Preflight in Adobe Acrobat Pro Version 10.1.10
- PDF/A-Manager Version V6.1121853 von PDFTron
- 3-Heights™ PDF Validator Version 4.3 von PDF Tools AG
- pdfaPilot Version 5.1.211 von Callas

Alle vier Validatoren identifizierten das Dokument als valides PDF/A-1b-Dokument.

## **3 Fazit der Analysen**

Wenn ein mit den aktuellen Tools erzeugtes und als valid geprüftes PDF/A-Dokument nicht korrekt druckbar ist, haben die Archive ein (noch unbekanntes, aber potentiell grosses) Problem, für das eine Lösung dringend nottut.

Für die Archive ist es inakzeptabel, dass in einem validen PDF/A ganze Textpassagen mit Adobe Acrobat Pro und Adobe Reader nicht gedruckt werden. Von einem Erhalten der visuellen Reproduzierbarkeit unabhängig der verwendeten Systeme kann nicht annähernd die Rede sein.

Die fehlerhaften Unicode-Zeichen – [U+2013], ‘ [U+02BB] und ’ [U+02BC] sind geläufige Zeichen.

Die erstellenden Applikationen (Microsoft® Word 2010 mit Acrobat Distiller 11.0) sind aktuell, und entsprechend muss davon ausgegangen werden, dass weitere PDF-Dokumente mit fehlerhafter Einbettung existieren und weiterhin produziert werden.

Entsprechend wurden folgende Preservation-Planning-Massnahmen definiert.

#### 4 Preservation-Planning-Massnahmen

Das beschriebene Problem muss auf vier Ebenen gleichzeitig<sup>5</sup> angegangen werden. Deshalb werden im Oktober 2014 folgende Massnahmen in die Wege geleitet:

- A. Die Hersteller des analysierten Dokumentes<sup>6</sup> werden über das Problem informiert und gebeten, die Reproduzierbarkeit des Fehlers abzuklären.
- B. Die Herstellerin des fehlerhaften PDF/A-Konverters wird angeschrieben und gebeten, den Fehler bei der Einbettung der TrueType-Schrift zu untersuchen und zu beheben.
- C. Die Hersteller der getesteten Validatoren werden angeschrieben und gebeten, die Prüfung der korrekten Einbettung zu erweitern, damit solche PDF/A-Dokumente nicht als valid gelten.
- D. Das Sekretariat des ISO/TC 171/SC 2/WG 5, welches für den ISO 19005 zuständig ist, wird informiert und gebeten, mit einem 'Corrigenda' oder einem anderen Mechanismus die Satzergänzung von ISO 19005-2 6.2.11.2 "as well as to the font specifications referenced by these provisions" auch in ISO 19005-1 6.3.2 nachzutragen<sup>7</sup>.

Die Massnahme C ist notwendig, damit solche PDF/A-Dokumente erkannt und korrigiert werden können. Da die herstellende Applikation eine aktuelle ist, kann mit der Massnahme B die Entstehung solcher PDF-Dokumente reduziert werden. Die Massnahme D soll dazu führen, dass die Hersteller von PDF/A-Software diesen Fehler entweder nicht produzieren oder ihre Validatoren entsprechend aufdatieren.

Erst wenn die PDF/A-Validatoren den Fehler erkennen, können die Archive die betroffenen PDF-Dokumente ermitteln und korrigieren.

---

<sup>5</sup> Alle vier Ebenen werden gleichzeitig angegangen, da unsere Erfahrung darin besteht, dass die beteiligten Ebenen sich gegenseitig abstützen, und es nicht zielführend ist, wenn nur eine Ebene angegangen wird.

<sup>6</sup> Sowohl die Autorin Cassie Findlay als auch die publizierende Institution, das Niederländische Nationalarchiv.

<sup>7</sup> Die Präzisierung in ISO 19005-2 6.2.11.2 hätte gleichzeitig mit einem Corrigenda in die Version 1 zurückfliessen sollen. Das Problem nur in den neuen Version anzugehen, greift zu kurz, da die bisherigen Versionen explizit gültig bleiben und bleiben sollen.