

Alphabet of open-source licenses



- LGPL – GNU Lesser General Public License. Allows commercial use. Changes must be provided though.
- GPL – GNU General Public License. Incorporating it into proprietary code will require that code to be GPL licensed.
- AGPL – Affero GPL. Addresses use of GPL code on servers.

Alphabet of open-source licenses



- Apache – Permissive license. Requires list of modifications.
- MIT – Just need to include original license and copyright.
- BSD – Berkeley source distribution. Similar to Apache but not so clear on patents.
- MPL – Mozilla public license. Permissive. Changes to code must be made available.

Open-source PDF rendering (plus more) solutions



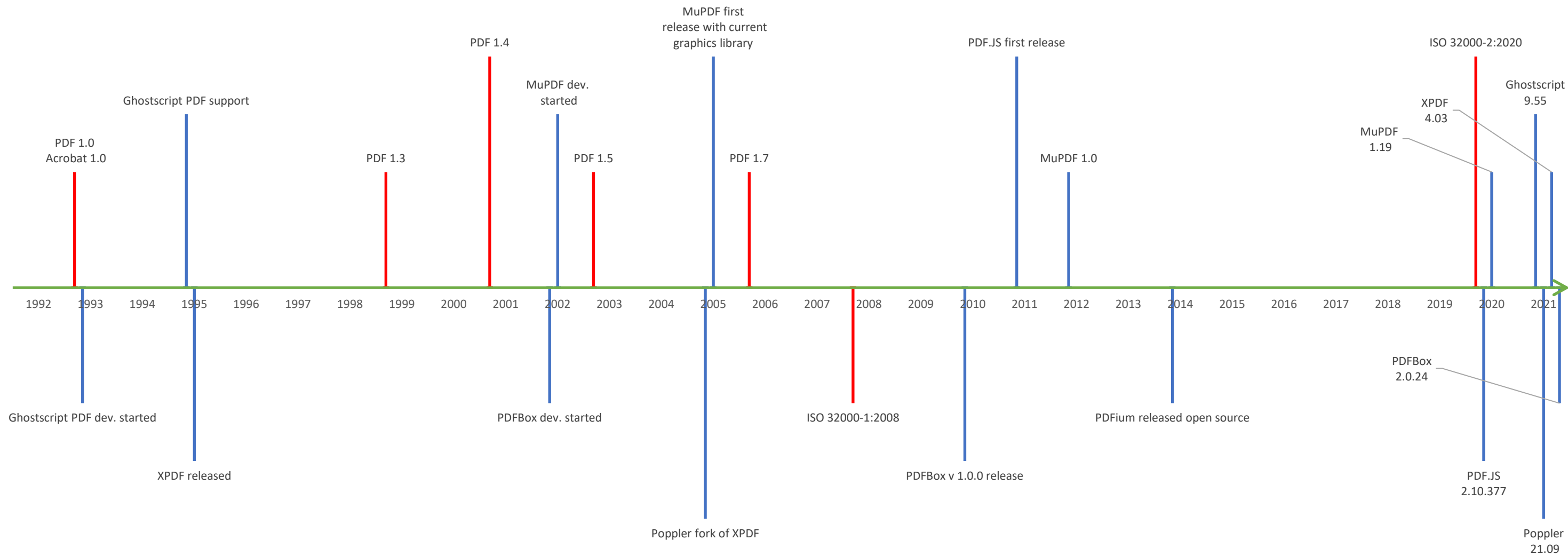
- Ghostscript (AGPL v3)
- XPDF (GPL)
- Apache PDFBox (Apache License 2.0)
- MuPDF (AGPL v3)
- Poppler (GPL v3)
- PDF.JS (Apache License 2.0)
- PDFium (Apache License 2.0)

Open-source PDF creation/manipulation solutions



- iText Java and .NET PDF generation/manipulation solution (AGPL)
- OpenPDF is a fork of Java iText (LGPL v3 and MPL v2)
- QPDF allows limited PDF manipulation C++ (Apache License, Version 2.0)
- PoDoFo C++ library. Parse and modify existing PDF files create new ones (LGPL v3)
- TCPDF PHP class for generating PDF documents (LGPL v3)

Timeline of open-source rendering and PDF



Ghostscript



- Written in C
- First released in 1986 as PostScript interpreter, rendering engine
- PDF interpreter work began in 1993, released in Jan 1995 Ver. 3.24
- Ver. 9.55 September 2021
- <https://ghostscript.com>
- GNU Affero GPL v3 / Commercial from Artifex Software
- Supports overprint
- Renders to Gray, RGB, CMYK, DeviceN, RGBA
- Does not handle interactive features
- Has anti-aliasing control
- Provides high level conversion between PDL languages (XPS, PostScript, PCL)
- API in C, Python, Java, and C# (Mono and .NET)



- First released in Dec. 1995
- Written in C++
- <https://www.xpdfreader.com/download.html>
- GPL v2 and GPL v3 / Commercial from Glyph & Cog
- Renders to RGB or Gray
- Version 4.03 in Jan 2021
- Does not support overprint or interactive features
- Has anti-aliasing control
- Python third-party bindings



- Written in Java
- Development started in 2002
- Release 1.0.0 Feb. 2010
- Ver. 2.0.24 June 2021
- Apache License 2.0
- <https://pdfbox.apache.org/>
- Renders to Gray, RGB, or ARGB
- Does not support overprint
- Cannot turn off anti-aliasing
- Exports and imports AcroForm data



- Written in C
- Work started in 2002, ver. 1.0 August 2012
- Ver. 1.19.0 September 2021
- <https://mupdf.com>
- GNU Affero GPL v3 / Commercial from Artifex Software
- Supports overprint, interactive features, and AcroForms
- Renders to Gray, RGB, CMYK, RGB+Spots, CMYK+Spots, all with or without alpha
- Has anti-aliasing control
- API in C, Python, Java, C++, WASM, plus Appkits for iOS and Android
- Includes JavaScript open-source interpreter that provides script interface



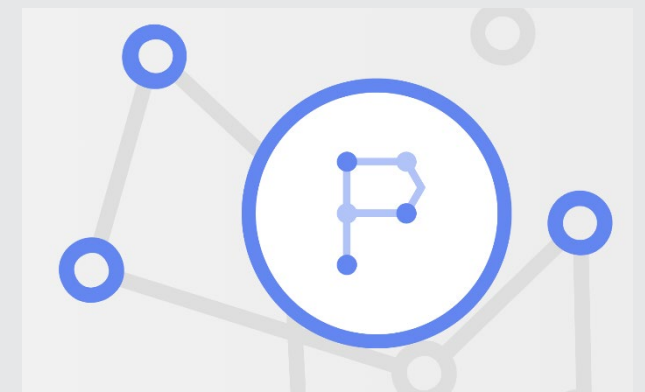
- Fork of XPDF in 2005 (XPDF 3.0)
- Written in C++
- Release 21.09.0 September 2021
- <https://poppler.freedesktop.org/>
- GPL v2 or v3
- Renders to Gray, RGB, or CMYK
- Can convert to PostScript
- Has anti-aliasing control
- Supports AcroForms, annotations, and overprint
- Third party Python bindings

The word 'Poppler' in a blue serif font, with a horizontal line passing through the middle of the letters, and a faint reflection of the word below it.

- Written in JavaScript
- Initial release July 2011
- <https://mozilla.github.io/pdf.js/>
- Apache License 2.0
- Ver. 2.10.377 July 2021
- Renders PDF onto web standards-compliant HTML5 Canvas
- Does not support overprint, outputs RGB
- Cannot turn off anti-aliasing
- Supports annotations, AcroForms



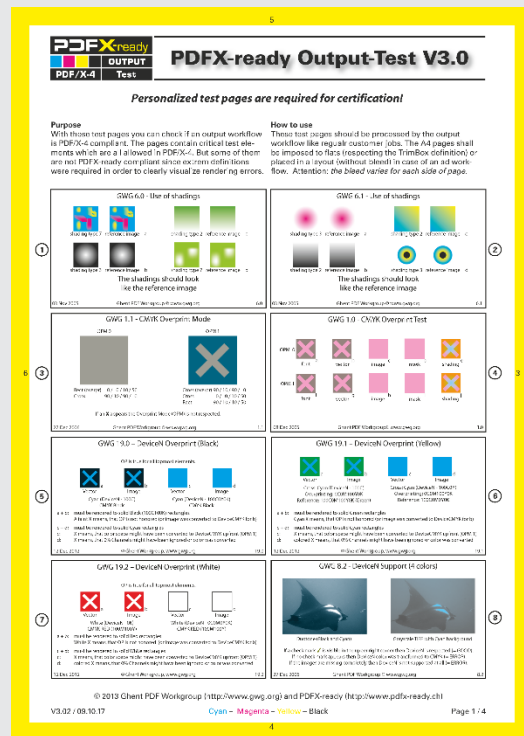
- 2014 Fork from Foxit PDF SDK
- A Google open-source project
- Apache License 2.0
- <https://opensource.google/projects/pdfium>
- Do not do formal releases
- Written in C++
- Renders to Gray, RGB, and RGBA
- Has anti-aliasing control
- Does conversions to PostScript
- Does not support overprint
- Supports AcroForms, XFA, interactive JavaScript via V8
- Several third-party C# (.NET) solutions
- Python third-party solutions
- Github has projects that build for iOS, Android, macOS, WASM
- PSPDFKit uses PDFium as the rendering engine for their products



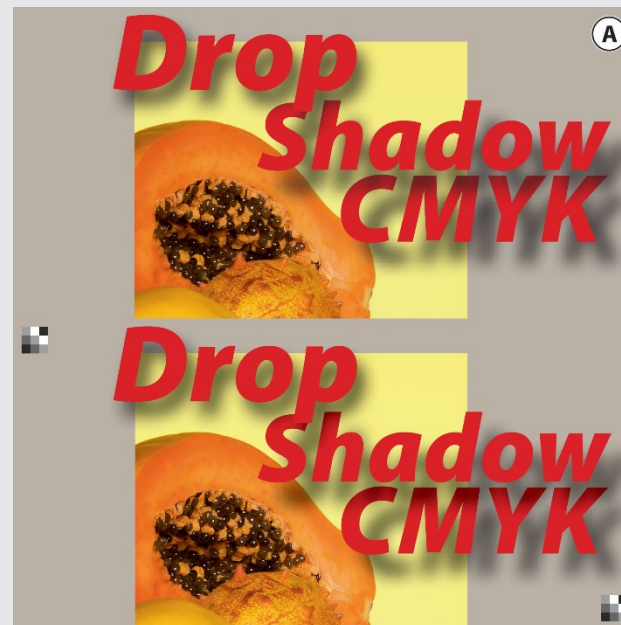
- Projects built on Windows
 - PDFBox Java binary used pdfbox-app-2.0.24.jar
 - PDF.JS HTML/JavaScript PDF to PNG convertor written, run in Firefox
- Four core Intel i7 960 @ 3.20 GHz 12GB RAM
- Clocked down to 1.6 GHz to avoid thermal throttle
- Each project evaluated on
 - Rendering accuracy
 - Processing time
 - Memory usage

Testing

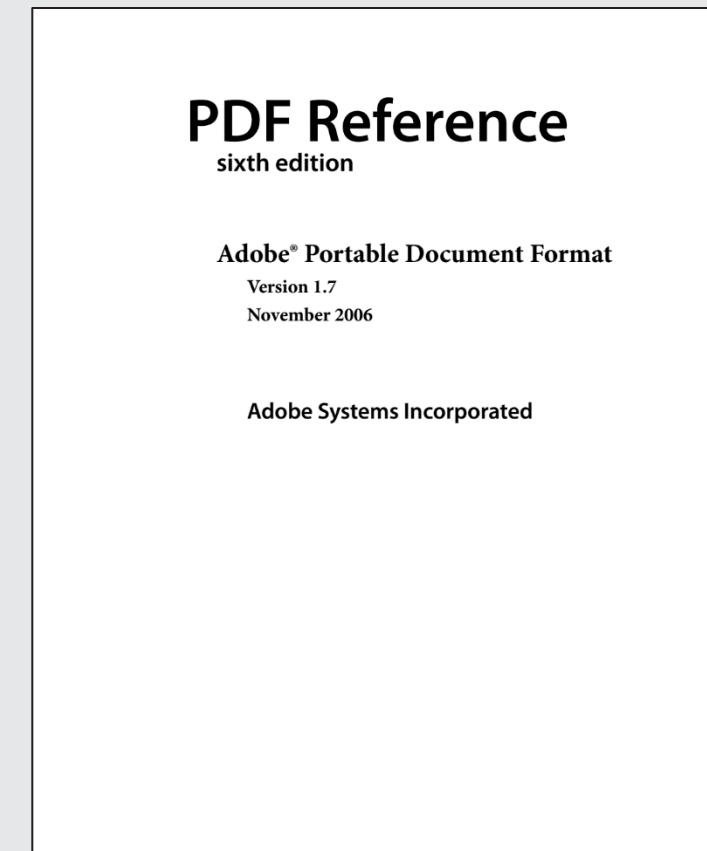
- Three files – Ghent, Altona, PDF 1.7 Specification



Four pages
Rendered at 600dpi

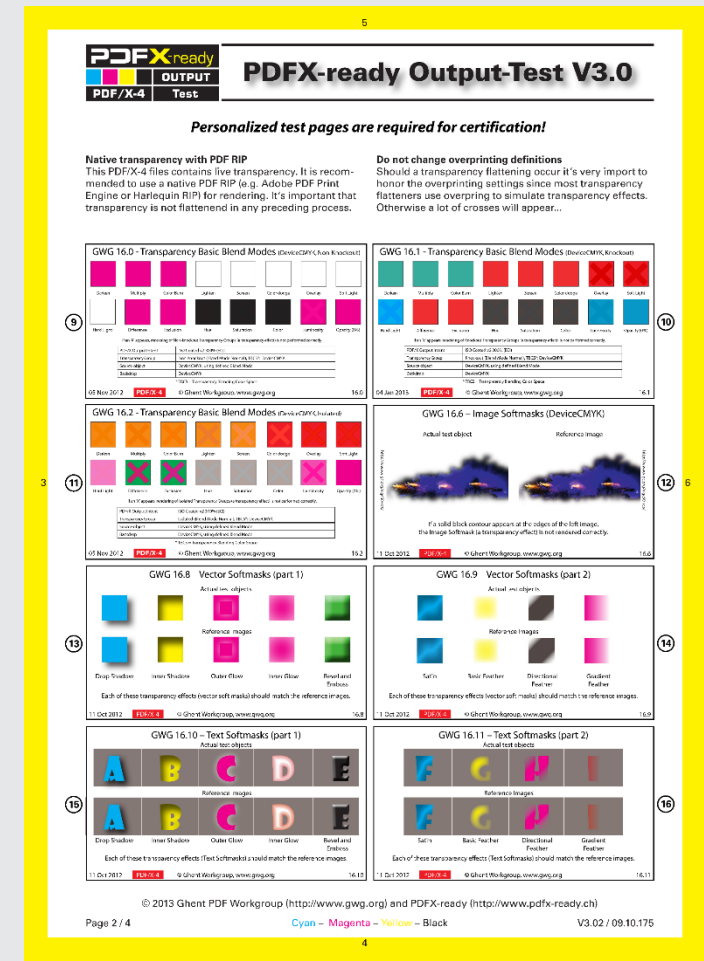
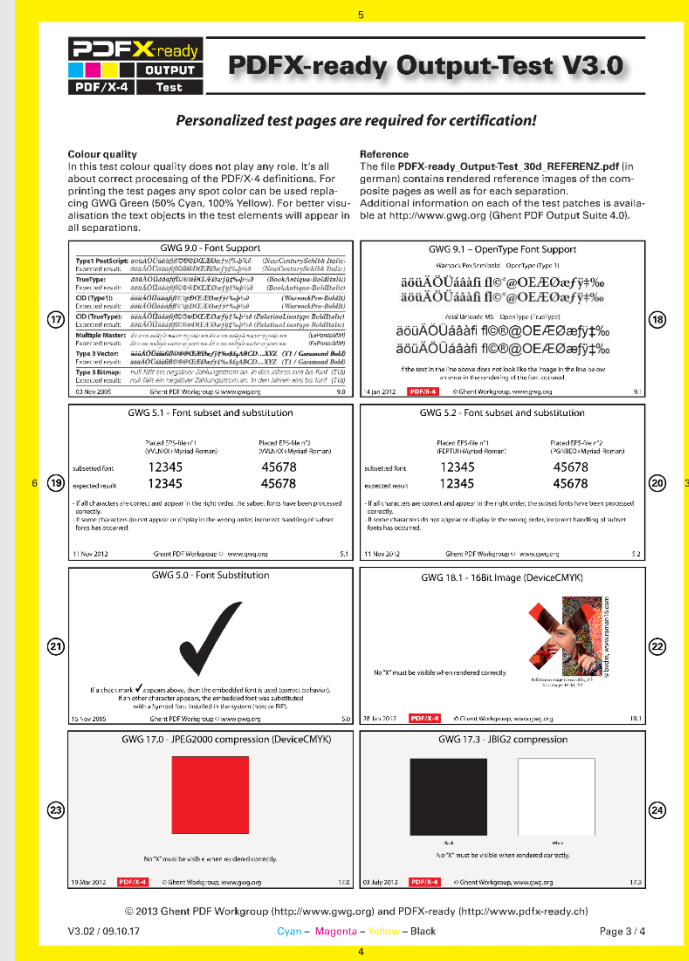


Seventeen pages
Rendered at 600dpi



1310 pages
Rendered at 150dpi

Sampling of rendering issues

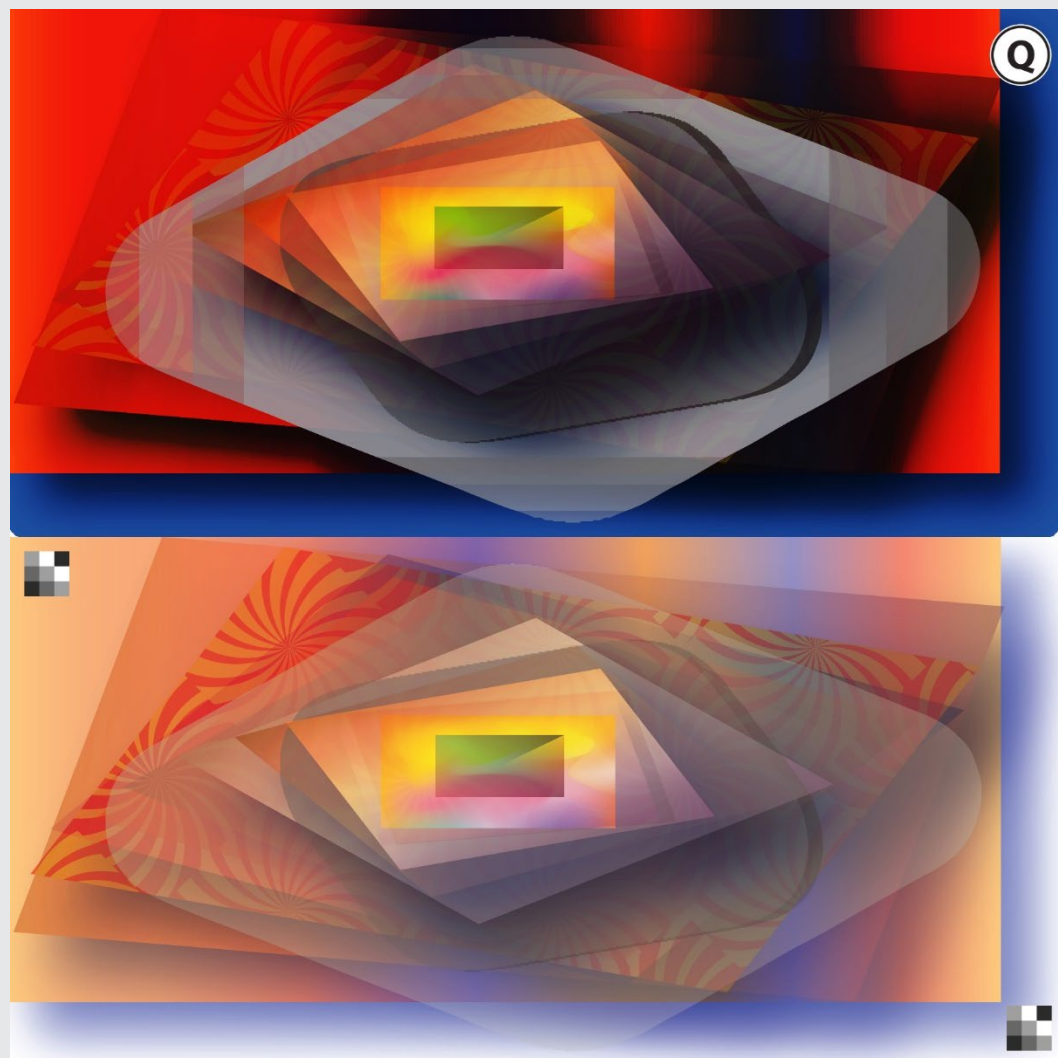


PDFBox: Ghent

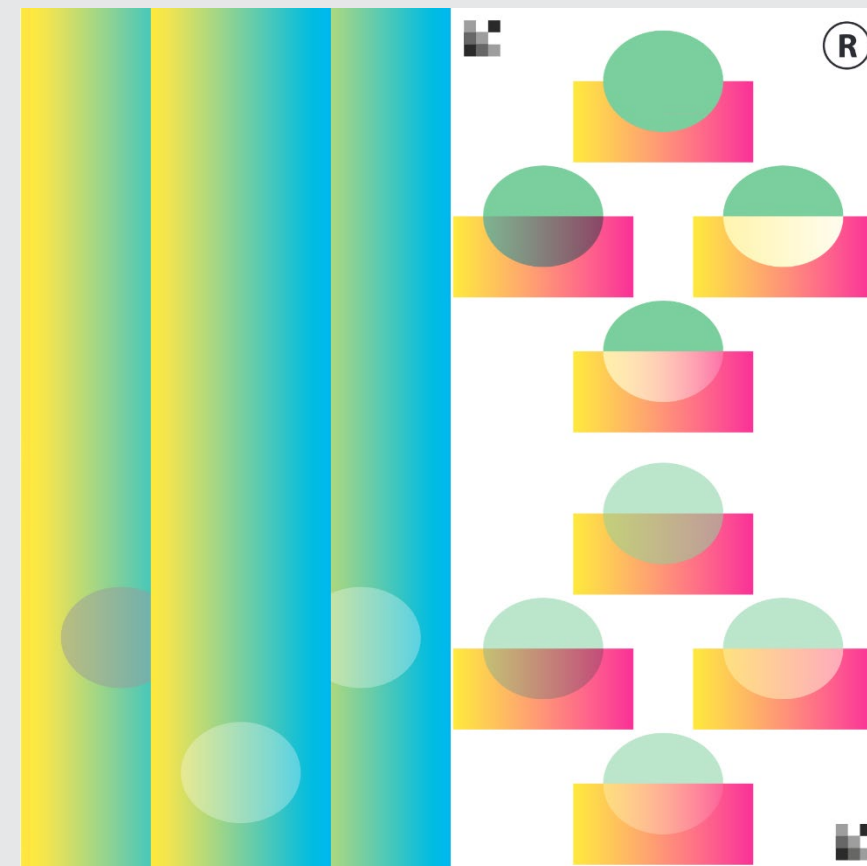
Poppler: Ghent

XPDF: Ghent

Sampling of rendering issues

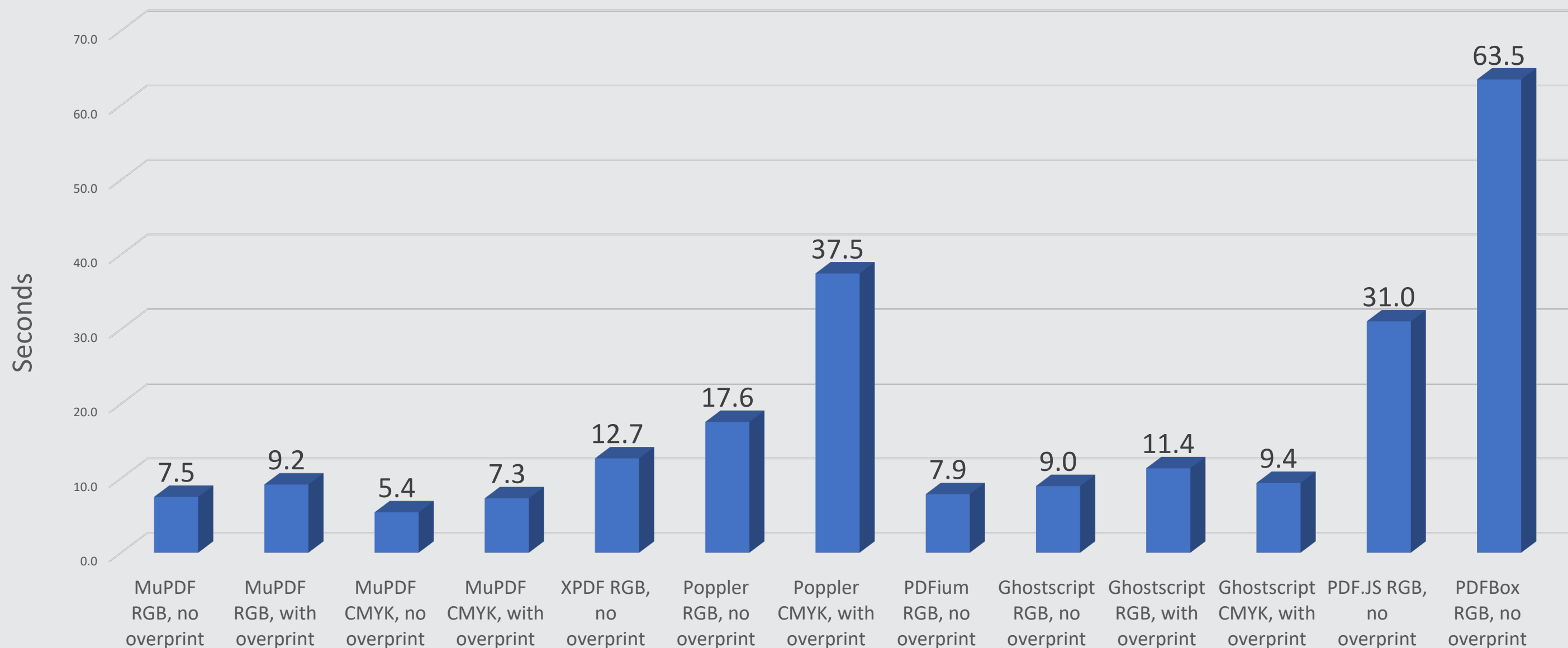


PDFium: Altona

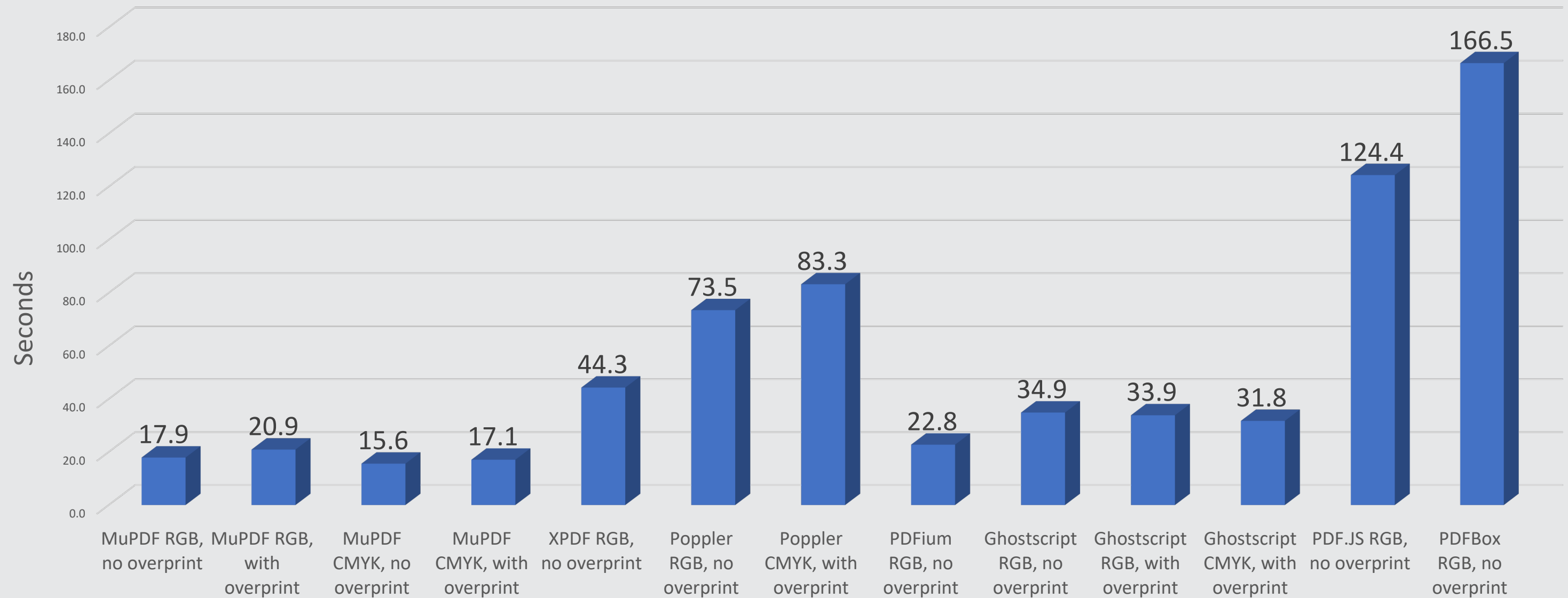


PDF.js: Altona

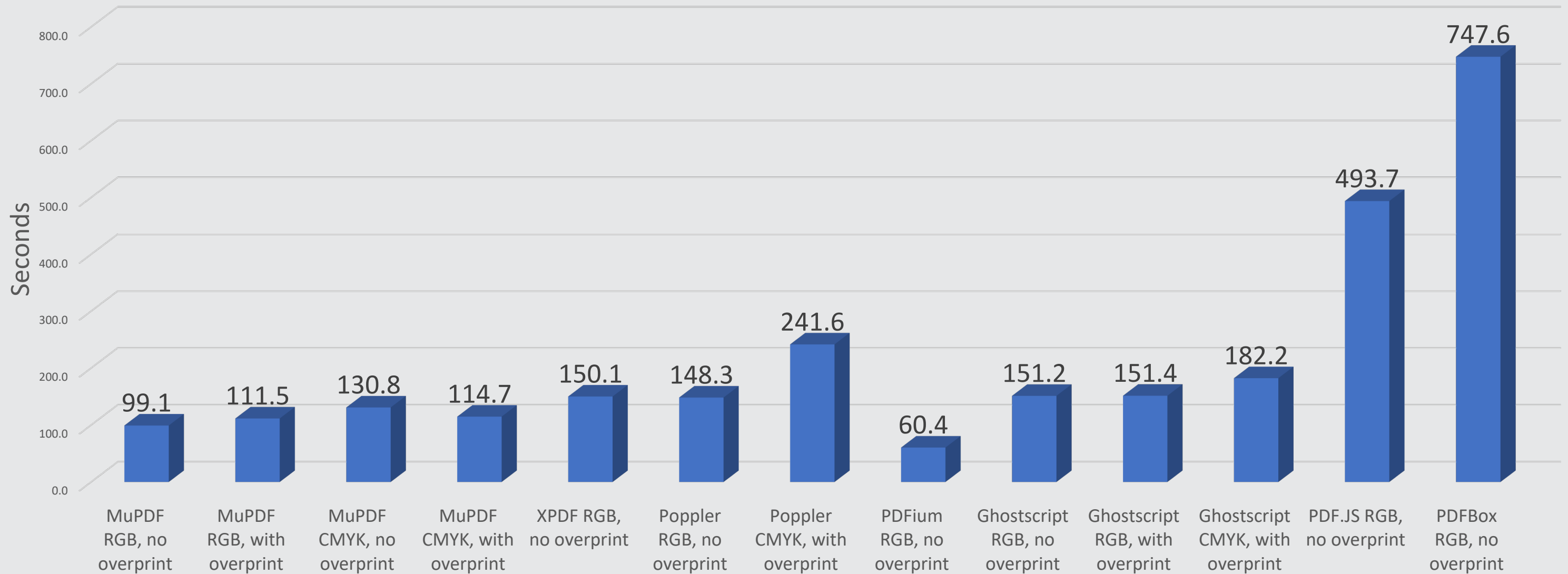
Timing performance, Ghent 600dpi



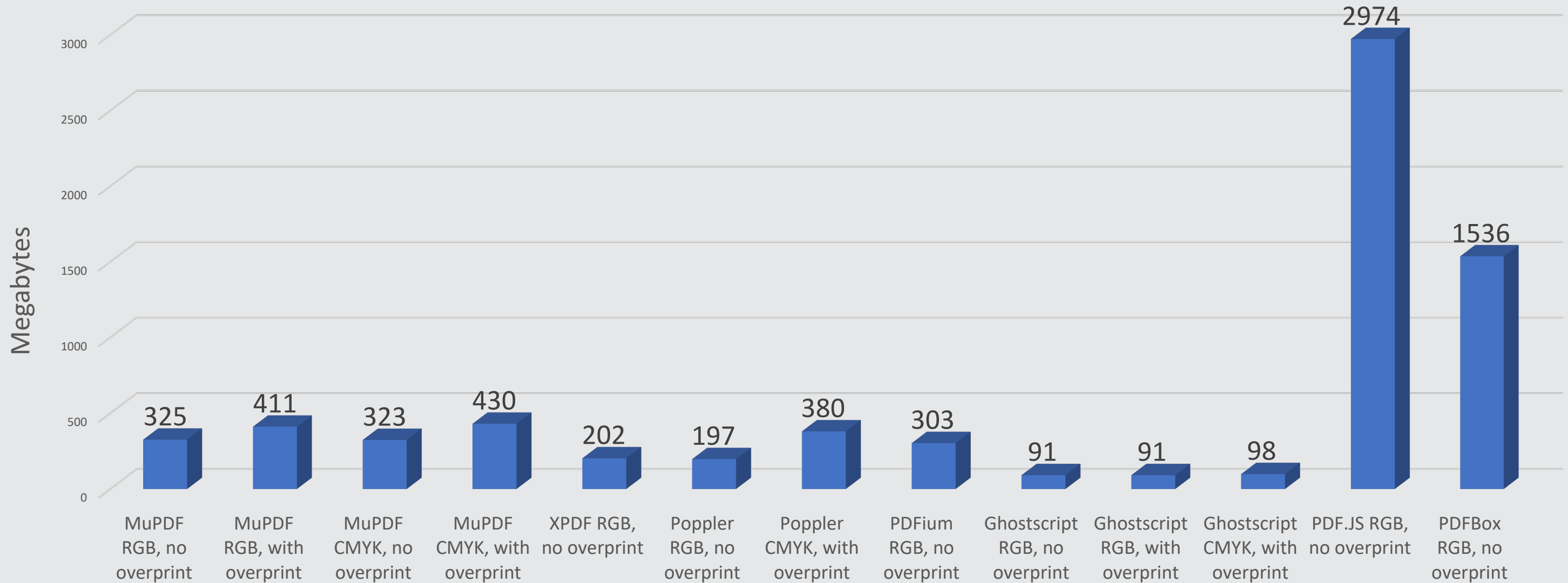
Timing performance, Altona 600dpi



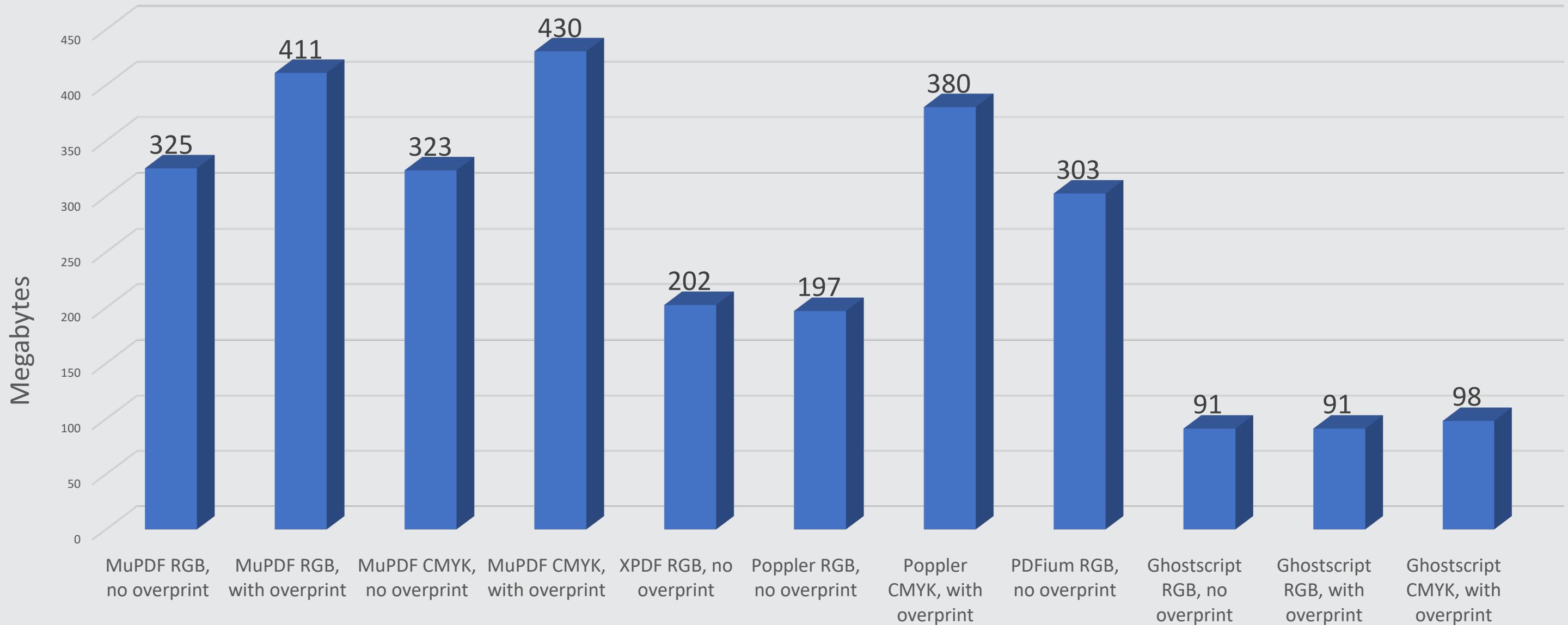
Timing perf., PDF 1.7 specification 150dpi



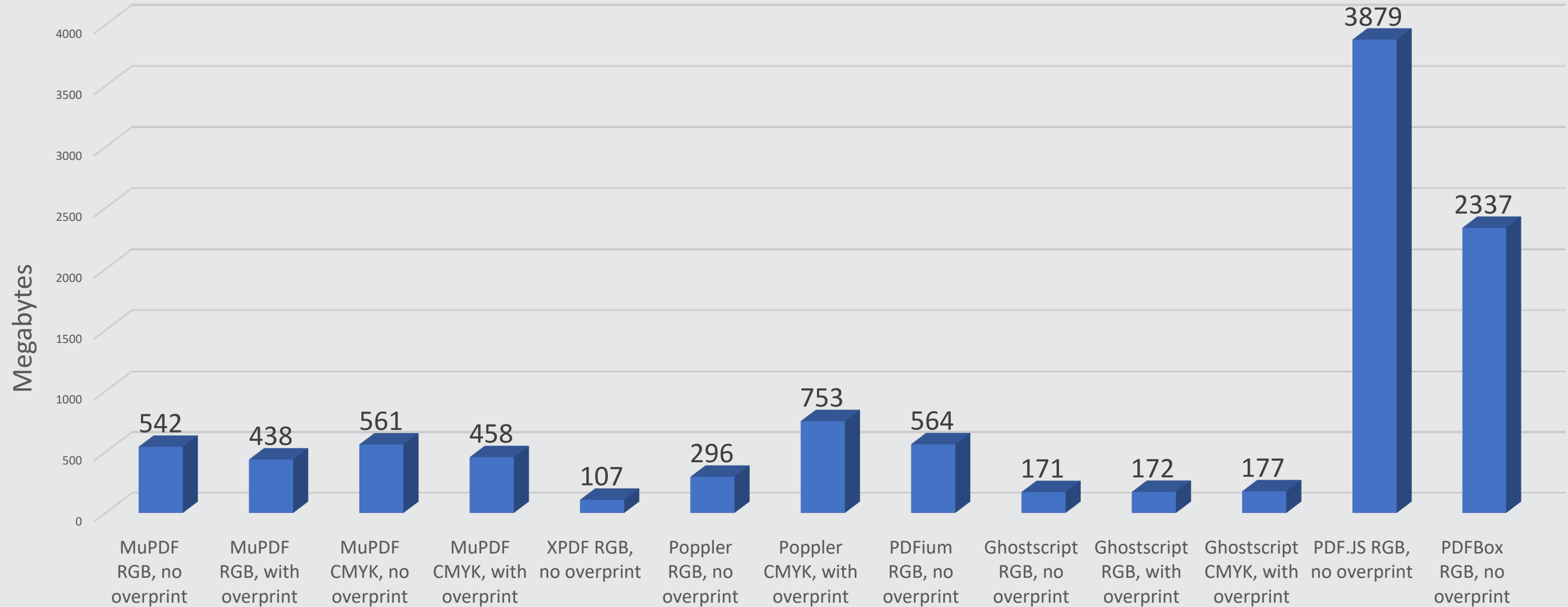
Memory performance, Ghent 600dpi



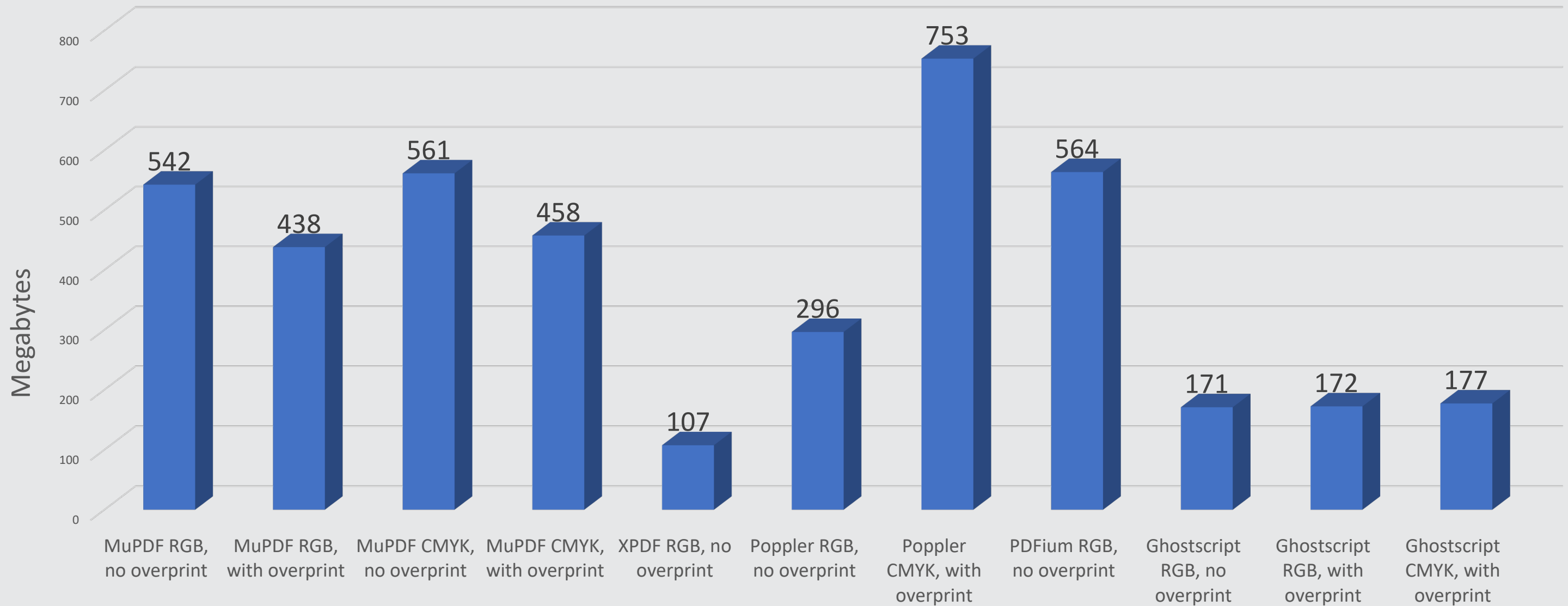
Memory performance, Ghent 600dpi



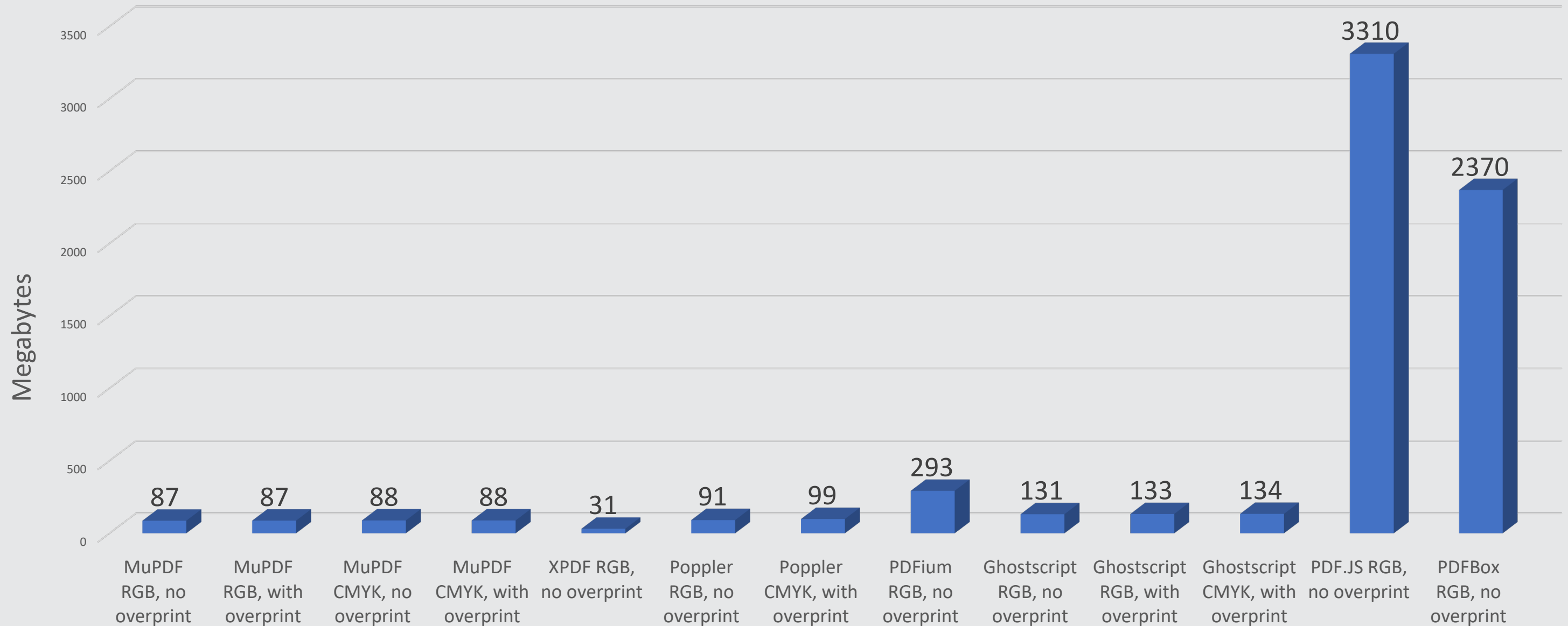
Memory performance, Altona 600dpi



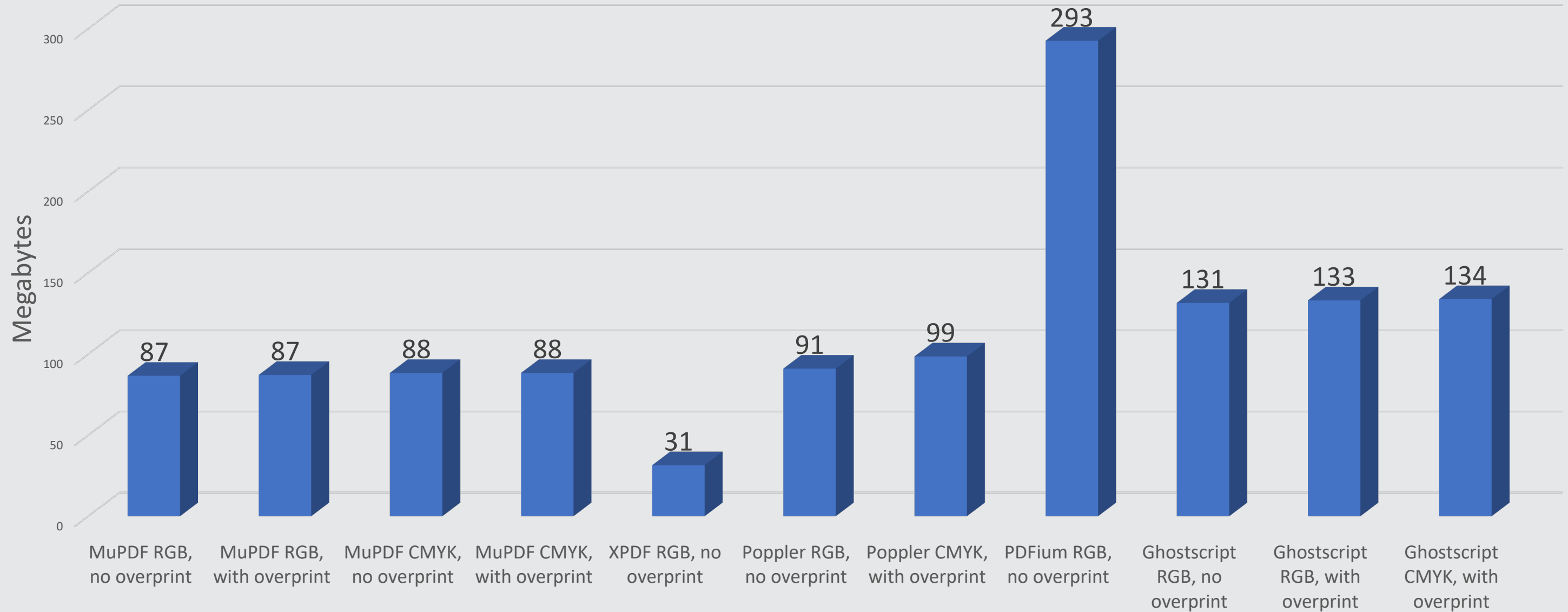
Memory performance, Altona 600dpi



Memory performance, PDF 1.7 spec. 150dpi



Memory performance, PDF 1.7 spec. 150dpi



Summary



Several options in terms of license types available.

GPL : Ghostscript, MuPDF, XPDF, Poppler

Apache: PDFBox, PDF.JS, PDFium

The native language solutions had similar performance.

Native: Ghostscript, MuPDF, XPDF, Poppler, PDFium

Virtual Machine: PDFBox, PDF.JS

Some issues seen in corner cases, but all did well in text document rendering quality

Transparency and overprint.

About Artifex Software



- Roots in open-source with Ghostscript and MuPDF
- Solutions for PDL processing/conversion
PDF, PS, PCL, XPS
- Solutions for Office documents
(doc, docx, xls, xlsx, ppt, pptx)
- Solutions for embedded, cloud,
enterprise, print, mobile
- Partners with over 150 OEM leaders such as
Google, Blackberry, Garmin...
- Engineering staff in U.S., Europe, and Asia
able to provide prompt support
- Over 100 consecutive quarters of profitability

