

Preservation Action Plan: Structured Data/Spreadsheets National Archives and Records Administration (NARA)

DRAFT 20190801

Template: 201907

Electronic Record or Digital Surrogate Types and Associated Formats

A spreadsheet is an electronic document in which data is arranged in grid-like rows and columns can be manipulated and be acted upon by formulae. Spreadsheet software may allow for multiple interacting sheets, AKA a workbook, and can display data as text, numerals, symbols, or in graphical form. So spreadsheet files may consist of not only the data, but also contain charts or visualizations based on the data and formulae. Each cell may contain either raw data or the results of formulas that automatically calculate and display a value based on the contents of other cells from the same or other pages/sheets, as well as external data sources.

Library of Congress Format Description for XLSX Transitional (Office Open XML), ISO 29500:2008-2016, ECMA-376, Editions 1-5
(<https://www.loc.gov/preservation/digital/formats/fdd/fdd000398.shtml>)

Library of Congress Format Description for CSV, Comma Separated Values
(<https://www.loc.gov/preservation/digital/formats/fdd/fdd000323.shtml>)

Library of Congress Format Description for OpenDocument Spreadsheet Document Format (ODS), Version 1.2 (<https://www.loc.gov/preservation/digital/formats/fdd/fdd000439.shtml>)

Essential Characteristics of this Record Type

Spreadsheets pose a challenge due to the various ways they can be used. They may contain only data and/or formulas, but they may be used to present information in a table format, combining text, numeric data, and possibly other visual cues such as color. They can also contain visual presentations created by the spreadsheet software itself such as charts, graphs and tables.

Due to the variability in use of spreadsheet applications, some of the questions to consider when preserving textual formats are relevant for spreadsheets as well:

- Would a change in the record's appearance alter its meaning?
- Does changing the record's appearance diminish its value? For example, if the records have been appraised as permanent for their informational value, and not evidential, then appearance characteristics may not need to be preserved.
- Would a change in the record's technical structure alter its appearance?
- Would a change in the record's technical structure affect its possible behaviors?

- Does deletion of a sheet/page always materially affect the content?

Built in tools such as macros with built in reports or external links are related to other record or data types which have their own essential characteristics.

Appearance

Name	Definition	Function Description
Fonts	Includes characteristics of type used in the document such as: <ul style="list-style-type: none"> • Typeface (Arial, Times New Roman, etc.) • Size (10 pt, 18 pt, etc.) • Pitch • Spacing • Emphasis (bold, italic, strikethrough, underline, etc.) 	There will always be a font in order to have text, however, font is only a core characteristic if it conveys meaning.
Color	Identification of the use of color in text and layout elements, e.g. borders, boxes. <ul style="list-style-type: none"> • Hue: color family or name • Saturation: purity or sharpness • Brightness: shade or tint • Contrast: range of optical density or tone 	Color is essential if it bears meaning and/or value.
Formatting	Includes features of the document that determine how information is presented, such as : <ul style="list-style-type: none"> • conditional formatting that provides visualization of data • color coding • formatting data in a table layout • themes or templates 	Not all formatting is essential to understanding the records. If the formatting conveys information, such as visualization of data, it may be core.
Annotations (comments)	Annotations included in the document.	

Graphics	<p>Images embedded in the document, or graphics created using a spreadsheet graphics feature.</p> <ul style="list-style-type: none"> ● WordArt ● SmartArt diagrams ● 3D shapes ● Pictures ● Shapes 	Not all graphics features are supported by all programs, but graphics often provide visual information that must be maintained as part of the record.
Pivot Tables	A pivot table is a table of statistics that summarize the data of a more extensive table. This summary might include sums, averages, or other statistics, which the pivot table groups together in a meaningful way.	Pivot tables can be used to draw attention to important information and therefore can be core.
Freeze panes	The ability to lock rows and/or columns into place so that when the page is scrolled, those rows/columns remain in place.	Indicates how the spreadsheet was viewed while in use, but not likely a core characteristic.

Structure

Name	Definition	Function Description
Schema	Record layout is typically embedded, but like databases, code lists and data dictionaries may be necessary to understand data.	
Linkage	Connection between or within records or worksheets. (See also Hyperlinks)	If connections exist, then they are core.
Character Encoding	Encoding schema, e.g., US-ASCII, EBCDIC, UTF-8.	Required for the proper parsing and rendering of the record content.
Column Count	Total number of columns with content in the document	Valuable for evaluating the completeness of the content after transformations.

Row Count	Total number of rows in the document	Valuable for evaluating the completeness of the content after transformations.
-----------	--------------------------------------	--

Behavior

Name	Definition	Function Description
Macros	A set of automated tasks that apply to content in the spreadsheet.	Generally not a core feature of a spreadsheet as macros are used to create content and appearance, but the instructions in and of themselves are ultimately not significant.
Formulas	An instruction in a cell or column that allows for automating calculations.	The final data in the cell itself is the core feature. Since information should not be changed once a spreadsheet is accessioned, maintaining formulas is likely not a core feature.
Hyperlinks	Links within the file, to external files, or to external data sources.	Hyperlinks are generally core features. The biggest risk is links to external files that may not be part of the series or to external websites that may not remain active.

Context

Name	Definition	Function Description
Related Files	A group of related or linked files that are referenced in the spreadsheet.	

Current NARA Transfer Guidance for this Record Type

Preferred:

- Comma Separated Value (CSV)
- OpenDocument Format Spreadsheet (ODS)
- ASCII Text

Acceptable:

- Microsoft Excel Office Open XML

- Microsoft Excel 97 Binary Document Format

Current NARA Public Access/Reference Format(s) for this Record Type

This Plan references existing public access file formats for electronic records at NARA, determined with a survey of the available public access formats in the National Archives Catalog. These references do not represent recommended public access formats under NARA policies. They are intended for informational purposes only.

Reference Format: In general, the records are delivered to researchers in the formats in which they are preserved.

Public Access Format: CSV, ASCII/plain text, and, where available, are downloadable in the native format. Some datasets extracted from spreadsheets are made searchable at a row level via the Access to Archival Databases (AAD) function of the National Archives Catalog.

Comments and Notes

In general, NARA accessions spreadsheets and databases in formats defined in 36 CFR 1235 and [NARA Bulletin 2014-04](#). The formats defined in those issuances drive the formats we preserve records in and provide access to them.

Microsoft has documented some of the differences between ODS and XLSX (<https://support.office.com/en-us/article/Differences-between-the-OpenDocument-Spreadsheet-ods-format-and-the-Excel-xlsx-format-4311c54f-ee86-4197-bd2d-5ecc35deb138>) as related to opening and saving ODS files in various versions of Excel.

Formats that can be opened in OpenOffice are identified here:

https://wiki.openoffice.org/wiki/Documentation/OOo3_User_Guides/Getting_Started/File_formats.

Lotus 1-2-3 Worksheet version 1.0 and 1A

NARA Format ID: NF00228

Extension(s):

- wks
- 123

Documentation

- Lotus 1-2-3 is a discontinued spreadsheet program from Lotus Software (later part of IBM). It was the spreadsheet standard throughout the 1980s and into the 1990s, part of an unofficial set of three stand-alone office automation products that included dBase and WordPerfect, to build a complete business platform.
- [ArchiveTeam entry for Lotus 1-2-3](#)
- [Wikipedia entry for Lotus 1-2-3](#)
- [IANA entry for Lotus 1-2-3](#)
- [Fileextension.org entry for WKS](#)
- [Fileextension.org entry for 123](#)

Risk and Prioritization Analysis

- Low Risk
 - Moderate Risk
 - High Risk
- 3 Numeric Risk Rating
-5 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Lotus 1-2-3 is a proprietary format but the records can be transformed into a non-proprietary format or the Open Office proprietary but open format.

Preferred Processing and Transformation Tool(s)

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can convert WK1, WKS, and 123.
- Open Source converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) retrieve cell's content with cell's format and some graphics for WK1-4 and 123.
- Lotus 1-2-3 may be required to open and convert WK2, WK3, and WK4 files.

Fileextension.org documents that you can use Lotus 1-2-3 to open the WK2 or later files. It also notes that some spreadsheets programs can still open the files, however since Microsoft Excel 2007, Excel no longer supports WK1, WK2, WK3, or WK4. Even using earlier versions of Excel may result in the loss of some of the formatting.

Preferred Viewer/Access Software

- NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Lotus 1-2-3 Worksheet 2.0

NARA Format ID: NF00229

Extension(s):

- wk1
- wk2

Documentation

- Lotus 1-2-3 is a discontinued spreadsheet program from Lotus Software (later part of IBM). It was the spreadsheet standard throughout the 1980s and into the 1990s, part of an unofficial set of three stand-alone office automation products that included dBase and WordPerfect, to build a complete business platform.
- [ArchiveTeam entry for Lotus 1-2-3](#)
- [Wikipedia entry for Lotus 1-2-3](#)
- [IANA entry for Lotus 1-2-3](#)
- [Fileextension.org entry for WK1](#)
- [Fileextension.org entry for WK2](#)

Risk and Prioritization Analysis

- Low Risk
 - Moderate Risk
 - High Risk
- 3 Numeric Risk Rating
-5 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Lotus 1-2-3 is a proprietary format but the records can be transformed into a non-proprietary format or the Open Office proprietary but open format.

Preferred Processing and Transformation Tool(s)

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can convert WK1, WKS, and 123.
- Open Source converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) retrieve cell's content with cell's format and some graphics for WK1-4 and 123.
- Lotus 1-2-3 may be required to open and convert WK2, WK3, and WK4 files.

Fileextension.org documents that you can use Lotus 1-2-3 to open the WK2 or later files. It also notes that some spreadsheets programs can still open the files, however since Microsoft Excel 2007, Excel no longer supports WK1, WK2, WK3, or WK4. Even using earlier versions of Excel may result in the loss of some of the formatting.

Preferred Viewer/Access Software

- NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Lotus 1-2-3 Worksheet 3.0

NARA Format ID: NF00230

Extension(s):

- wks
- 123

Documentation

- Lotus 1-2-3 is a discontinued spreadsheet program from Lotus Software (later part of IBM). It was the spreadsheet standard throughout the 1980s and into the 1990s, part of an unofficial set of three stand-alone office automation products that included dBase and WordPerfect, to build a complete business platform.
- [ArchiveTeam entry for Lotus 1-2-3](#)
- [Wikipedia entry for Lotus 1-2-3](#)
- [IANA entry for Lotus 1-2-3](#)
- [Fileextension.org entry for WK3](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 6 Numeric Risk Rating
- 8 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Lotus 1-2-3 is a proprietary format but the records can be transformed into a non-proprietary format or the Open Office proprietary but open format.

Preferred Processing and Transformation Tool(s)

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can convert WK1, WKS, and 123.
- Open Source converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) retrieve cell's content with cell's format and some graphics for WK1-4 and 123.
- Lotus 1-2-3 may be required to open and convert WK2, WK3, and WK4 files.

Fileextension.org documents that you can use Lotus 1-2-3 to open the WK2 or later files. It also notes that some spreadsheets programs can still open the files, however since Microsoft Excel 2007, Excel no longer supports WK1, WK2, WK3, or WK4. Even using earlier versions of Excel may result in the loss of some of the formatting.

Preferred Viewer/Access Software

- NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Lotus 1-2-3 Worksheet 4.0

NARA Format ID: NF00231

Extension(s):

- wk4

Documentation

- Lotus 1-2-3 is a discontinued spreadsheet program from Lotus Software (later part of IBM). It was the spreadsheet standard throughout the 1980s and into the 1990s, part of an unofficial set of three stand-alone office automation products that included dBase and WordPerfect, to build a complete business platform.
- [ArchiveTeam entry for Lotus 1-2-3](#)
- [Wikipedia entry for Lotus 1-2-3](#)
- [IANA entry for Lotus 1-2-3](#)
- [Fileextension.org entry for WK4](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 6 Numeric Risk Rating
- 8 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Lotus 1-2-3 is a proprietary format but the records can be transformed into a non-proprietary format or the Open Office proprietary but open format.

Preferred Processing and Transformation Tool(s)

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can convert WK1, WKS, and 123.
- Open Source converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) retrieve cell's content with cell's format and some graphics for WK1-4 and 123.
- Lotus 1-2-3 may be required to open and convert WK2, WK3, and WK4 files.

Fileextension.org documents that you can use Lotus 1-2-3 to open the WK2 or later files. It also notes that some spreadsheets programs can still open the files, however since Microsoft Excel 2007, Excel no longer supports WK1, WK2, WK3, or WK4. Even using earlier versions of Excel may result in the loss of some of the formatting.

Preferred Viewer/Access Software

- NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Lotus 1-2-3 Worksheet 5.0

NARA Format ID: NF00232

Extension(s):

- wk5

Documentation

- Lotus 1-2-3 is a discontinued spreadsheet program from Lotus Software (later part of IBM). It was the spreadsheet standard throughout the 1980s and into the 1990s, part of an unofficial set of three stand-alone office automation products that included dBase and WordPerfect, to build a complete business platform.
- [ArchiveTeam entry for Lotus 1-2-3](#)
- [Wikipedia entry for Lotus 1-2-3](#)
- [IANA entry for Lotus 1-2-3](#)
- <http://www.fileextension.org/WK5>

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 6 Numeric Risk Rating
- 8 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Lotus 1-2-3 is a proprietary format but the records can be transformed into a non-proprietary format or the Open Office proprietary but open format.

Preferred Processing and Transformation Tool(s)

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can convert WK1, WKS, and 123.
- Open Source converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) retrieve cell's content with cell's format and some graphics for WK1-4 and 123.
- Lotus 1-2-3 may be required to open and convert WK2, WK3, and WK4 files.

Fileextension.org documents that you can use Lotus 1-2-3 to open the WK2 or later files. It also notes that some spreadsheets programs can still open the files, however since Microsoft Excel 2007, Excel no longer supports WK1, WK2, WK3, or WK4. Even using earlier versions of Excel may result in the loss of some of the formatting.

Preferred Viewer/Access Software

- NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Visicalc

NARA Format ID: NF00435

Extension(s):

- VC
- VCS

Documentation

- VisiCalc (for "visible calculator") was the first spreadsheet computer program for personal computers, originally released for the Apple II by VisiCorp. When Lotus 1-2-3 was launched in 1983, VisiCalc sales ended almost overnight and the company was soon insolvent. Lotus Development purchased the company in 1985 and immediately ended sales of VisiCalc and the company's other products.
- <https://en.wikipedia.org/wiki/VisiCalc>
- <http://fileformats.archiveteam.org/wiki/VisiCalc>

Risk and Prioritization Analysis

- Low Risk
 - Moderate Risk
 - High Risk
- 15 Numeric Risk Rating
-17 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: TBD, preferably CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Visicalc has been obsolete since 1985 and ran exclusively on the TRS-80, Apple II, and Amiga platforms. It will require extensive testing to determine which tool can migrate this format.

Preferred Processing and Transformation Tool(s)

- TBD at this time.

Preferred Viewer/Access Software

Tools, in preferred order, to be used for viewing/accessing record type. Such tools could also potentially be considered for use in providing reference access to records of this type.

- NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Microsoft Excel Worksheet 2.0 and 2.1

NARA Format ID: NF00259

Extension(s):

- xls

Documentation

- Microsoft Excel is a spreadsheet tool and format developed by Microsoft for use across multiple environments. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It is the most ubiquitous format for spreadsheets, especially since the release of version 5 in 1993; it replaced Lotus 1-2-3 as the most common tool/format for spreadsheets. Excel is part of the Microsoft Office suite.
- [OpenOffice.org's documentation of the Excel File Format](#)
[Wikipedia entry for Excel](#)
[ArchivesTeam entry for XLS](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 3 Numeric Risk Rating
- 1 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: The formats are no longer supported in current versions of the software.

Preferred Processing and Transformation Tool(s)

- Excel 2106 can open early Excel versions.

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

So far no current software has been identified that can open Excel 2.0 or 3.0 files, so an older version of Excel may be required to access these files and migrate them to either ODS or XLSX.

Preferred Viewer/Access Software

- None in the marketplace for Excel 2 or 3.
- Current Excel software can open early Excel versions.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

Microsoft Excel Worksheet 2.0 and 2.1

NARA Format ID: NF00259

Extension(s):

- xls

Documentation

- Microsoft Excel is a spreadsheet tool and format developed by Microsoft for use across multiple environments. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It is the most ubiquitous format for spreadsheets, especially since the release of version 5 in 1993; it replaced Lotus 1-2-3 as the most common tool/format for spreadsheets. Excel is part of the Microsoft Office suite.
- [OpenOffice.org's documentation of the Excel File Format](#)
[Wikipedia entry for Excel](#)
[ArchivesTeam entry for XLS](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 3 Numeric Risk Rating
- 1 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: The formats are no longer supported in current versions of the software.

Preferred Processing and Transformation Tool(s)

- Excel 2106 can open early Excel versions.

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

So far no current software has been identified that can open Excel 2.0 or 3.0 files, so an older version of Excel may be required to access these files and migrate them to either ODS or XLSX.

Preferred Viewer/Access Software

- None in the marketplace for Excel 2 or 3.
- Current Excel software can open early Excel versions.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

Microsoft Excel Worksheet 3.0

NARA Format ID: NF00261

Extension(s):

- xls

Documentation

- Microsoft Excel is a spreadsheet tool and format developed by Microsoft for use across multiple environments. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It is the most ubiquitous format for spreadsheets, especially since the release of version 5 in 1993; it replaced Lotus 1-2-3 as the most common tool/format for spreadsheets. Excel is part of the Microsoft Office suite.
- [OpenOffice.org's documentation of the Excel File Format](#)
[Wikipedia entry for Excel](#)
[ArchivesTeam entry for XLS](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 3 Numeric Risk Rating
- 1 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: The formats are no longer supported in current versions of the software.

Preferred Processing and Transformation Tool(s)

- Excel 2106 can open early Excel versions.

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

So far no current software has been identified that can open Excel 2.0 or 3.0 files, so an older version of Excel may be required to access these files and migrate them to either ODS or XLSX.

Preferred Viewer/Access Software

- None in the marketplace for Excel 2 or 3.
- Current Excel software can open early Excel versions.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

Microsoft Excel Worksheet 4.0

NARA Format ID: NF00262

Extension(s):

- xls

Documentation

- Microsoft Excel is a spreadsheet tool and format developed by Microsoft for use across multiple environments. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It is the most ubiquitous format for spreadsheets, especially since the release of version 5 in 1993; it replaced Lotus 1-2-3 as the most common tool/format for spreadsheets. Excel is part of the Microsoft Office suite.
- [OpenOffice.org's documentation of the Excel File Format](#)
[Wikipedia entry for Excel](#)
[ArchivesTeam entry for XLS](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 8 Numeric Risk Rating
- 6 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: The formats are no longer supported in current versions of the software.

Preferred Processing and Transformation Tool(s)

- Excel 2106 can open early Excel versions.

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

So far no current software has been identified that can open Excel 2.0 or 3.0 files, so an older version of Excel may be required to access these files and migrate them to either ODS or XLSX.

Preferred Viewer/Access Software

- None in the marketplace for Excel 2 or 3.
- Current Excel software can open early Excel versions.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

Microsoft Excel Worksheet 97

NARA Format ID: NF00263

Extension(s):

- xls
- xlt
- xlm

Documentation

- Microsoft Excel is a spreadsheet tool and format developed by Microsoft for use across multiple environments. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It is the most ubiquitous format for spreadsheets, especially since the release of version 5 in 1993; it replaced Lotus 1-2-3 as the most common tool/format for spreadsheets. Excel is part of the Microsoft Office suite.
- [OpenOffice.org's documentation of the Excel File Format](#)
[Wikipedia entry for Excel](#)
[ArchivesTeam entry for XLS](#)

Risk and Prioritization Analysis

- Low Risk
 - Moderate Risk
 - High Risk
- 17 Numeric Risk Rating
17 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format:
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Excel 97 is an Acceptable format as per NARA Transfer Guidance.

Preferred Processing and Transformation Tool(s)

- Excel 2016 can open Excel 97.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

So far no current software has been identified that can open Excel 2.0 or 3.0 files, so an older version of Excel may be required to access these files and migrate them to either ODS or XLSX.

Preferred Viewer/Access Software

- None in the marketplace for Excel 97.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

Microsoft Excel Worksheet 2000-2003

NARA Format ID: NF00260

Extension(s):

- xls

Documentation

- Microsoft Excel is a spreadsheet tool and format developed by Microsoft for use across multiple environments. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It is the most ubiquitous format for spreadsheets, especially since the release of version 5 in 1993; it replaced Lotus 1-2-3 as the most common tool/format for spreadsheets. Excel is part of the Microsoft Office suite.
- [OpenOffice.org's documentation of the Excel File Format](#)
- [Wikipedia entry for Excel](#)
- [ArchivesTeam entry for XLS](#)

Risk and Prioritization Analysis

Low Risk

Moderate Risk

High Risk

23 Numeric Risk Rating

21 Numeric Prioritization Rating

Proposed Preservation Plan

Retain file format in its existing format.

Transform file to a new format.

Selected Format: XLSX

Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.

Procure/develop tools to transform the format to the preferred normalized form.

Provide Additional Information so that the record type remains understandable/usable over time.

Explore Additional Options

Justification: Excel 2000-2003 is an Acceptable format as per NARA Transfer Guidance.

Preferred Processing and Transformation Tool(s)

- Excel 2016 can open Excel 97.

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

So far no current software has been identified that can open Excel 2.0 or 3.0 files, so an older version of Excel may be required to access these files and migrate them to either ODS or XLSX.

Preferred Viewer/Access Software

- None in the marketplace for Excel 97.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open early Excel files.

Microsoft Excel for Macintosh (3.0, 4.0, 98, 2001, 2002, 2004, v.X) (xls)

NARA Format ID:

- NF00265 (2001)
- NF00266 (2002)
- NF00267 (2004)
- NF00268 (3)
- NF00269 (4)
- NF00270 (98)
- NF00271 (X)

Extension(s):

- xls

Documentation

- Microsoft Excel is a spreadsheet tool and format developed by Microsoft for use across multiple environments. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It is the most ubiquitous format for spreadsheets, especially since the release of version 5 in 1993; it replaced Lotus 1-2-3 as the most common tool/format for spreadsheets. Excel is part of the Microsoft Office suite.
- [Wikipedia entry for Excel](#)
- [ArchivesTeam entry for XLS](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 1 Numeric Risk Rating
- 1 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Some versions are no longer supported by current software. Transforming allows for maintaining the files in a more sustainable format.

Preferred Processing and Transformation Tool(s)

- So far no current software has been identified that can open Excel 3.0 files, so an older version of Excel may be required to access these files and migrate them to either ODS or XLSX.
- Current Excel software should be able to open Excel 4.0 and later files.
- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open Excel 4.0 and later files.

Preferred Viewer/Access Software

- None in the marketplace for Excel 2 or 3.
- Current Excel software can open Excel 4.0 files but nothing earlier.
- Universal Office Converter - Commandline library that converts between any document format which supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>) can open Excel 4.0 files.

Microsoft Excel Office Open XML

NARA Format ID: NF00272

Extension(s):

- xlsx

Documentation

- Open Office XML is an open, XML-based container format developed by Microsoft for representing spreadsheets, charts, presentations and word processing documents. As of Excel 2007, the xlsx format replaced the binary .xls format, although it does not support Excel macros for security reasons.
- [Format specification](#)
- [OpenOffice.org's documentation of the Excel File Format](#)
- [ArchivesTeam entry for Excel](#)

Risk and Prioritization Analysis

- Low Risk**
 - Moderate Risk**
 - High Risk**
- 29 **Numeric Risk Rating**
29 **Numeric Prioritization Rating**

Proposed Preservation Plan

- Retain** file format in its existing format.
- Transform** file to a new format.
Selected Format:
- Procure/develop tools** to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools** to transform the format to the preferred normalized form.
- Provide Additional Information** so that the record type remains understandable/usable over time.
- Explore Additional Options**

Justification: The format is considered “acceptable” per NARA transfer guidance.

Preferred Processing and Transformation Tool(s)

- The format can be opened in current versions of Excel or OpenOffice.

Preferred Viewer/Access Software

- None - NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Microsoft Excel Backup

NARA Format ID: NF00264

Extension(s):

- xlk

Documentation

- XLK files were used by Microsoft Excel 2000 and earlier version for backups of XLS workbook files. Current versions of Microsoft Excel can still open XLK files.
- [Wikipedia entry for Excel](#)
[Fileextension.org entry for XLK](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 4 Numeric Risk Rating
- 6 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: The file format is a backup version of XLS files. If it is determined to be a unique file, it should be converted to the more standard XLS format.

Preferred Processing and Transformation Tool(s)

- Microsoft Excel

Preferred Viewer/Access Software

- None - NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Microsoft Excel Workspace

NARA Format ID: NF00273

Extension(s)

- xlw

Documentation

- The XLW files contains layout information for open workbooks and their exact screen position and window sizes, as well as references to the spreadsheet workbook files. XLW files do not include the actual workbook data, but only document the layout on the screen. In order to open the XLW files you must have all the workbooks referenced by the file. It is not clear if these files in any way enhance the understanding of spreadsheet records in their native format.
- [Fileextension.org entry of XLW](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 2 Numeric Risk Rating
- 4 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: TBD
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: This file type can no longer be created, and cannot be converted to any other file format by Excel. Case-by-case assessment is needed to determine if these files should be retained as Record.

Preferred Processing and Transformation Tool(s)

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>). XLW files can be opened in LibreOffice.

Preferred Viewer/Access Software

- None - NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Quattro Pro Spreadsheet for Windows

NARA Format ID: NF00391

Extension(s):

- wb1
- wb2
- wb3
- qpw

Documentation

- Quattro Pro is a spreadsheet program developed by Borland and now sold by Corel, most often as part of Corel's WordPerfect Office suite.
- [ArchivesTeam entry](#)
- [Wikipedia entry for Quattro Pro](#)
- [Fileextension.org entry for WB1](#)
- [Fileextension.org entry for WB2](#)
- [Fileextension.org entry for WB3](#)
- [Fileextension.org entry for QPW](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 1 Numeric Risk Rating
- 3 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Quattro Pro is still sold as part of the Word Perfect Suite, but it is no longer in common use; the files can be transformed to a more sustainable format.

Preferred Processing and Transformation Tool(s)

- Universal Office Converter - Commandline library that converts between any document format which is supported by LibreOffice/OpenOffice (<https://github.com/dagwieers/unoconv>). WB2 files are supported by LibreOffice.
- A converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) can retrieve a cell's content with cell's format and some graphics for WB1-3 files and QPW v. 9 files.

Preferred Viewer/Access Software

- None - NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Quattro Pro Spreadsheet for DOS

NARA Format ID: NF00390

Extension(s):

- wq1
- wq2

Documentation

- Quattro Pro is a spreadsheet program developed by Borland and now sold by Corel, most often as part of Corel's WordPerfect Office suite.
- [ArchivesTeam entry](#)
- [Wikipedia entry for Quattro Pro](#)
- [Fileextension.org entry for WQ1](#)
- [Fileextension.org entry for WQ2](#)

Risk and Prioritization Analysis

- Low Risk
- Moderate Risk
- High Risk
- 2 Numeric Risk Rating
- 4 Numeric Prioritization Rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: CSV or XLSX
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Quattro Pro for DOS is no longer sold or supported. Testing is needed to confirm if current Quattro Pro for Windows versions can open these files for transformation.

Preferred Processing and Transformation Tool(s)

- LibreOffice Calc appears to be able to open WQ1 and WQ2 files, in which case the files could then be saved to ODS (with unknown risk), but that needs to be tested.
- A converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) can retrieve a cell's content with cell's format and some graphics for WQ1 and WQ2.
- Quattro Pro software may be required to open WQ1 and WQ2 files, which can then possibly be converted to XLS and then ODS, but it's unclear if anything will be lost in the process.

Preferred Viewer/Access Software

- None - NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Microsoft Multiplan 4.0

NARA File Format ID: NF00278

Extension(s):

- mp

Documentation

- Multiplan was a spreadsheet program released by Microsoft in 1982 on several mainframe and early desktop computer platforms including MS-DOS and Apple Macintosh.
- [Fileextension.org entry for MP](#)
- [Wikipedia entry Multiplan](#)

Risk and Prioritization Analysis

Supply the Risk Level and Numeric Rating and the Prioritization Numeric Rating as generated with the Format Risk and Prioritization Matrix for this file format.

- Low Risk
 - Moderate Risk
 - High Risk
- 3 Numeric Risk rating
-11 Numeric Prioritization rating

Proposed Preservation Plan

- Retain file format in its existing format.
- Transform file to a new format.
Selected Format: TBD
- Procure/develop tools to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools to transform the format to the preferred normalized form.
- Provide Additional Information so that the record type remains understandable/usable over time.
- Explore Additional Options

Justification: Additional research is needed to determine if there are any tools in the marketplace that can open these files.

Preferred Processing and Transformation Tool(s)

The MP format contains data used by the application and cannot be opened except by the Multiplan software, nor can it be converted to another format. Multiplan files could perhaps be converted by Excel 5, but this is unverified. It is possible that the MP format is a software file and not the actual spreadsheet file. Further investigation is required.

- A converter based on libwps (<http://libwps.sourceforge.net/convertWPS.html>) can retrieve cell contents, but it's unclear if this format is actually data.

Preferred Viewer/Access Software

- None - NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).

Open Document Format Spreadsheet version 1.0, 1.1, 1.2

NARA Format ID: NF00349

Extension(s):

- fods
- ods

Documentation

- The Open Document Format for Office Applications (ODF), also known as OpenDocument, is an open ZIP-compressed XML-based file format specification for spreadsheets and other office productivity files.
- <https://en.wikipedia.org/wiki/OpenDocument>
- http://fileformats.archiveteam.org/wiki/Open_Document_Spreadsheet
- <https://www.oasis-open.org/standards#opendocumentv1.2>

Risk and Prioritization Analysis

- Low Risk**
- Moderate Risk**
- High Risk**
- 45 Numeric Risk Rating**
- 45 Numeric Prioritization Rating**

Proposed Preservation Plan

- Retain** file format in its existing format.
- Transform** file to a new format.
Selected Format:
- Procure/develop tools** to preserve, manage and provide access to records of this type in their existing form.
- Procure/develop tools** to transform the format to the preferred normalized form.
- Provide Additional Information** so that the record type remains understandable/usable over time.
- Explore Additional Options**

Justification: The format is considered Preferred as per NARA transfer guidance.

Preferred Processing and Transformation Tool(s)

- Retain. The format can be opened in current versions of Excel or OpenOffice.

Preferred Viewer/Access Software

- None - NARA does not currently provide viewer/access software. Some data is provided via Access to Archival Databases (AAD).
- The format can be opened in current versions of Excel or OpenOffice.