



# Delo s PDF dokumenti

## Working with PDF documents

Matjaž Kljun [matjaz.kljun@upr.si](mailto:matjaz.kljun@upr.si)

---

# Oblika delavnice / Workshop format

- 4 urna delavnica (predstavitev, delo s premori)
  - 17. 3. 2025
- Kraj izvedbe: UP PEF

## Vsebina:

- Na kratko o PDF
- Prebiranje PDF
- Ustvarjanje PDF
- Pretvarjanje PDF v ...
- Stiskanje PDF
- Zaščita PDF
- OCR PDF
- Obrazci
- Podpisovanje PDF
- Razdeljevanje in združevanje PDF
- Popravljanje besedila
- PDF in umetna inteligenca

- 4-hour workshop (presentation, hands-on work with breaks)
  - 17 March 2025
- Venue: UP PEF

## Content:

- PDF at a glance
- Reading PDF
- Create PDF
- Convert PDF to ...
- Compress PDF
- Protect PDF
- OCR PDF
- Forms
- Sign PDF
- Split and merge PDF
- Editing text
- PDF and artificial intelligence

Na kratko o PDF

---

PDF at a glance

# Kaj je PDF? / What is PDF?

PDF (Portable Document Format) je vsestranski datotečni format,

- ki ga je razvil Adobe Systems v zgodnjih devetdesetih letih prejšnjega stoletja,
- da bi olajšal izmenjavo oblikovanih dokumentov
- ne glede na sistem pošiljatelja in prejemnika.

Format temelji na postscriptu in je v bistvu njegova poenostavitev,

- Postscript je programski jezik, ki ga je desetletje prej razvil Adobe za opisovanje dokumentov za tiskanje na tehnološko nevtralen način.

PDF (Portable Document Format) is a versatile file format developed

- by Adobe Systems in the early 1990s
- to make it easy to share formatted documents
- regardless of the sender and receiver's systems.

The format was based on and is essentially a simplification of postscript,

- PS is a programming language developed by Adobe 10 years earlier to describe documents for printing in a technology-neutral way.

- PDF (in postscript) uporablja kodo za opisovanje položaja besedila, slik in grafičnih elementov na strani.
- Rasterizer kodo spremeni v pike na zaslonu ali pike na izpisu.
- Datoteke PDF so zelo pogoste in se uporabljajo za različne namene:
  - kot arhivski format za skenirane in digitalizirane dokumente različnih vrst,
  - kot standardni format za znanstvene članke,
  - za digitalne pogodbe in
  - za shranjevanje za poznejše tiskanje, če naštejemo le nekaj primerov.
- PDF (and postscript) uses code to describe where text, images, and graphic elements should be placed on a page.
- A rasterizer turns the code into either pixels on a screen or dots in a printout.
- PDF files are extremely common and are used for all sorts of things:
  - as an archive format for scanned and digitized documents of various kinds,
  - as a standard format for scientific articles,
  - for digital contracts and
  - to save something for later printing, just to name a few examples.

## Prednosti / Advantages

- Univerzalna združljivost (deluje na vseh napravah).
- Fiksno oblikovanje (ob odprtju v različnih programih se postavitev ne spremeni).
- Varnostne funkcije (zaščita z geslom, šifriranje, digitalni podpisi).
- Podpira večpredstavnost (besedilo, slike, videoposnetke ...) in interaktivne elemente (obrazce, gumbe ...).
- Universal compatibility (works on all devices).
- Fixed formatting (no layout changes when opened on different software).
- Security features (password protection, encryption, digital signatures).
- Supports multimedia (text, images, videos ...) and interactive elements (forms, buttons ...).

# Vrste PDF dokumentov / PDF document types

- PDF/A - Dolgoročno arhiviranje.
- PDF/X - Profesionalno tiskanje.
- PDF/UA - Skladnost z zahtevami za dostopnost.
- PDF/E - Inženiring in CAD.
- PDF/VT - Tiskanje velikih količin transakcij.
- PDF/H - Zdravstvena in medicinska dokumentacija.
- PDF/B - Splošna izmenjava dokumentov.



- PDF/A – Long-term preservation.
- PDF/X – Professional printing.
- PDF/UA – Accessibility compliance.
- PDF/E – Engineering and CAD.
- PDF/VT – High-volume transactional printing.
- PDF/H – Healthcare and medical records.
- PDF/B – General document exchange.

## PDF/A – Arhiviranje / Archiving

PDF/A (Portable Document Format - Archival) je specializirana različica PDF, standardizirana po standardu ISO, **namenjena dolgoročnemu shranjevanju** in dostopnosti.

Zagotavlja, da bodo datoteke ostale berljive daljše časovno obdobje, saj morajo vsebovati vse potrebne elemente. Oblika tudi omejuje funkcije, ki bi lahko povzročile težave pri upodabljanju na zaslonu.

PDF/A (Portable Document Format – Archival) is an ISO-standardized version of PDF designed for **long-term preservation** and accessibility.

It ensures that files will remain readable over long period of time by embedding all necessary elements and restricting features that might cause rendering issues.



## Značilnosti / Characteristics

- Samostojen - vse pisave, slike in metapodatki so vgrajeni.
- Neodvisen od naprave - dosledno upodabljanje na različnih platformah.
- Brez zunanje vsebine ali šifriranja - Zagotavlja stabilnost in dostopnost.
- Zahteve za metapodatke - uporablja metapodatke XMP za boljšo organizacijo.
- Upravljanje barv - profili ICC zagotavljajo natančno predstavitev barv.
- Omejena prozornost in plasti - preprečuje neskladnosti pri upodabljanju.
- Self-Contained – All fonts, images, and metadata are embedded.
- Device-Independent – Consistent rendering across different platforms.
- No External Content or Encryption – Ensures stability and accessibility.
- Metadata Requirements – Uses XMP metadata for better organization.
- Color Management – ICC profiles ensure accurate color representation.
- Restricted Transparency and Layers – Prevents rendering inconsistencies.

# Vrte PDF/A / PDF/A types

## PDF/A-1 (2005 - na podlagi PDF 1.4)

- Dva podtipa: PDF/A-1a (dostopen) in PDF/A-1b (osnovni)
- Ne podpira prosojnosti, slojev ali stiskanja JPEG2000.
- Primeren za preprosto hrambo dokumentov.

## PDF/A-2 (2011 - temelji na PDF 1.7)

- Podpira prosojnost, plasti in stiskanje JPEG2000.
- Omogoča vstavljanje pisav OpenType.
- Trije podtipi:
  - PDF/A-2a (popolnoma dostopni, označeni dokumenti)
  - PDF/A-2b (zagotovljena vizualna ponovljivost, ne pa tudi dostopnost)
  - PDF/A-2u (dodan preslikavo besedila Unicode za iskanje)

## PDF/A-3 (2012 - temelji na PDF 1.7)

- Iste lastnosti kot PDF/A-2, vendar omogoča vstavljanje katere koli vrste datotek.
- Uporabno za shranjevanje povezanih dokumentov v enem samem arhivu.

## PDF/A-4 (2020 - temelji na PDF 2.0)

- Uvaja eno samo stopnjo skladnosti (ni več razlik med „a“ in „b“).
- Omogoča vstavljene datoteke, vendar s strožjimi pravili.
- Izboljša ravnanje z digitalnim podpisom in šifriranjem.

## PDF/A-1 (2005 – Based on PDF 1.4)

- Two subtypes: PDF/A-1a (Accessible) & PDF/A-1b (Basic)
- Does not support transparency, layers, or JPEG2000 compression.
- Suitable for simple document preservation.

## PDF/A-2 (2011 – Based on PDF 1.7)

- Supports transparency, layers, and JPEG2000 compression.
- Allows embedding of OpenType fonts.
- Three subtypes:
  - PDF/A-2a (Fully accessible, tagged documents)
  - PDF/A-2b (Ensures visual reproducibility, but not accessibility)
  - PDF/A-2u (Adds Unicode text mapping for searchability)

## PDF/A-3 (2012 – Based on PDF 1.7)

- Same features as PDF/A-2 but allows embedding of any file type.
- Useful for storing related documents within a single archive.

## PDF/A-4 (2020 – Based on PDF 2.0)

- Introduces a single conformance level (no more ‘a’ or ‘b’ distinctions).
- Allows for embedded files but with stricter rules.
- Improves digital signature and encryption handling.

# Primerjava / Comparison

Type	Based On	Transparency & Layers	Unicode Text Search	Embedded Files	Best For
<b>PDF/A-1a</b>	PDF 1.4	No	Yes	No	Accessible documents
<b>PDF/A-1b</b>	PDF 1.4	No	No	No	Basic archiving, scanned documents
<b>PDF/A-2a</b>	PDF 1.7	Yes	Yes	No	Fully accessible documents
<b>PDF/A-2b</b>	PDF 1.7	Yes	No	No	Modern PDFs with layers, transparency
<b>PDF/A-2u</b>	PDF 1.7	Yes	Yes	No	Searchable, accessible documents. Best balance between preservation and accessibility.
<b>PDF/A-3a</b>	PDF 1.7	Yes	Yes	Yes	Documents requiring embedded files
<b>PDF/A-3b</b>	PDF 1.7	Yes	No	Yes	General archiving with attachments
<b>PDF/A-3u</b>	PDF 1.7	Yes	Yes	Yes	Searchable documents with embedded files
<b>PDF/A-4</b>	PDF 2.0	Yes	Yes	Yes	Future-proofing with PDF 2.0 features

## PDF/B (PDF Baseline)

- Namen: Poenostavljena, osnovna različica PDF, zasnovana za zanesljivo izmenjavo in prikaz v različnih napravah.
- Ključne funkcije:
  - Zagotavlja dosledno prikazovanje in tiskanje, brez dodatnih pisav ali virov.
  - Ne uveljavlja arhivskih zahtev ali zahtev, specifičnih za tiskanje, kot sta PDF/A ali PDF/X.
- Primer uporabe: Uporaba: Splošna izmenjava dokumentov, kjer napredne funkcije niso potrebne.
- Purpose: A simplified, baseline version of PDF designed for reliable exchange and display across different devices.
- Key Features:
  - Ensures consistent viewing and printing without requiring additional fonts or resources.
  - Does not enforce archival or print-specific requirements like PDF/A or PDF/X.
- Use Case: General-purpose document exchange where advanced features aren't needed.

## PDF/E – Inženirstvo / Engineering

PDF/E je različica PDF, standardizirana po standardu ISO, namenjena inženirskemu sektorju, zlasti geoprostorskim, gradbenim in proizvodnim delovnim tokovom.

Podpira interaktivne medije, vključno s 3D vsebino, in je namenjena omogočanju varne, zanesljive in bogate izmenjave vsebin v inženirski in tehnični industriji.

PDF/E is an ISO-standardized version of PDF designed for the engineering sector, particularly for geospatial, construction, and manufacturing workflows.

It supports interactive media, including 3D content, and is aimed at enabling secure, reliable, and rich content exchange within the engineering and technical industries.

## PDF/X – Tiskanje / Printing

PDF/X je različica PDF, ki jo je ravnotako standardiziral ISO in je prilagojena potrebam grafične industrije, zlasti za izmenjavo grafike pred tiskom.

PDF/X zagotavlja, da se lahko datoteke natisnejo in reproducirajo točno tako, kot je bilo predvideno, z vsemi potrebnimi vsebinami in barvnimi informacijami, ki so vgrajene v datoteko.

PDF/X is another ISO-standardized version of PDF that is tailored for the needs of the graphic arts industry, specifically for prepress graphics exchange.

PDF/X ensures that files can be printed and reproduced exactly as intended, with all necessary content and color information embedded in the file.

- Ključne lastnosti:
    - Zagotavlja barvno natančnost (prek CMYK, neposrednih barv in profilov ICC).
    - Zahteva vgradnjo vseh pisav.
    - Prepoveduje prosojnost, JavaScript in plasti, da se preprečijo napake pri tiskanju.
  - Običajne različice:
    - PDF/X-1a: Podpira CMYK in vključuje vse potrebne elemente za tiskanje.
    - PDF/X-3: Podpira barvne profile RGB in ICC.
    - PDF/X-4: Omogoča prosojnost in plasti, hkrati pa ohranja varnost tiskanja.
  - Primeri uporabe: založništvo, revije, časopisi in komercialni tisk.
- Key Features:
    - Ensures colour accuracy (via CMYK, spot colours, and ICC profiles).
    - Requires all fonts to be embedded.
    - Prohibits transparency, JavaScript, and layers to prevent printing errors.
  - Common Variants:
    - PDF/X-1a: CMYK-based, ensures all necessary elements are included for printing.
    - PDF/X-3: Supports RGB and ICC color profiles.
    - PDF/X-4: Allows transparency and layers while maintaining print safety.
  - Use Case: publishing, magazines, newspapers, and commercial printing.
-

## PDF/UA – Dostopnost / Accessibility

PDF/UA, znan tudi kot PDF za univerzalno dostopnost, je standardizirana oblika PDF po standardu ISO, ki je zasnovana tako, da so dokumenti dostopni in vodljivi za osebe s posebnimi potrebami.

Zagotavlja smernice za ustvarjanje datotek PDF, ki so dostopne s tehnologijo, kot so bralniki zaslona, brajlovi zasloni in druga pomožna orodja. Ta standard je ključnega pomena za zagotavljanje enakega dostopa do informacij vsem uporabnikom, ne glede na njihove sposobnosti.

PDF/UA, also known as PDF for Universal Accessibility, is an ISO-standardized form of PDF designed to make documents accessible and navigable for individuals with disabilities.

It provides guidelines for creating PDF files that are accessible through technology such as screen readers, braille displays, and other assistive tools. This standard is crucial for ensuring that all users, regardless of ability, have equal access to information.



# Primerjava standardov / Comparison of standards

Standard	Purpose	Key Features	Best For
<b>PDF/A</b>	Archival & Long-Term Preservation	Embeds all content, restricts external dependencies	Government, legal, library archives
<b>PDF/X</b>	Prepress & Printing	Ensures color accuracy, embedded fonts, no transparency issues	Publishing, magazines, newspapers
<b>PDF/UA</b>	Universal Accessibility	Requires tagged PDFs, structured content for screen readers	Government, legal, educational documents
<b>PDF/E</b>	Engineering & CAD	Supports 3D models, layers, technical drawings	Architecture, engineering, manufacturing
<b>PDF/VT</b>	Variable & Transactional Printing	Optimized for mass, data-driven printing (invoices, direct mail)	High-volume personalized printing
<b>PDF/H</b>	Healthcare & Medical	Compliant with medical standards (DICOM, HL7)	Medical imaging, patient records
<b>PDF/B</b>	Basic Document Exchange	Simplified for reliable exchange and display	General-purpose PDFs

Prebiranje PDF

---

Reading PDF

# Programi / Software

- MS Edge, Google Chrome, Mozilla Firefox ...
  - Adobe Acrobat / Adobe Reader DC
    - Izvirni program
    - Velik in okoren
    - Reader je zmogljiv a plačljiv
  - Sumatra PDF
    - <https://www.sumatrapdfreader.org/>
    - Brezplačen, hiter, majhen, opremljen s funkcijami, prilagodljiv
  - Okular <https://okular.kde.org/en-gb/>
    - Odprtokodni
- MS Edge, Google Chrome, Mozilla Firefox ...
  - Adobe Acrobat / Adobe Reader DC
    - Original
    - Big and cumbersome
    - Reader is very versatile but costly
  - Sumatra PDF
    - <https://www.sumatrapdfreader.org/>
    - Free, fast, small, packed with features, customizable
  - Okular <https://okular.kde.org/en-gb/>
    - Opensource
-

- Foxit <https://www.foxit.com/>
    - Brezplačna poskusna različica
  - PDF Gear <https://www.pdfgear.com/>
    - Brezplačen
    - Branje, urejanje, pretvarjanje, združevanje in podpisovanje datotek PDF
    - Integrirana umetna inteligenca
  - PDF-Xchange <https://pdf-xchange.eu/>
    - Brezplačne, profesionalne in druge različice
  - Mnogi drugi (MiniTool PDF Editor, etc.)
- Foxit <https://www.foxit.com/>
    - Free trial
  - PDF Gear <https://www.pdfgear.com/>
    - Free
    - Read, edit, convert, merge, and sign PDF files
    - AI integrated
  - PDF-Xchange <https://pdf-xchange.eu/>
    - Free, pro and other versions
  - Plenty of others (MiniTool PDF Editor, etc.)
-

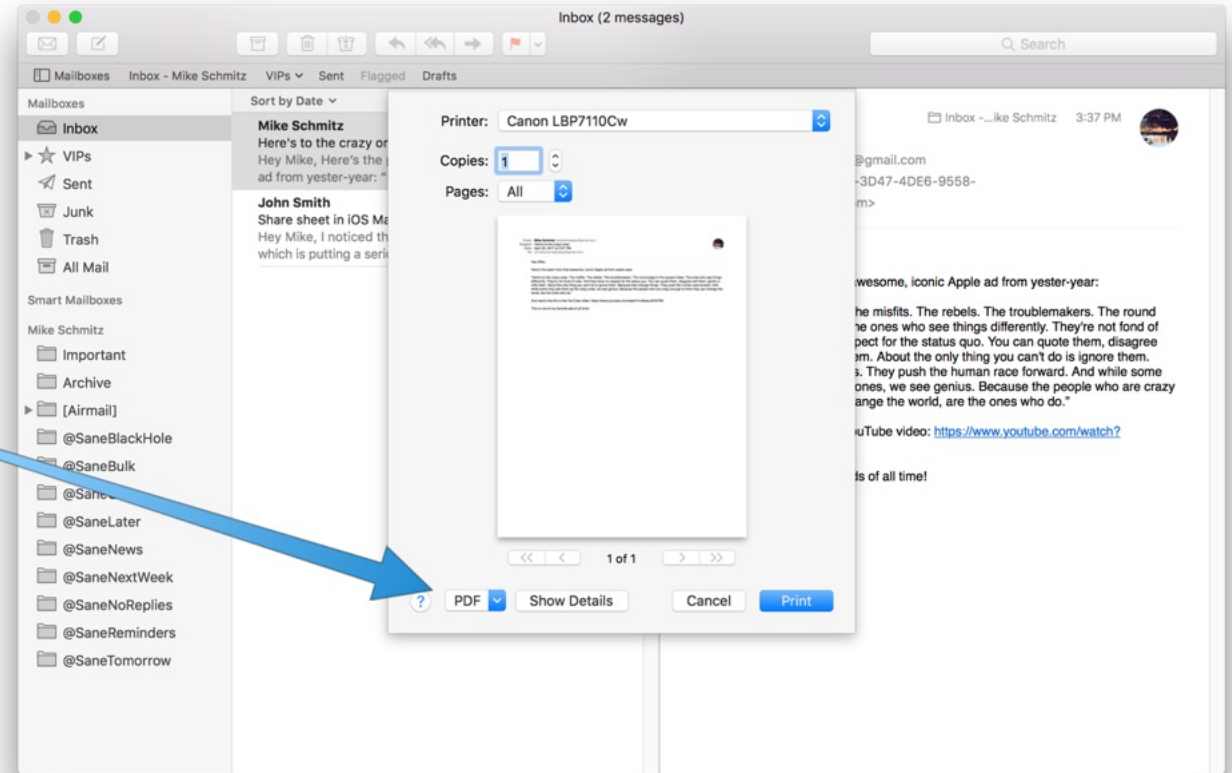
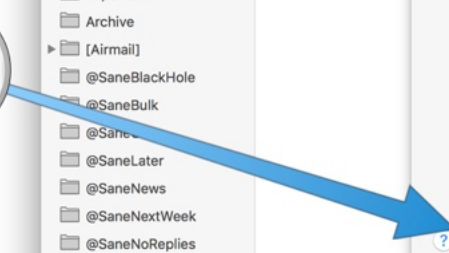
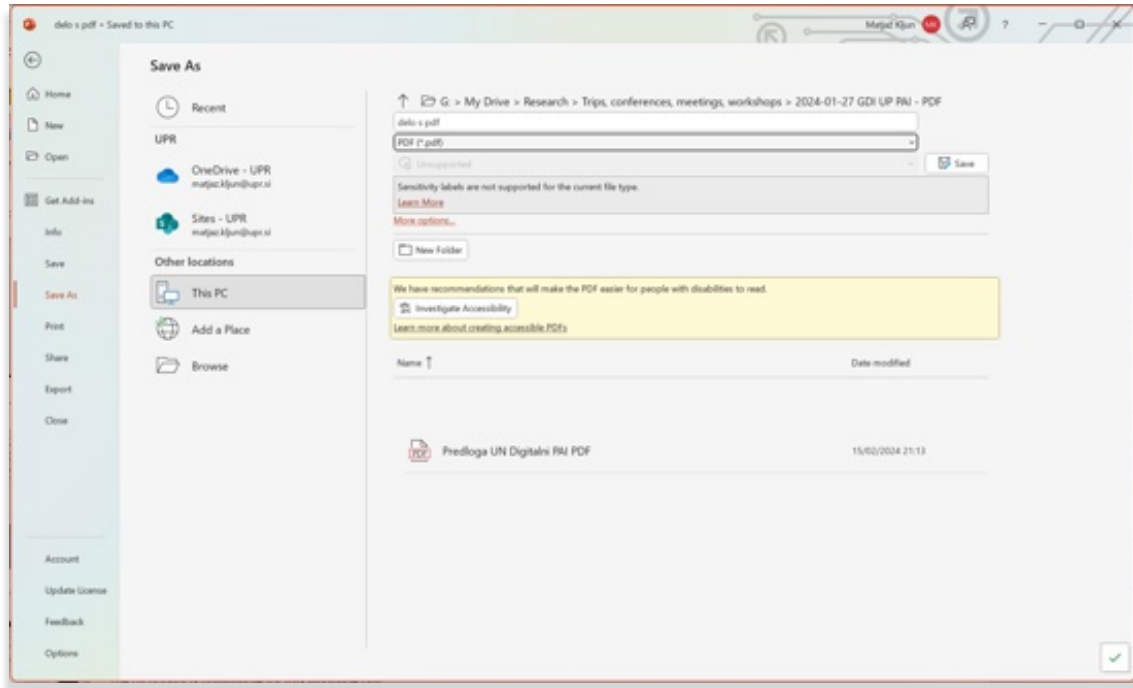
# Ustvarjanje PDF

---

## Create PDF

# Načini / Ways of doing it

- Skeniranje dokumentov z optičnim bralnikom
    - Slike
  - Fotografiranje dokumentov
    - Android: MS Lens, PDFgear Scan
    - iPhone: MS Lens, PDFgear Scan
  - Shrani kot PDF / Izvozi kot PDF
    - To podpirajo nekateri programi
      - MS office, Libre Office, Safari
  - Tiskanje v PDF
    - MS Windows 10 in 11:
      - Izberemo “tiskalnik” Microsoft print to PDF
        - Ne ustvari vedno datotek z besedilom, ki ga je mogoče izbrati, preiskati in označiti.
      - Cute PDF writer:  
<https://www.cutepdf.com/products/cutepdf/writer.asp>
    - macOS: Izberemo PDF v oknu za tiskanje
- Scanning documents with a scanner
    - Images
  - Taking photos of documents
    - Android: MS Lens, PDFgear Scan
    - iPhone: MS Lens, PDFgear Scan
  - Save as PDF / Export as PDF
    - Some software support this
      - MS office, Libre Office, Safari
  - Print to PDF
    - MS Windows 10 and 11:
      - Select Microsoft print to PDF “printer”
        - It does not always produce files with text that can be selected, searched, and labelled
      - Cute PDF writer:
    - macOS: Select PDF in the print dialog box



Pretvarjanje PDF v ...

---

Convert PDF to ...



## Načini / Ways of doing it

- Najpogostejše pretvorbe so v Microsoft Office (Word, Excel, Powerpoint) in slike (JPEG, PNG), obstajajo pa tudi druge, na primer pretvorba PDF v EPUB za knjige.

- Spletni pretvorniki

- Acrobat online

<https://www.adobe.com/acrobat/online.html>

- I  PDF

<https://www.ilovepdf.com>

- Online PDF converter

<https://www.freepdfconvert.com/>

- Pomislimo, če je document, ki ga nalagamo na splet, občutljive narave!!!

- The most common conversions are to Microsoft Office (Word, Excel, Powerpoint) and images (JPEG, PNG) but there are also others, such as PDF to EPUB for books.

- Online

- Acrobat online

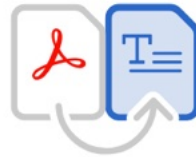
- I  PDF

- Online PDF converter

- Think if the document of a sensitive nature before uploading it to the web!!!



# Convert PDF to Word



Drag and drop a PDF file to use our PDF to Microsoft Word converter.

[Select a file](#)



Your files will be securely handled by Adobe servers and deleted unless you sign in to save them. ⓘ  
By using this service, you agree to the Adobe [Terms of Use](#) and [Privacy Policy](#).

# Težave / Problems

- Oblikovanje se ne bo vedno ohranilo
  - Nikoli ne bomo dobili nazaj izvirnika
    - pogoste napake
      - težave s pisavo,
      - neuskrajene črke in številke,
      - napačne besede,
      - napake pri krepkem, podčrtanem, ležečem besedilu,
      - Napake pri lomljenih besedah čez vrstice,
      - izginjajoče povezave,
      - nepravilna postavitev strani
      - izkrivljene grafike itd.
    - Dobro je shraniti izvirno datoteko skupaj z datoteko PDF
  - Dokumenti kot slike lahko ostanejo le slike
- Formatting cannot be always preserved
  - We will never get back the original
    - common errors
      - font problems,
      - disjointed letters and numbers,
      - wrong words,
      - bold/underline/Italics errors,
      - hyphenation confusion,
      - disappearing links,
      - incorrect page layout
      - distorted graphics, etc.
    - Good idea to save both the original file and PDF
  - Documents as images might just remain images

Stiskanje PDF

---

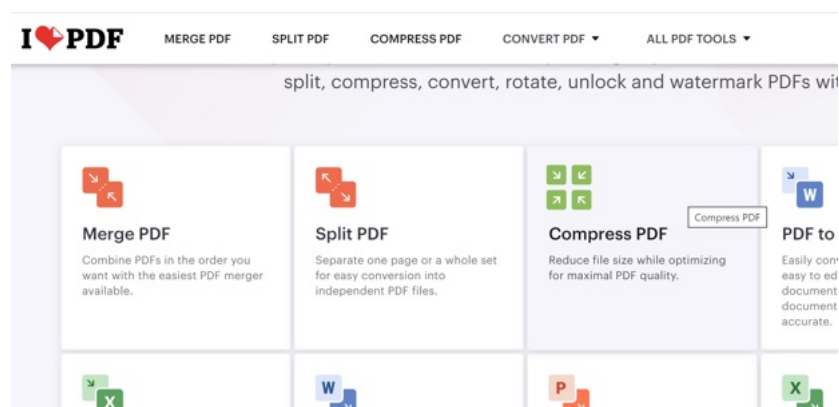
Compress PDF

# Zmanjšanje velikosti datoteke / Reduce the file size

- Če ne potrebujemo barvnih slik visoke ločljivosti, čeprav bomo datoteko natisnili, jo lahko stisnemo.
- Acrobat online
- I ❤️ PDF
- If we don't need high-resolution color images even though we will print the file, we can compress it.
- Acrobat online
- I ❤️ PDF

<https://www.adobe.com/acrobat/online.html>

<https://www.ilovepdf.com>



Zaščita PDF

---

Protect PDF

# Načini / Ways of doing it

- S šifriranjem in geslom.
  - Uporablja šifriranje AES in je varno, če je geslo dolgo.
- Velika večina programov za PDF lahko odpira in shranjuje šifrirane datoteke.
- Če imamo datoteko, zaščiteni z geslom, vendar ne želimo vsakič znova vnašati le-to, jo lahko odpremo in natisnemo v PDF, pri čemer naredimo novo kopijo brez gesla, ali pa uporabimo program, ki to podpira.
  - Na primer:
    - Adobe Acrobat Pro
    - Adobe PDF Online
    - PDF Gear
- With encryption and a password.
  - It uses the AES encryption and is secure provided the password is long.
- The vast majority of PDF programs can open and save encrypted files.
- If we have a password-protected file but we don't want to have to fill in the password every time, we can either open it and print it to PDF, making a new password-free copy of it, or use a third-party program that supports it.
  - For example:
    - Adobe Acrobat Pro
    - Adobe PDF Online
    - PDF Gear

delo s.pdf.pdf - PDFgear

Home Annotate Edit Form Page Tools Help

Convert Compress Merge PDF Split PDF Signature Password Remove

Thumbnail

Delo s PDF dokumenti  
Working with PDF documents  
matjaz.kljun@upr.si

1

Oblika datoteke / Workshop format

2

Na kratko o PDF  
PDF at a glance

3

Kaj je PDF? / What is PDF?

4

Delo s PDF dokumenti  
Working with PDF documents

Matjaž Kljun [matjaz.kljun@upr.si](mailto:matjaz.kljun@upr.si)

PDFgear Copilot - Chat with PDF

PDFgear - Powered by ChatGPT  
**Chat, Learn, Earn**

With PDFgear Copilot, your documents are becoming intelligent. It works great for commanding operations on PDFs and obtain precise information from them using natural language. Try chatting with your files right now.

Chat Now

1/45 125%



OCR PDF

---

OCR PDF

## Težava in rešitev / The problem and solution

- Pri skeniranju dokumentov dobimo slike in ne digitalno besedilo
  - besedila ni mogoče označiti in kopirati ali po njem iskati.
- Rešitev je programska oprema s funkcijo OCR.
- To ima vgrajeno več brezplačnih programov
  - PDF Gear
  - PDF-Xchange.
  - OCRmyPDF - najboljši, vendar dela v ukazni vrstici, postopek namestitve pa je nekoliko zapleten.
    - Za pogostejšo uporabo (vprašajte IKT službo)
  - I ❤️ PDF
- When documents are scanned, they are images and not digital text
  - it is not possible to highlight text and copy or search it.
- The solution is software with OCR functionality.
- Several free programs have this built in
  - PDF Gear
  - PDF-Xchange.
  - OCRmyPDF – The best but it is a command line, and the installation procedure is somewhat complicated.
    - For frequent use (ask your ICT office)
  - I ❤️ PDF

<https://www.ilovepdf.com>



### PDF to JPG

Convert each PDF page into a JPG or extract all images contained in a PDF.

### JPG to PDF

Convert JPG images to PDF in seconds. Easily adjust orientation and margins.

### Sign PDF

Sign yourself or request electronic signatures from others.

### Watermark

Stamp an image or text over your PDF in seconds. Choose the typography, transparency and position.

### Rotate PDF

Rotate your PDFs the way you need them. You can rotate multiple PDFs at once.



### HTML to PDF

Convert webpages in HTML to PDF. Copy and paste the URL of the page you want and convert it to PDF with a click.



### Unlock PDF

Remove PDF password security, giving you the freedom to use your PDFs as you want.



### Protect PDF

Protect PDF files with a password. Encrypt PDF documents to prevent unauthorized access.



### Organize PDF

Sort pages of your PDF file however you like. Delete PDF pages or add PDF pages to your document at your convenience.



### PDF to PDF/A

Transform your PDF into ISO-standardized PDF/A for long-term archival. PDF/A will preserve format and content for future access.



### Repair PDF

Repair a damaged PDF and recover data from corrupt PDF. Fix PDF files with our Repair tool.



### Page numbers

Add page numbers into PDFs with ease. Choose your positions, dimensions, typography.



### Scan to PDF

Capture document scans from your mobile device and send them instantly to your browser.



### OCR PDF

Easily convert scanned PDF into searchable and selectable documents.

New!

OCR PDF

# Iskanje po PDF / Search PDF

- Ker se dokumenti PDF pogosto uporabljajo za dolge dokumente, je pogosto treba iskati besedilo.
- Zaradi načina delovanja formata PDF to ni vedno enostavno kot pri Wordovih dokumentih.
  - Besedilo ne leži kot blok neformatiranega besedila, za katerega se uporablja oblikovanje, temveč kot niz znakov skupaj z informacijami o tem, kje so nameščeni na strani, kakšno pisavo uporabljajo in tako naprej.
- Pri datotekah PDF, shranjenih neposredno iz Worda ali drugih programov, je to redko težava, pri skeniranih datotekah in datotekah, ki se berejo z OCR, pa ja.
  - To še posebej velja za besedila v stolpcih, tabelah, seznamih ali besedilnih poljih.
  - Izbira besedila se najpogosteje nadaljuje na napačnem mestu.
  - To lahko povzroči težave tudi funkciji iskanja - na primer, ko je beseda lomljena med vrsticami in nima več vseh znakov drug ob drugem.
    - Če ne dobimo zadetka s celotno besedo, poskusimo vnesti samo del besede, po možnosti del, ki je običajno pred ali za prelomom vrstice in je čim bolj edinstven.
- Since PDFs are often used for large documents, it is common to have to search them.
- Because the way the PDF format works, this is not always as easy as in Word documents.
  - The text does not lie as a block of unformatted text to which formatting is applied, but as a stream of characters together with information about where they are placed on the page, what font they have and so on.
- For PDFs saved directly from Word or other programs this is rarely a problem, but for scanned and OCR-read files it can be.
  - This is especially true when text is in columns, tables, lists, or text boxes.
  - Selection of text more often than not continues in the wrong place.
  - It can also cause problems for the search function — for example when a word has been hyphenated and no longer has all the characters next to each other.
    - If we don't get a hit with a whole word, try entering just part of the word, preferably a bit that normally comes before or after any line breaks and is as unique as possible..

Obrazci

---

Forms

# Interaktivni obrazci / Interactive forms

- Datoteke PDF so lahko interaktivne in vključujejo vnosna polja za besedilo, kot sta ime in naslov, potrditvena polja in včasih celo menuje z več izbirami.
- PDFs can be made interactive and include input fields for text such as name and address, tick boxes, and sometimes even multiple choice menus.

The image shows a screenshot of a PDF form titled "AVTAL OM UPPLÄTELSE AV BOSTADSRÄTTSLÄGGENHET I ANDRA HÄND" (Agreement on assignment of tenancy rights to another person). The form is displayed in a web browser window. The form contains several sections with input fields and checkboxes:

- BOSTADSRÄTTSLÄGGENHET:** Fields for name, address, and other personal information.
- UPPLÄTSELSEN:** Fields for the name and address of the assignee.
- LÄSARENHET SOM UPPLÄTSEL I ANDRA HÄND:** A section with checkboxes for "I ägarens namn" and "I hyresvärdens namn".
- INOMHÅLLET:** A section with checkboxes for "Lämnat till ägarens namn", "Lämnat till hyresvärdens namn", "I gemensamt ägande", and "Etablerat för hyresvärdens räkning".
- BÄRARE OCH P-PLÅT:** A section with checkboxes for "Pantstypens bärare" and "Böringsbärare".

The form is dated 2018-10-21 and is page 1 of 1. The browser window shows a file named "Bostad\_rattslagenhet.pdf" and a "Sign in" button in the top right corner.

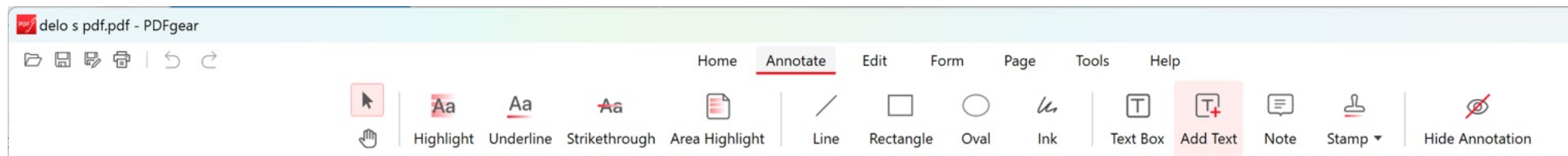
# Neinteraktivni obrazci / Noninteractive forms

- Kako dobro delujejo interaktivni obrazci v programih PDF, ki niso Adobe Acrobat (vključno s programom Reader DC), je vprašljivo.
- Zato je dobro, da imamo na računalniku nameščen program Acrobat Reader DC, tudi če ga nismo nastavili kot privzeti program za odpiranje datotek PDF.
- Če prednastavljena polja niso interaktivna, lahko z orodjem za dodajanje besedila sami dodamo besedilo.

OBRAZEC DR-02 Priloga 1  
Prijava za vpis fizične osebe v davčni register

1. Davčna številka (pri prvi prijavi izpolni FURS)	
2. Osebo ime	Ime <input type="text"/> Priimek <input type="text"/>
3. Datum rojstva	Kraj rojstva
4. Spol (M-moški, Ž-ženska)	EMŠO
5. Državljanstvo	
6. Podatki o rezidentskem statusu	
Rezident Republike Slovenije:	
<input type="checkbox"/> da	
<input type="checkbox"/> ne	Država rezidentstva (za nerezidenta)
	Številka za davčne namene v državi rezidentstva (za nerezidenta)
7. Naslov stalnega prebivališča	
Država	
Ulica in hišna številka	
Naselje	
Pošta	
8. Naslov začasnega prebivališča	
Država	
Ulica in hišna številka	
Naselje	
Pošta	
9. Naslov v Republiki Sloveniji za vročanje	
Ulica in hišna številka	
Naselje	
Pošta	

- How well interactive forms work with PDF programs other than Adobe Acrobat (including Reader DC) varies.
- Therefore, it may be a good idea to have Acrobat Reader DC installed on our computer even if we do not have it set as the default program for opening PDF files.
- If the present fields are not interactive, use the “add text” tool to add text ourself.



# Podpisovanje PDF

---

Sign PDF



# S podpisom (sliko) / With a signature (image)

The screenshot displays the PDFgear application window. The main document is a form titled "VLOGA ZA PRIZNANJE IZPITA" (Statement for Exam Recognition) from the University of Primorska. The form contains personal details of the student JOŽICA NOVAK and information about the exam recognition process. A blue arrow points to the "Signature" option in the "Tools" menu. A context menu is open over a signature on the document, showing options: "Flatten", "Delete", and "Place On Multiple Page". The interface includes a top menu bar (Home, Annotate, Edit, Form, Page, Tools, Help), a toolbar with icons for various PDF operations, a thumbnail view on the left, and a chat window on the right.

# S certifikatom / With certificate

V Adobe Acrobat v levem meniju izberemo "View More" in "Use a certificate". Na novem meniju izberemo "Digitally sign" in na dokumentu s potegom označimo kje bo podpis. Izberemo certifikat in vnesemo geslo, če je potrebno.

In Adobe Acrobat, select "View More" and "Use a certificate" from the left menu. In the new menu, select "Digitally sign" and drag and drop where the signature will appear on the document. Select the certificate enter the password if required.

The screenshot illustrates the process of signing a document in Adobe Acrobat. It shows the 'Use a certificate' sub-menu with 'Digitally sign' selected (marked with a red '2.'). A separate menu shows 'Use a certificate' selected (marked with a red '1.'). The main document area displays a form with fields for name, program, and date. A dialog box titled 'Sign as "Matjaž Kljun"' is open, showing the digital signature 'Matjaž Kljun' and the date 'Date: 2024.03.01 00:08:13 +01'00''. A red '3.' is placed over the 'Sign' button in the dialog box.

## Lastno kreiran certifikat / Own created certificate

- Certifikat lahko naredimo sami v Adobe Reader-ju (slika na naslednji strani):
  - Izberemo menu (levo zgoraj)
  - Izberemo Properties
  - Izberemo Signatures
  - Kliknemo na gumb More pri Identities & Trusted certificates
  - Kliknemo na Add ID
  - Izberemo "A new digital ID I want to create now" in kliknemo next
  - Izberemo "New PKCS#12 digital ID file" in kliknemo Next.
  - Izpolnimo podatke in kliknemo Next.
  - Izberemo geslo in kliknemo Next.
  - Certifikat je ustvarjen. Datoteko podpišemo kot smo pokazali prej.
- **Ni nobene avtoritete, ki bi garantirala veljavnost certificata, kot je to SIGEN-CA!!!**
- We can create the certificate ourself in Adobe Reader (see image on the next page):
  - Select menu (top left)
  - Select Properties
  - Select Signatures
  - Click on the More button under Identities & Trusted certificates
  - Click on Add ID
  - Select "A new digital ID I want to create now" and click Next
  - Select "New PKCS#12 digital ID file" and click Next.
  - Fill in your details and click Next.
  - Select a password and click Next.
  - The certificate is created. Sign the file as shown earlier.
- **There is no authority that guarantees the validity of a certificate like SIGEN-CA!!!**

1.

**All tools**

- Export a PDF
- Edit a PDF
- Create a PDF
- Combine files
- Organize pages
- AI Assistant (beta)
- Generative summary (beta)
- Request e-signatures
- Scan & OCR
- Protect a PDF
- Redact a PDF
- Compress a PDF
- Prepare a form

View more

Convert, edit and e-sign forms & agreements.

**Free 7-day trial**



2.

Preferences

Categories:

- Commenting
- Documents
- Full Screen
- General
- Page Display
- Accessibility
- Adobe Online Services
- Email Accounts
- Forms
- Identity
- Internet
- JavaScript
- Language
- Measuring (2D)
- Measuring (3D)
- Measuring (Geo)
- Multimedia & 3D
- Multimedia (legacy)
- Multimedia Trust (legacy)
- Reading
- Reviewing
- Search
- Security
- Security (Enhanced)
- Signatures
- Spelling
- Tracker
- Trust Manager
- Units

Digital Signatures

Creation & Appearance

- Control options for signature creation
- Set the appearance of signatures within a document

Verification

- Control how and when signatures are verified

Identities & Trusted Certificates

- Create and manage identities for signing
- Manage credentials used to trust documents

Document Timestamping

Digital ID and Trusted Certificate Settings

- Digital IDs
  - Roaming ID Accounts
  - Digital ID Files
    - MatjažKljun.pfx
  - Windows Digital IDs
  - PKCS#11 Modules and Tokens
  - Trusted Certificates

3.

4.

Add Digital ID

Add or create a digital ID to sign and encrypt documents. The certificate that comes with your digital ID is sent to others so that they can verify your signature. Add or create a digital ID using:

- My existing digital ID from:
  - A file
  - A roaming digital ID accessed via a server
  - A device connected to this computer
- A new digital ID I want to create now**

Cancel < Back Next >

5.

IZPOLNI IZVAJALEC/KA PREDMATA

Izvajalec/ka predmeta: izr. pro...

učnega načrta proučil vlogo za

Izpit iz predmeta (naziv predm...

Valid from: 2024/03/01 00:04:16 +01'00'

Valid to: 2029/03/02 00:04:16 +01'00'

Intended usage: Digital Signature, Encrypt Document, Key Agreement

# Ostala orodja, ki podpirajo certifikate

## Odprta koda:

- **JSignPDF** – napisan v Javi.
- **Okular** (Linux/KDE)
- ...

## Komercialna programska oprema

- **Foxit PDF Editor**
- **Nitro PDF Pro**
- ...

## Spletna orodja za podpisovanje PDF

- **DigiSigner**
- **DocuSign** - rešitev za elektronsko in digitalno podpisovanje na ravni podjetja.
- **SignRequest**
- ...

## Open source:

- **JSignPDF** – Java-based, opensource.
- **Okular** (Linux/KDE) – Supports digital signing.

## Commercial Software

- **Foxit PDF Editor**
- **Nitro PDF Pro**
- ...

## Online PDF Signing Tools

- **DigiSigner**
- **DocuSign** – Enterprise-level electronic and digital signature solution.
- **SignRequest**
- ...



Digitally signed by your common name here  
 DN: your distinguished name here  
 Reason: your signing reason here  
 Location: your signing location here  
 Date: 2014-09-09 10:55:50

Jane Doe



Digitally signed by Jane Doe  
 Date: 2020.05.19 12:23:05 -05'00'



Nathan Nitro

Digitally signed by Nathan Nitro  
 DN: cn=Nathan Nitro, gn=Nathan Nitro, c=AU, Australia, l=AU, Australia, o=Customer Experience, ou=Customer Experience, e=nathan.nitro@outlook.com  
 Reason: I am the author of this document  
 Location:  
 Date: 2022-05-10 12:50+10:00

Digitally Signed by Pablo Marcos the 25/03/2021  
 Email: me@example.com  
 Reason: Sample Signed Document



Harlean

Digital Signer: Harlean  
 DN: C=US, E=jiangxiaoheng@foxmail.com, OU="", O="", CN=Harlean  
 Date: 2020.01.09 12:00:09 +08:00

Signed by:

*Sample Signature*

Signer Name: Sample Signature  
 Signing Reason: I approve this document  
 Signing Time: 10/12/2021 | 9:27:36 PM PDT

7CE73AB2FC2E4D9DA6568C1BE1BA6354

John Doe

Digitally signed by James Bond  
 DN: cn=James Bond  
 Reason: I'm the author  
 Location: New York  
 Date: 2011.07.24 19:54:03 +0200

# Razdeljevanje in združevanje PDF

---

## Split and merge PDF

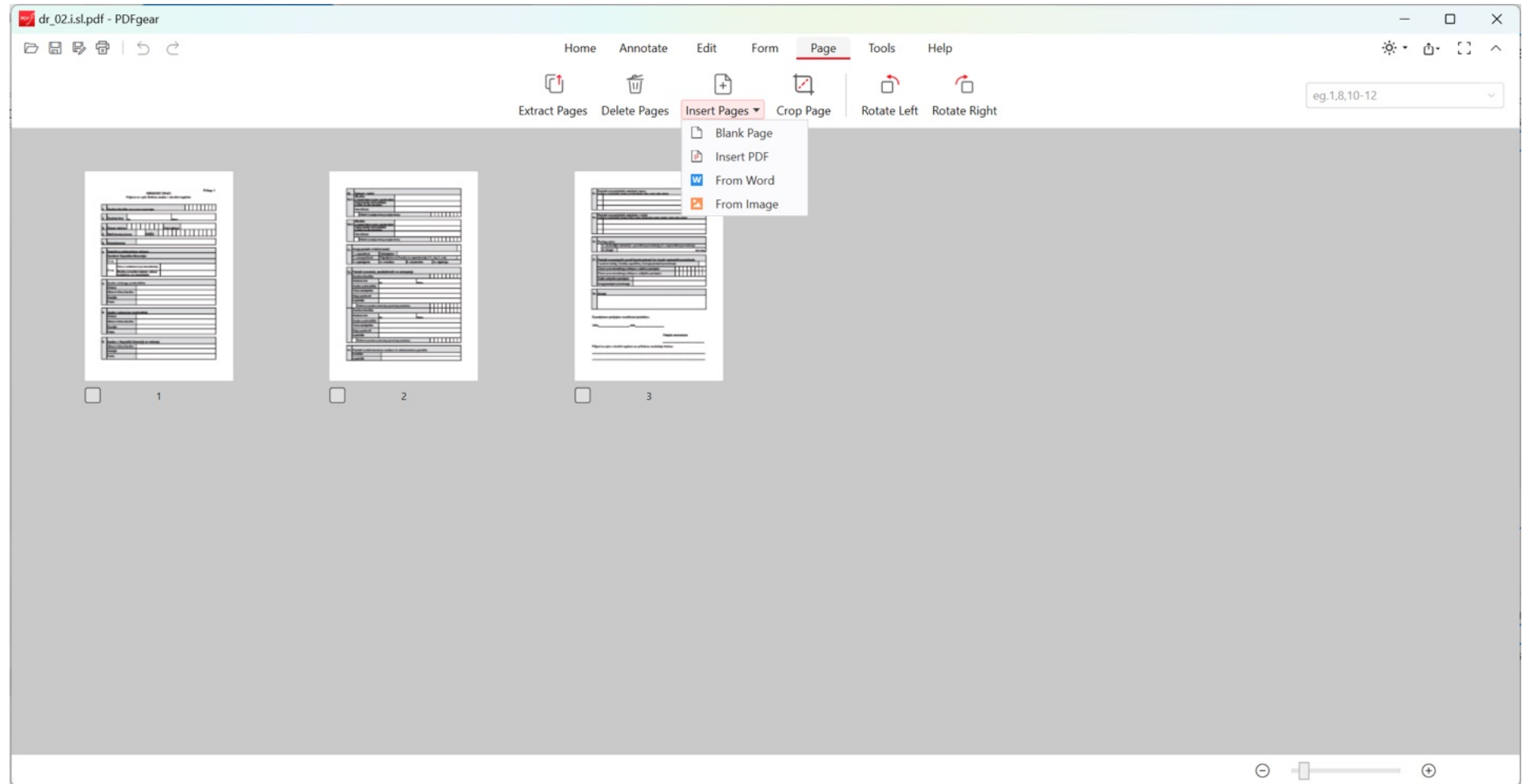
# Samodejno / Automatically

The screenshot shows the 'PDFsam Basic - Home' window. On the left is a sidebar with a 'Home' button and a 'TOOLS' list: Extract, Split, Split by bookmarks, Split by size, Alternate Mix, Insert pages multiple times, Merge, Rotate, Workspace, Logs, News, About, Settings, and Exit. The main area is titled 'Split tools' and contains four cards: 'Extract' (Extract pages from PDF documents), 'Split' (Split a PDF document at the given page numbers), 'Split by bookmarks' (Split a PDF document at bookmarked pages by specifying a bookmark level), and 'Split by size' (Split a PDF document in files of the given size (roughly)). Below this is the 'Merge tools' section with 'Alternate Mix' (Merge two or more PDF documents taking pages alternately in natural or reverse order) and 'Insert pages multiple times' (Insert the pages of a PDF document A into another document B, repeating it after a certain number of pages, resulting in B1 A B2 A B3 A etc.). At the bottom is the 'Other tools' section with 'Rotate'. A 'New version available' notification is present at the bottom right, stating 'PDFsam 5.2.2 is available for download' with a 'Download' button.

The screenshot shows the 'PDFsam Basic - Split' window. It features a 'Select PDF' button and a 'Clear' dropdown at the top. Below is a search bar. The 'TOOLS' sidebar on the left is identical to the home screen, with 'Split' highlighted. The main area is titled 'Split settings' and includes: 'Split after' (radio button selected) with a dropdown menu set to 'Every page'; 'Split after the following page numbers' (radio button unselected) with a text input 'Page numbers to split at (n1,n2,n3,...)'; and 'Split by every "n" pages' (radio button unselected) with a text input 'Number of pages'. The 'Output settings' section includes a 'Select a directory' input with a 'Browse' button and an 'Overwrite if already exists' checkbox. Below is a 'Show advanced settings' section. The 'File names settings' section includes a 'Generated PDF documents name prefix' input with the value 'PDFsam\_'. A blue 'Run' button is located at the bottom left. A 'New version available' notification is present at the bottom right, identical to the one in the home screen.



# Ročno / Manually



# Popravljanje besedila

---

## Editing text

## Načini / Ways of doing it

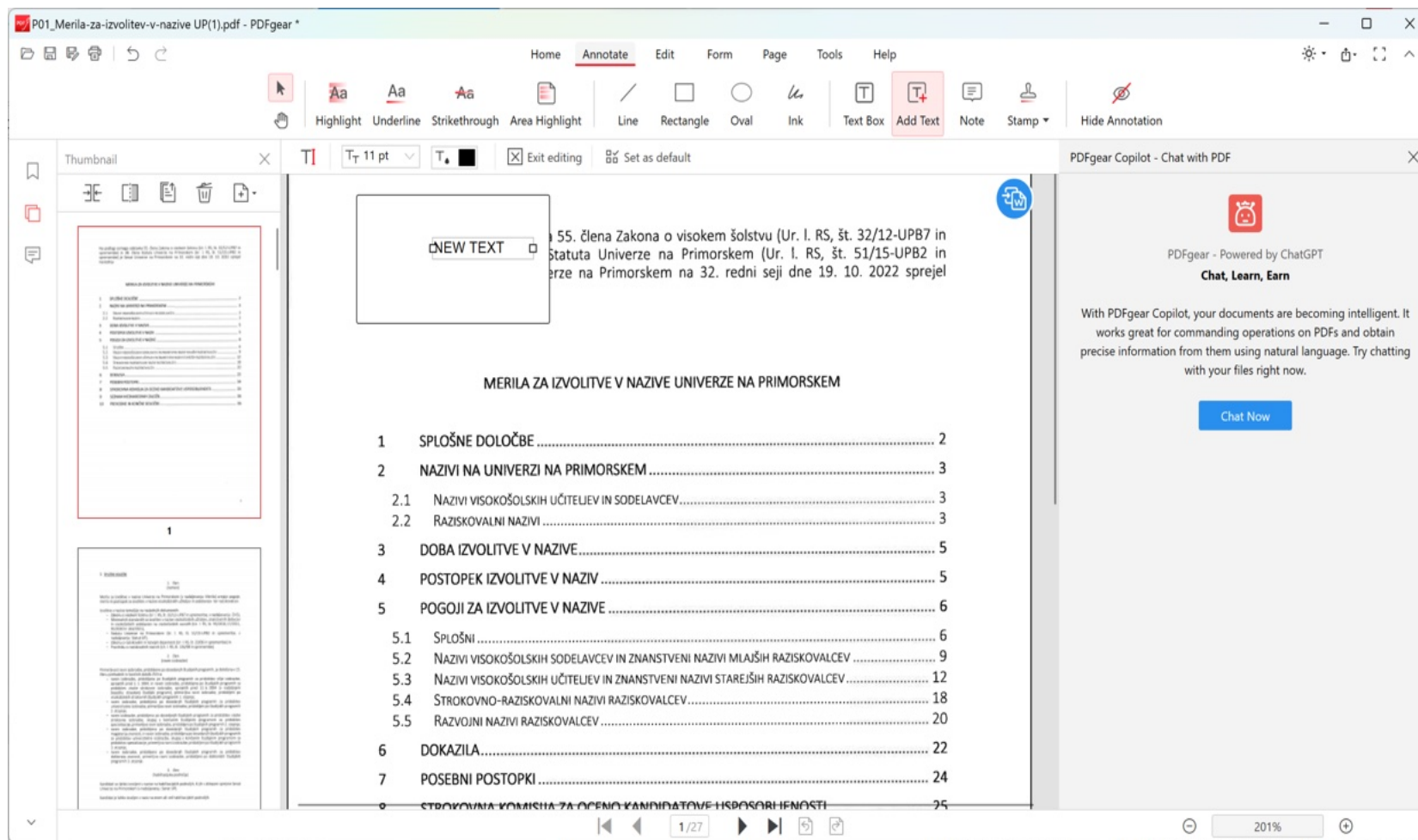
- Pretvorba v Word, Excel, Powerpoint, Gdocs
- Prekrivanje besedila s belimi pravokotniki in pisanje čez le-te.
- LibreOffice Draw
  - Najboljše rezultate dobimo, če je dokument bil narejen (Shrani kot) z MS Word-om ali LibreOffice Writerjem ali drugim urejevalnikom besedil.
  - Nezanosljivi rezultati pri skeniranih dokumentih.
- Priporočilo:
  - Če imamo izvorno datoteko napisano v oblikovalniku besedila, popravljamo le-to in so ponovno shranimo v PDF.
  - Zaposimo za izvorno datoteko.
- Conversion to Word, Excel, Powerpoint. Gdpcs
- Overlaying text with white rectangles and writing over them.
- LibreOffice Draw
  - Best results are obtained if the document has been made (Save As) with MS Word or LibreOffice Writer or another word processor.
  - Unreliable results for scanned documents.
- Recommendation:
  - If we have a source file written in a word processor, edit it and save it again as a PDF.
  - Request the source file.

# Prekrivanje besedila s pravokotniki / Covering text with rectangles

Besedilo pod belim pravokotnikom je še vedno mogoče izbrati, vendar ni vidno! Je samo prekrito. Novo besedilo lahko postavimo nad beli pravokotnik.

---

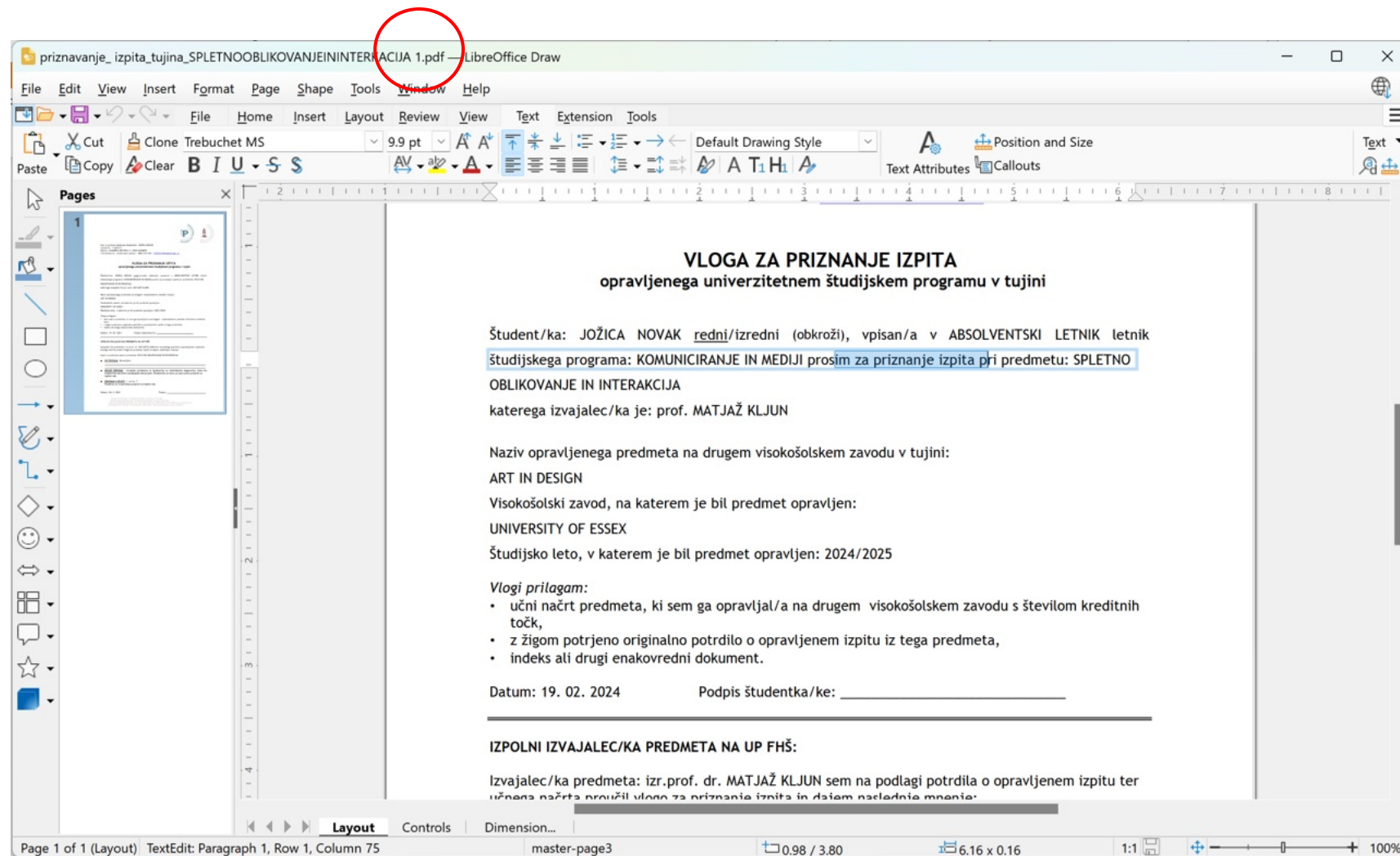
The text under the white rectangle can still be selected but it's not visible! It is just covered. We can place the new text over the white rectangle.



# PDF dokument narejen v Wordu / PDF document made in Word

LibreOffice Draw nam omogoča popraviljanje besedila vrstico za vrstico kot ločene dele besedila.

LibreOffice Draw enables to edit the text line by line, which are considered separate blocks of text.

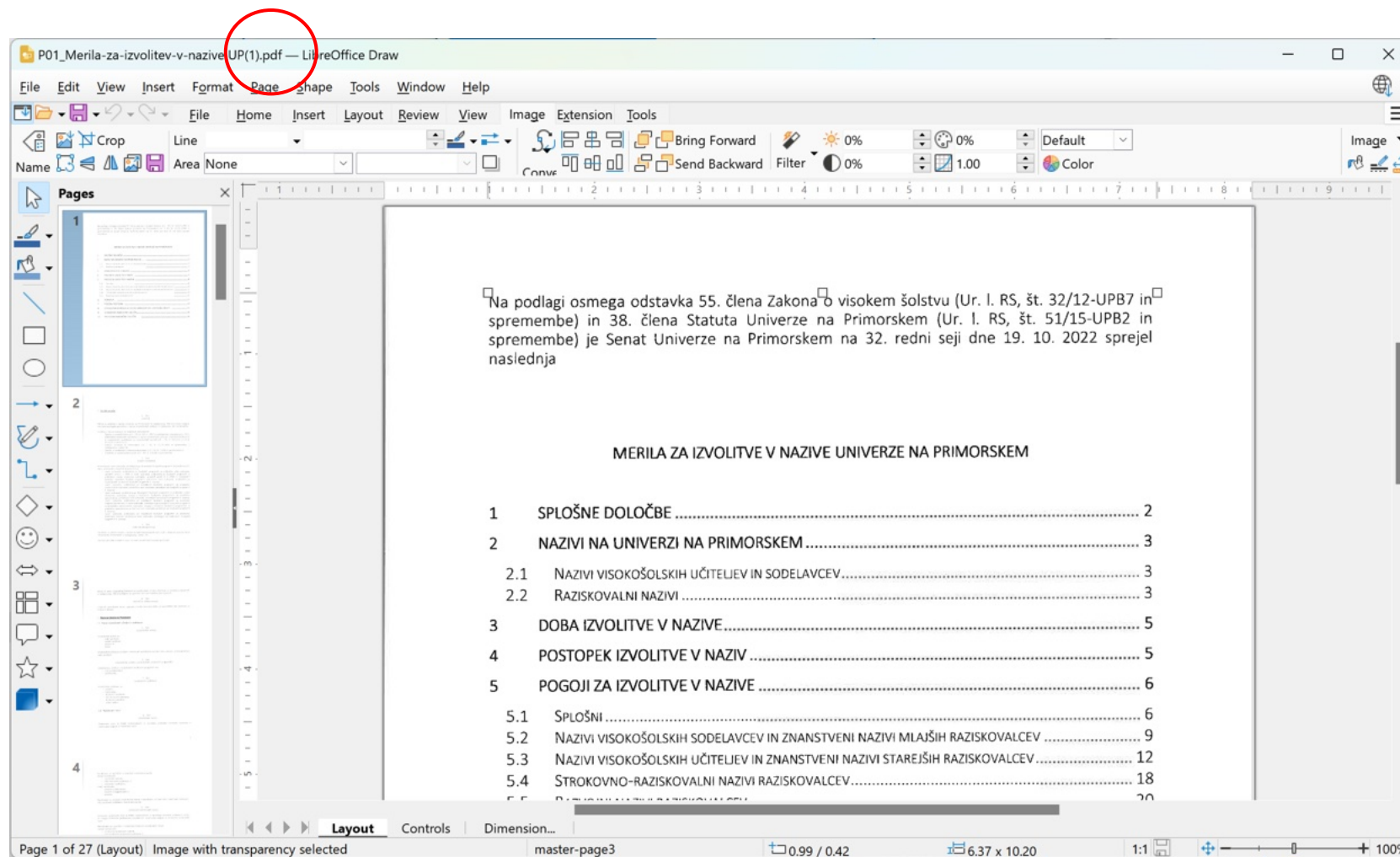


# Skeniran PDF dokument / Scanned PDF document

LibreOffice Draw nam slike še vedno odpre kot slike.

---

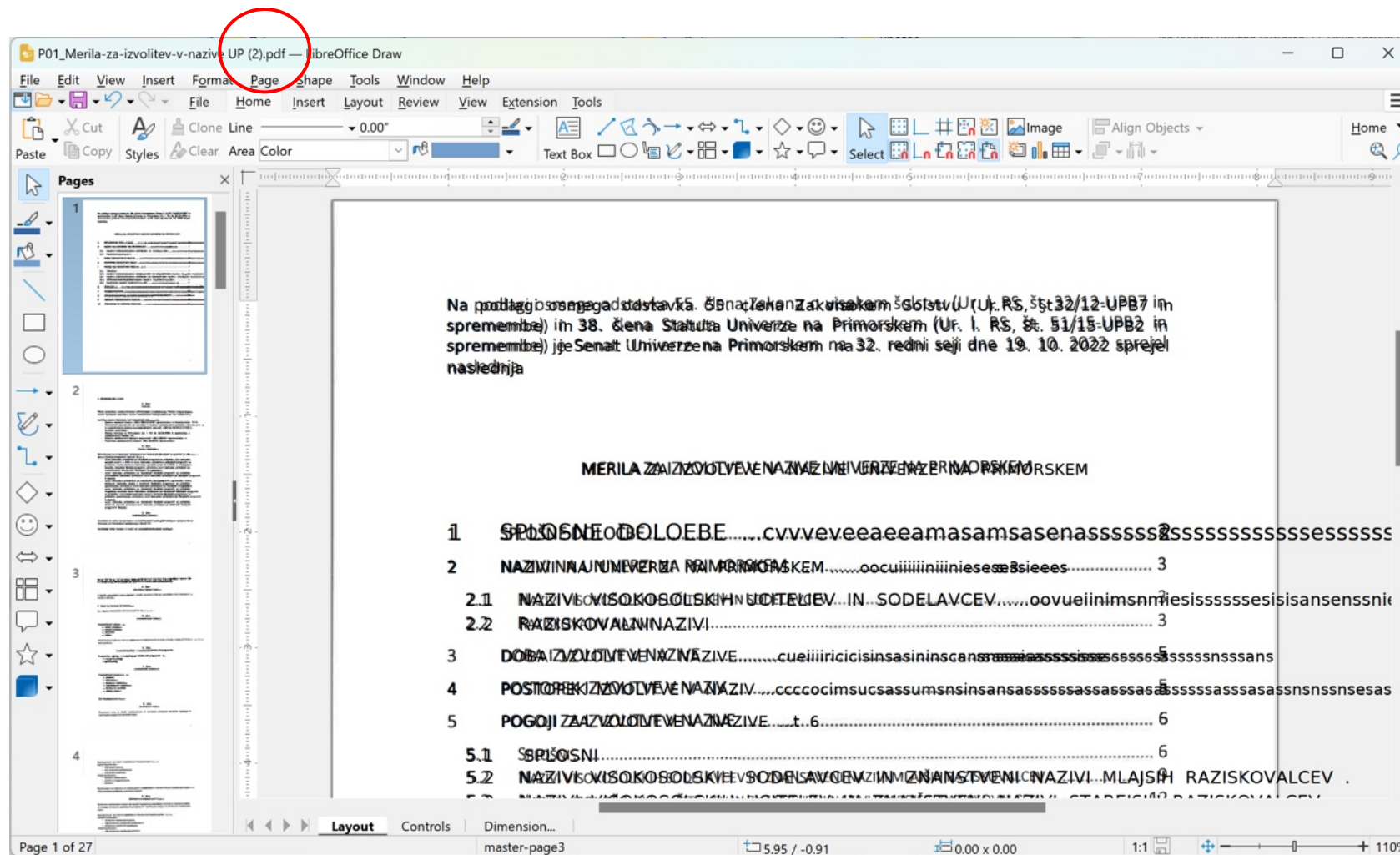
LibreOffice Draw treats images as images and not text.



# Skeniran in OCRjan PDF / Scanned and OCRred PDF

LibreOffice Draw nam slike še vedno odpre kot slike in doda besedilo čez.

LibreOffice Draw treats images as images and not text, but adds text over as well.



# Ostala orodja / Other tools

## Odprtokodne rešitve

- PDFEdit
- LibreOffice Draw - Urejanje besedila, slik in obrazcev.
- Okular - Podpira osnovno urejanje besedila.

## Spletna orodja za urejanje PDF

- Sejda PDF Editor - Spletni, omogoča neposredno urejanje besedila.
- PDFescape - omejene možnosti brezplačnega urejanja besedila.

## Komercialna programska oprema

- Adobe Acrobat Pro DC - polno funkcionalno urejanje in pretvorba besedila.
- Foxit PDF Editor - napredne funkcije za urejanje besedila in OCR.
- Nitro PDF Pro - Urejanje, podpisovanje in pretvorba datotek PDF s spreminjanjem besedila.
- PDFelement by Wondershare - Intuitiven vmesnik z robustnimi orodji za urejanje besedila.
- PDF Studio by Qoppa - Podpira urejanje celotnega besedila in OC

## Open-Source Solutions

- PDFEdit
- LibreOffice Draw – Edit text, images, and forms.
- Okular – Supports basic text editing.

## Online PDF Editing Tools

- Sejda PDF Editor – Web-based, allows direct text editing.
- PDFescape – Limited free text editing capabilities.

## Commercial Software

- Adobe Acrobat Pro DC – Full-featured text editing and conversion.
- Foxit PDF Editor – Advanced text editing and OCR features.
- Nitro PDF Pro – Edit, sign, and convert PDFs with text modification.
- PDFelement by Wondershare – Intuitive interface with robust text editing tools.
- PDF Studio by Qoppa – Supports full text editing and OCR.



PDF in UI

---

PDF and AI

# Kaj lahko naredimo / What can be done

UI predelai vsebino PDF in odgovori na naša vprašanja.

- **Izveček in povzemanje besedila** - Izpiše ključne ugotovitve.
- **OCR** - Pretvori skenirane datoteke PDF v besedilo, ki ga je mogoče urejati in po njem iskati.
- **Analiza vsebine** - Analizira strukturo dokumentov, izpiše ključne informacije.
- **Avtomatizirano prevajanje** - Prevede besedilo, pri čemer lahko celo ohrani celovitost postavitve.
- **Iskanje podatkov** - Prepozna tabele, strukturirane podatke in izlušči informacije za nadaljnjo obdelavo.
- **Ustvarjanje in dopolnjevanje besedila** - Ustvari manjkajoče dele, izboljša oblikovanje dokumenta in samodejno dopolni besedilo.
- **Pogovorno iskanje in druga dejanja** - Poizvedovanje po vsebini PDF in druga dejanja (zaščita z geslom, prikaz PDF kot diapozitiv ...).
- **Druge možnosti** - Zaščita z geslom, prikaz PDF kot prosojnice ...

Odvisno od orodja

AI “scans” the content of the PDF and answers our questions.

- **Text Extraction & Summarization** – Generate key takeaways.
- **OCR (Optical Character Recognition)** – Convert scanned PDFs into editable and searchable text.
- **Content Analysis** – Analyse document structures, detect key information.
- **Automated Translations** – Translate text while maintaining layout integrity.
- **Data Extraction** – Identify tables, structured data, and extract information for further processing.
- **Text Generation & Completion** – Generate missing sections, improve document formatting, and auto-complete text.
- **Conversational Search** – Query PDF content
- **Other possibilities** – Password protect, display PDF as slides ...

Depends on the tool

Home Annotate Edit Form Page Tools Help

Print 100% Single Double Continuous Auto Scroll Slide Show Screenshot OCR Background Find

the remains of the satellite into the ocean far from human settlements<sup>1</sup> or (2) continue TRMM operation until the fuel runs out in 2010 or 2011 and accept the added risk of an uncontrolled reentry because of the operational and scientific benefits of doing so.

### THE ROLE OF THE NATIONAL ACADEMIES

In August 2004 NASA Administrator Sean O’Keefe requested that the National Academies provide advice on the anticipated scientific and operations contributions from extending TRMM beyond 2004. The charge to the Committee on the Future of the Tropical Rainfall Measuring Mission in the first phase of its work was to conduct a workshop and prepare an interim report to be delivered in December 2004 on how best to use the remaining TRMM spacecraft life. The Academies were able to begin the study process in October 2004, and immediately assembled the committee and planned an information-gathering workshop. The committee and workshop participants were asked to consider

- scientific and research contributions of TRMM to date and those expected if TRMM is continued;
- operational contributions of TRMM to date and those expected if TRMM is continued;
- assessment of expected benefits of continuing TRMM operation until (1) fuel is depleted to the level needed for a controlled reentry (around December 2005), and (2) all fuel is depleted (estimated to be 2010-2011).

A second phase of the committee’s work will focus on needs for satellite based measurements of tropical rainfall beyond TRMM (see Appendix B).

The committee hosted its phase I workshop in Washington, D.C., on No

#### PDFgear Copilot - Chat with PDF

PDFgear Copilot is an AI tool designed to assist you in summarizing PDF, searching PDF content, editing PDF, converting PDF, managing PDF, and more using natural language commands.

Examples:

Summarize this PDF file.

Compress this PDF using a high compression level.

Display this PDF in slide mode.

Add a password to protect this file.

Zoom in this PDF view.

Click the examples above to give them a try, or input your command.

[Clear Chat History](#) [Export](#)

Make a detailed summary of this article.

Send

# NotebookLM

The screenshot displays the NotebookLM interface for a document titled "AI Code Generation in Programming Education: A Review". The interface is divided into three main panels: Sources, Chat, and Studio.

- Sources Panel:** Contains a list of three PDF sources: "2310.05292v5.pdf", "2401.11314v2.pdf", and "3626252.3630958.pdf". Each source has a checkmark indicating it is selected.
- Chat Panel:** Features a chat window with a robot icon. The title "AI Code Generation in Programming Education: A Review" is displayed. Below the title, it says "3 sources". The chat content includes a paragraph: "The first source presents **HypoCompass**, an LLM-based tutor designed to help students debug code by interacting with an agent that provides buggy snippets, requests test cases and explanations, and offers feedback. It also discusses **CodeAid**, an LLM-powered programming assistant evaluated in a large course, which offers various features to aid learning without directly giving code solutions. The paper details the **design, deployment, and student/educator perspectives** on CodeAid, identifying key design considerations for future AI educational tools. The second source is a **systematic literature review** examining the use of LLM-based code generation models in programming education, highlighting their applications in teaching and learning, model characteristics, performance evaluation, and important considerations like **academic integrity and potential risks**." Below the chat content are buttons for "Save to note", "Add note", "Audio Overview", and "Briefing doc". At the bottom of the chat panel is a text input field with "Start typing..." and a "3 sources" indicator.
- Studio Panel:** Contains an "Audio Overview" section with a "Deep Dive conversation" card (Two hosts (English only)) and buttons for "Customize" and "Generate". Below this is a "Notes" section with an "Add note" button and four note templates: "Study guide", "Briefing doc", "FAQ", and "Timeline". At the bottom of the Studio panel, it says "Saved notes will appear here" and "Save a chat message to create a new note, or click Add note above."

At the bottom of the interface, a disclaimer reads: "NotebookLM can be inaccurate; please double check its responses."

# Konec / The End

Matjaž Kljun

[matjaz.kljun@upr.si](mailto:matjaz.kljun@upr.si)

@\_mkljun\_

Vprašanja?

Matjaž Kljun

[matjaz.kljun@upr.si](mailto:matjaz.kljun@upr.si)

@\_mkljun\_

Questions?

