

EXPERT RECOMMENDATION ON NEN-ISO 19005-1:2005 (PDF/A-1)

Standardisation Forum

EXPERT RECOMMENDATION ON NEN-ISO 19005-1:2005 (PDF/A-1)

Standardisation Forum

Erika Hokke

5 June 2008

Status Final

Version

Contents

Management summary	1
1 Introduction	2
1.1 Background and history	2
1.2 Composition of the expert group	2
1.3 The standard	2
2 Recommendation of the expert group concerning the context of the standard	4
2.1 Area of application of the standard	4
2.2 Organisational scope of the standard	5
3 Evaluation of the standard based on the criteria	6
3.1 Openness	6
3.2 Usability	Fout! Bladwijzer niet gedefinieerd.
3.3 Potential	8
3.4 Impact	8
4 Considerations of the expert group regarding the standard	10
5 Expert group recommendation to the Standardisation Forum on the standard	11
6 Other comments by the expert group	12

Management summary

Expert recommendation to the Standardisation Forum on inclusion of ISO 19005-1: Document Management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1) in the list of standards of the Standardisation Board:

- The PDF/A-1 standard meets the criteria with respect to openness, usability and potential.
- Introduction of the PDF/A-1 standard will mainly impact the document creation process. Organisations will have to draw up rules for document creation in order to allow the documents to be saved in the PDF/A-1 format. Secondly, standardisation in PDF/A-1 leads to improved long-term preservation and availability of documents and simplification of conversion and migration projects. Finally, the introduction of PDF/A-1 enhances the implementation of the NEN-ISO 15489:1 and the *Besluit Kwaliteit Rijksoverheidswebsites* (National Government Websites Quality Decree).
- Based on the assessment of the openness, usability, potential, and impact, the expert group recommends inclusion of PDF/A-1 in the list of open standards.
- With regard to the organisational scope of PDF/A-1, the expert group recommends alignment with the organisational scope of ODF. Both standards focus on the same object: documents.
- The area of application of the standard includes the final versions of documents created or received by organisations within this organisational scope during the execution of government duties. Within this area of application, a distinction is made between files only realising the visual representation of the document (in accordance with PDF/A-1 b) and files also accurately representing structure and semantics (in accordance with PDF/A-1 a). This means the following:
 - In accordance with PDF/A-1 a, the following are stored: all new 'digital-born' documents created by the government in the execution of government duties.
 - In accordance with PDF/A-1 b (or PDF/A-1 a, if possible), the following are stored:
 - Analogue (paper) documents and pictures that are digitised (scanned);
 - Existing / legacy digital documents that cannot be accurately converted into PDF/A-1 a because they were created with legacy software, because they include legacy fonts, because they include mathematical equations, or for any other reason that makes conversion into PDF/A-1 a impossible.
 - The documents received by the government in the execution of government duties.
 - In all other cases, the organisation must make a choice between PDF/A-1 a and PDF/A-1 b.

1 Introduction

1.1 Background and history

On Monday, 17 September 2007, the Dutch State Secretary of Economic Affairs sent the action plan for open standards and open-source software to the Lower House. The purpose of the action plan is to make the information supply more accessible, achieve independence from IT suppliers, and enable innovation.

Part of the action plan is compiling a list of standards governed by the 'Comply or Explain' principle. The Standardisation Board decides which standards will be included in the list, based in part on an expert evaluation of the standard. The experts are part of a balanced expert group that evaluates the standard based on a number of criteria. These criteria – and their elaboration in the form of specific questions – are specified and discussed in this expert report. For additional information on the criteria used, and the follow-up steps based on this expert recommendation, please refer to the VKA report *Open standaarden. Het proces om te komen tot een lijst met open standaarden* (Open standards – the process for creating a list of open standards). This report was approved by the Standardisation Board on 14 May 2008.

1.2 Composition of the expert group

The PDF/A-1 expert group consists of the following members:

- R. de Rooij (ICTU)
- W. Kool (National Library of the Netherlands)
- G. Leffelaar (National Library of the Netherlands)
- A. Goedewagen (CJIB, Ministry of Justice)
- R. van Fucht (Elsevier)
- C. van Oosterhout (Adobe)
- R. Kerpershoek (SABIC)
- D. Roorda (DANS, KNAW)

The expert group met on 9 June 2008. This meeting was chaired by Mrs E. Hokke (VHIC) and supervised by L. Korsten and D. Jense from VKA. Before this expert meeting took place, the experts made a written assessment of the standard based on the above-mentioned criteria. In addition to the above experts, Mr K. Duijvelaar (VNG) responded in writing.

1.3 The standard

This recommendation concerns the following standard:

ISO 19005-1: Document Management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1), hereafter referred to as: PDF/A-1.

The PDF/A-1 standard is a subset of PDF (Portable Document Format). PDF was created in 1993 and has been accepted as a de facto standard for a number of years. PDF can be used as a format for the following:

1. Digital documents stored directly as PDF or converted from another format into PDF.
2. Scanned (previously analogue) documents.

PDF/A-1 is based on PDF version 1.4 and is intended for archiving documents. In this context, archiving means keeping documents accessible for an extended period of time. The visual representation of these documents remains guaranteed, regardless of the applications or systems required to manage these documents. The accurate visual representation is important to ensure the authenticity and reliability of documents.

The PDF/A-1 standard contains two application levels. Level B (PDF/A-1 b) is restricted to an accurate visual representation of the document over time. Level A (PDF/A-1 a) has more attributes, focusing on the structure and semantics of the documents. Documents that are stored in accordance with PDF/A-1 a not only look the same as the original, but can, for instance, also be searched in a structured manner.

The development of PDF/A-2, based on PDF version 1.7, has started.

2 Recommendation of the expert group concerning the context of the standard

2.1 Area of application of the standard

2.1.1 Considerations of the expert group

The expert group first assessed whether the existing PDF/A-1 standard should be examined, or whether to recommend waiting for the publication of PDF/A-2 or PDF 1.7. An argument for the latter is that PDF/A-1 entails a number of issues, which have been resolved in the new standards. However, since it is not yet known when these will be formally published by ISO, and since the current standard will be incorporated in the future version, the expert group has decided to provide a recommendation on PDF/A-1 now.

The expert group has assessed which level of PDF/A-1 qualifies for inclusion in the list: level A or level B. The expert group has not made a choice for level A or level B, but has observed that both levels of the standard have their own area of application. PDF/A-1 a is, for instance, not always suitable for scanning paper documents (in bulk). PDF/A-1 b is suitable for this purpose, as 1 b is restricted to visual representation. The current digital-born documents cannot always be stored in PDF/A-1 a. This is typically the case for legacy documents such as old digital documents that were created using old software. Documents containing old fonts cannot be stored in PDF/A-1 a.

Based on the above, the expert group concludes that both PDF/A-1 a and 1 b must be included in the organisational scope of the recommendation. The restriction to PDF/A-1 a is (technically) not feasible in all cases. By extending the recommendation to PDF/A-1 a and PDF/A-1 b, a wider spectrum of information types can be supported.

With regard to the area of application of both levels of the standard, different dimensions have been discussed. The expert group concludes that the area of application of PDF/A-1 b follows from the conclusion that, from a technical point of view, not all documents can be transferred to the higher level PDF/A-1 a. This primarily applies to analogue documents and pictures that are digitised (scanned). Second, this applies to existing digital-born documents that, for a wide variety of reasons, cannot be converted into the PDF/A-1 a format. Third, the expert group concludes that PDF/A-1 does not offer any possibilities for functionalities such as moving images, 3D and interactivity. Documents containing these functionalities are consequently excluded from the organisational scope of PDF/A-1. For now, this means a restriction in the area of application of PDF/A-1.

In addition to this 'technical' aspect, there is another issue. The obligation to use the standard can only apply to the documents created by the government, i.e. process-specific information generated during the execution of government duties. Documents received by the government organisation, for instance from citizens, may not comply with the requirements to transfer them to PDF/A-1 a. PDF/A-1 b on the other hand can be used for this purpose. Furthermore, in certain cases government bodies gather information sources in a focused manner when executing specific duties. This is, for instance, the case for the National Library of the Netherlands, a collection administration institution where information is collected, administered and made available. This type of specific collections lies outside the area of application.

This means that the area of application for PDF/A-1 a is focused on new digital-born documents created by government bodies during the execution of government duties. The expert group emphasises that these are not only documents that are exchanged within organisations or between government bodies, as PDF/A-1 also plays a role in the publication of documents via government websites.

The expert group believes that PDF/A-1 a must be used for storing final versions of digital-born documents, i.e. documents that have reached the last stage of the production process, the 'final' document. ODF is available for reviewing and exchanging documents.

2.1.2 Expert group recommendation

The expert group recommends using the PDF/A-1 standard for saving the final versions of documents. The expert group makes a distinction between documents created by the government and documents received by the government. Second, the expert group makes a distinction between originally analogue documents and digital-born documents. Within the group of digital-born documents, a distinction is made between documents that will be created after publication on the list and legacy documents. This leads to the following division within the area of application:

The expert group recommends using the PDF/A-1 a standard for:

1. The final versions of new digital-born documents created by the government during the execution of government duties. These include dynamically created documents (for instance, from a web server or a database).
This does not include documents containing moving images, 3D or interactivity, as this functionality is not yet supported by PDF/A-1.

The expert group recommends using PDF/A-1 b (or, if possible, PDF/A-1 a) for:

1. Converting analogue (paper) documents and pictures into digital documents.
2. Existing / legacy digital documents that cannot be accurately converted into PDF/A-1 a because they were created with legacy software, because they include legacy fonts, because they include mathematical equations, or for any other reason that makes conversion into PDF/A-1 a impossible.
3. The documents received by the government in the execution of government duties.

In all other cases, the organisation must make a choice between PDF/A-1 a and PDF/A-1 b.

2.2 Organisational scope of the standard

2.2.1 Considerations of the expert group

The members of the expert group see a close tie between the organisational scope for ODF and the organisational scope for PDF/A-1. Both standards are intended for documents – ODF for the creation, exchange and review of documents, PDF/A-1 for long-term storage, accessibility and unchangeability of documents. The expert group therefore proposes that the organisational scope of both standards should match each other.

2.2.2 Expert group recommendation

The expert group recommends that the organisational scope of PDF/A-1 should match the organisational scope of ODF: central government, provincial government, water boards and local government.

3 Evaluation of the standard based on the criteria

3.1 Openness

With respect to openness, the following criteria apply to the standard:

- *The standard has been approved and will be enforced by a non-profit organisation. Ongoing development is based on a decision-making procedure which is open to all stakeholders (consensus or decision by majority, etc.).*
- *The standard has been published and its specifications document is freely available or can be acquired at a nominal fee. Anyone must be able to copy and use the document and make it available for free, or at a nominal fee.*
- *The intellectual property – with respect to any patents that may exist – of (parts of) the standard is irrevocably made available on a royalty-free basis.*
- *There are no restrictions with respect to reuse of the standard.*

For specifications, the following criteria apply with respect to openness:

- *Open Specification: An open specification is a specification that is published. The document of this specification is freely available or available at a nominal fee. Anyone must be able to copy and use the document and make it available for free or at a nominal fee.*
- *Fee Specification: A free specification is an open specification that is free from any legal restrictions that impede usage and distribution. The intellectual property – with respect to any patents that may exist – of (parts of) the standard is irrevocably made available to anyone on a royalty-free basis.*

3.1.1 Considerations of the expert group

PDF/A-1 was ratified and published by the International Organization for Standardization (ISO) in June 2005. Further development and maintenance of PDF/A-1 will be realised by ISO/TC 171 in line with ISO's regular procedures.

The standard is published and can be obtained in the Netherlands at a nominal fee from the Netherlands Normalisation Institute.

The expert group addressed the question whether the original relationship with Adobe affects the openness of the standard. This is not the case. Adobe has transferred the standard to ISO. As indicated above, ISO is responsible for the further development of the standard. Since mid-2006, the PDF/A Competence Center has promoted the use and further development of the standard. In 2007, they started doing this in collaboration with AIM, the International Association for enterprise content management. Several suppliers have already developed software based on the standard.

3.1.2 Expert group recommendation

The expert group concludes that the PDF/A-1 standard meets the stated criteria with respect to openness.

3.2 Usability

To test the usability of the standard, the following questions must be addressed:

- *Is the standard sufficiently mature?*

- *How will further development and maintenance of the standard take place?*
- *How can compliance with the standard be assessed?*
- *Is there sufficient practical experience with use of the standard?*
- *Does the standard have sufficient support now, and will it also enjoy this support in the future?*
- *What are the expectations for future use of the standard?*
- *Which functional requirements does the standard have to meet within the proposed area of application?*
- *To what extent does the standard meet these requirements?*
- *How does this compare to competing standards?*

3.2.1 Considerations of the expert group

The standard will contribute to the establishment of good record management within government bodies and long-term storage and accessibility of digital documents. The standard meets the requirements of the NEN-ISO 15489-1 standard for archive and information management with regard to the creation of documents.

The expert group has also addressed the question whether the PDF/A-1 standard is sufficiently mature. Because the standard is published by ISO, the expert group concludes that the current standard is sufficiently mature. However, the expert group also wants to make a reference to the development of PDF/A-2 and PDF 1.7. In addition to the procedure for inclusion of standards in the list, the expert group suggests establishing a procedure for maintaining the list.

With respect to the usability of and the practical experience with the standard, the expert group concludes that, while PDF/A is relatively new, PDF has been in use for years. Furthermore, extensive experience has been gained with the associated standard PDF/X. Thanks to this standard, the publishing industry has been able to prevent many problems in the processing of PDF files by printers. Based on the wide use of PDF and the experiences with PDF/X, the experts expect broad support for PDF/A. Furthermore, PDF/A-1 is currently the only standard that fully complies with the web guidelines quality model. In accordance with the *Besluit Kwaliteit Rijksoverheidswebsites* (National Government Websites Quality Decree), the national government is obliged to apply this model. Including PDF/A-1 in the list of open standards will make it easier for organisations to comply with this decree. This means that synergy is possible at this point – the quality model and standard will enhance each other.

PDF/A is expected to receive sufficient support from market parties, both now and in the future. This expectation is based on the development of the PDF/A Competence Center, and the increasing number of organisations and experts becoming involved. The announced integration of PDF in Microsoft Office, and the importance of companies such as IBM and Adobe in guaranteeing long-term storage, further support the expectation that PDF/A will receive sufficient support.

In order to assess whether the standard is applied by the organisations within the standard's organisational scope, technical solutions including pre-flight tools and validation tools can be deployed. Assessment within specific business processes is also possible. This assessment method can only be applied afterwards. Although not ideal, this is considered a realistic way of measuring, as verification in advance is not always possible.

The expert group points out that the use of a PDF/A-compliant viewer is a precondition for optimal use of PDF/A-1.

Although other standards for document storage are available, these are not (yet) of the same quality as PDF/A-1. The reason can be that they are closed, relatively new, or not focused on accurate visual representation. Another

alternative is a combination of different standards, such as bitmap (jpeg2000) and XML. The disadvantage of this type of combination is that it is not available in one file. The expert group concludes that although several options are available, no other open standard can compete with PDF/A-1 when it comes to storage of digital-born documents requiring an accurate visual representation combined with indexing capabilities and a certain degree of interoperability.

3.2.2 Expert group recommendation

The expert group concludes that the PDF/A-1 standard meets the stated criteria with respect to usability.

3.3 Potential

To test the potential of the standard, the following questions must be addressed:

- *To what extent does inclusion of the standard in the list contribute to increased supplier independence?*
- *To what extent does inclusion of the standard in the list contribute to increased interoperability?*

3.3.1 Considerations of the expert group

The expert group has discussed the relationship between the PDF/A-1 standard and Adobe. Dependency between PDF and Adobe has diminished since PDF/A-1 became an ISO standard. Inclusion of PDF/A-1 in the list of standards will further increase the supplier independence. As outlined above, the number of members of the PDF/A Competence Center is growing. Several companies have developed software. There are also open-source initiatives.

Even though standards with higher interoperability are available, including XML-based standards, the experts expect inclusion of PDF/A-1 in the list of standards to enhance the interoperability. Interoperability was one of the original objectives of the development of PDF/A.

3.3.2 Expert group recommendation

The expert group concludes that the PDF/A-1 standard meets the stated criteria with respect to usability.

3.4 Impact

Outline the risks and opportunities of the standard with respect to the following areas (if applicable):

- *Continuity of the business process*
- *Financial aspects (cost and benefit)*
- *Organisational aspects*
- *Migration aspects*
- *Security aspects*
- *Privacy aspects*
- *Administrative burden*
- *Interoperability (with other processes, organisations)*
- *Supplier dependence*

3.4.1 Considerations of the expert group

After discussing the impact of mandatory use of PDF/A-1 for the long-term storage of digital-born documents, the expert group concludes that the impact will mainly involve the document creation process. Implementation of PDF/A-1 requires the author of a document to create a 'tidy' document. Documents must be well structured and metadata

must be added correctly. The findings of the expert group show that in government organisations this is not always the case. As a result, conversion of documents to PDF/A-1 will require an additional processing step. This can be viewed as a risk, but the expert group primarily focuses on the opportunities. First, there is the opportunity to further regulate the document creation process. This enhances the implementation of NEN-ISO 15489:1. Another positive effect of standardisation of PDF/A-1 is that long-term storage with guaranteed usability and global support can be ensured. Please note that the sustainability is affected by multiple factors. This is, however, not within the organisational scope of the assessment of this standard.

With respect to the financial aspects, the expert group points out that using PDF/A as a standard will most likely entail costs for new software or modifications in the existing software and business process. However, the expert group also points out that with PDF/A-1, future migration and conversion procedures can be avoided, information disclosure can be improved, and metadata management is relatively easy. Furthermore, the proliferation of document types is prevented (resulting in simplified conversion in the future), and the legacy diversity will disappear because of univocal storage.

The expert group concludes that there are almost no organisational risks or opportunities. The impact of PDF/A-1 is therefore neutral at an organisational level.

3.4.2 Expert group recommendation

The expert group concludes that the PDF/A-1 standard mainly impacts the document creation process. Organisations will have to draw up rules for document creation in order to allow the documents to be saved in the PDF/A-1 format. Secondly, standardisation in PDF/A-1 leads to improved long-term storage and availability of documents and simplification of conversion and migration projects.

4 Considerations of the expert group regarding the standard

The expert group concludes that the PDF/A-1 standard meets the criteria with respect to openness, usability, and potential. The standard impacts the document creation process and the long-term storage of documents.

The expert group further concludes that the organisational scope of PDF/A-1 should be aligned with the organisational scope of ODF. Both standards focus on the same object: documents.

The expert group defines the area of application as follows: storing final versions of documents created or received by the government in the execution of government duties. The further setup within the area of application is based on the two levels within the standard. It is not always possible to meet the more complex level of PDF/A-1 a. Furthermore, the government can only impose document creation requirements on themselves. They cannot expect citizens to submit documents that comply with these requirements. This means that documents received by the government may not be convertible into PDF/A-1 a. The expert group therefore recommends the following setup within the area of application:

PDF/A-1 b is applied for:

- a. Converting analogue (paper) documents and pictures into digital documents.
 - b. Existing / legacy digital documents that cannot be accurately converted into PDF/A-1 a because they were created with legacy software, because they include legacy fonts, because they include mathematical equations, or for any other reason that makes conversion into PDF/A-1 a impossible.
 - c. The documents received by the government in the execution of government duties.
- Organisations that can and want to store these documents in accordance with PDF/A-1 a are allowed to do so.

PDF/A-1 a is applied for the final versions of new digital-born documents created by the government in the execution of government duties. This not only applies to office documents. Dynamically created documents (for instance, from a web server or a database) are also included.

This does not include documents containing moving images, 3D or interactivity, as this functionality is not yet supported by PDF/A-1.

In all other cases, the organisation must make a choice between PDF/A-1 a and PDF/A-1 b.

The expert group does not have any additional considerations with respect to the PDF/A-1 standard.

5 Expert group recommendation to the Standardisation Forum on the standard

The PDF/A expert group recommends including PDF/A-1 in the list of recommended standards. The organisational scope of the standard matches the organisational scope of ODF – central government, provincial government, water boards and local government. The area of application of the standard includes the final versions of documents created or received by organisations within this organisational scope in the execution of government duties. Within this area of application, a distinction is made between files only realising the visual representation of the document (in accordance with PDF/A-1 b) and files also accurately representing structure and semantics (in accordance with PDF/A-1 a). This means the following:

- In accordance with PDF/A-1 a, the following is stored: all new digital-born documents created by the government in the execution of government duties.
- In accordance with PDF/A-1 b (or PDF/A-1 a, if possible), the following is stored:
 - Analogue (paper) documents and pictures that are digitised (scanned);
 - Existing / legacy digital documents that cannot be accurately converted into PDF/A-1 a because they were created with legacy software, because they include legacy fonts, because they include mathematical equations, or for any other reason that makes conversion into PDF/A-1 a impossible;
 - The documents received by the government in the execution of government duties.
- In all other cases, the organisation must make a choice between PDF/A-1 a and PDF/A-1 b.

6 Other comments by the expert group

The expert group has not addressed the question whether PDF/A-1 is only mandatory for government documents that qualify for permanent storage at a storage archive in accordance with the 1995 Archives Act and the corresponding selection instruments. The expert group refers to the new Archiving Regulations bill. This bill was not available to the members of the expert group. According to the recommendation¹ concerning this regulation, the digital archive files of the government must be saved by no later than the moment of transfer to a storage archive in accordance with a standard complying with the following requirements:

- a. The storage format is documented.
- b. It is an open standard, if possible.
- c. Compression is only allowed if it can be proved that no information loss occurs.
- d. The storage format does not include encryption.

The Archiving Regulations bill is expected to take effect in the course of 2008-2009, replacing the current 2000 Organised and accessible archive files regulation. PDF/A-1 meets the stated requirements and will therefore fit into the new regulation. It is currently unclear whether the organisational scope and area of application described in this PDF/A-1 recommendation are in line with the new archive regulation. The expert group therefore recommends asking the Nationaal Archief and/or the Dutch Ministry of Education, Culture and Science for a response during the consultation phase.

In addition to a procedure for including standards in the 'comply, explain or commit' list, the PDF/A expert group recommends developing a procedure for maintaining the list. At the moment of testing PDF/A-1, there is already limited insight into the development of PDF1.7 as ISO standard. The organisational scope and the area of application are expected to be wider than for PDF/A-1. A possible publication of the new standard by ISO raises the question of how organisations are to handle both standards.

¹ Recommendation Organised and accessible archive files regulation (Regulation 12), December 2006 via: http://www.nationaalarchief.nl/images/3_12707.pdf.