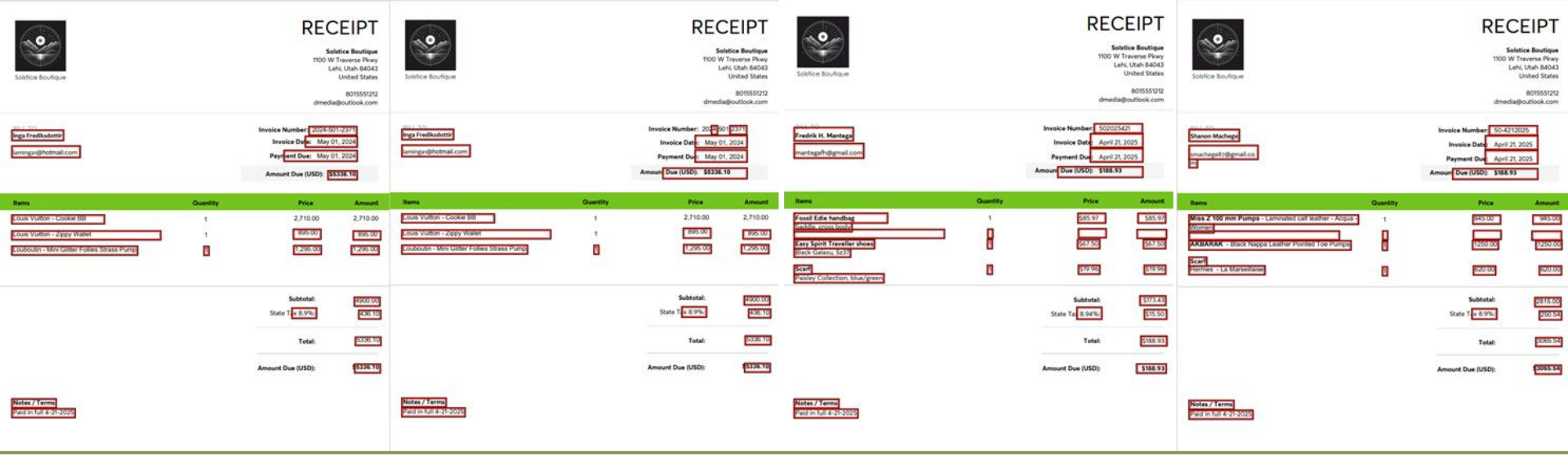


PDF Days Europe 2025

In Defense of the Incremental Save

PDF Forensics best friend?

Cherie Ekholm | Product Strategy Lead | Veriskll

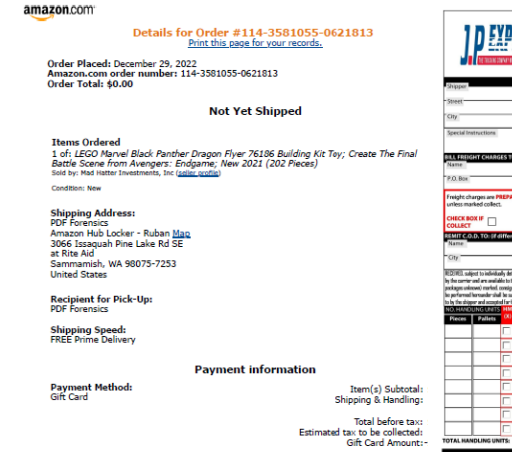
Quick bio

- Product Strategy Lead at Verisk focusing on Digital Media Forensics (Images and PDF)
- Co-Chair of the PDF Forensics LWG
- Former Project Leader for ISO 14289 and 32000
- 8+ years at Microsoft as a Software Lead in the Office Division focusing on international standards, including PDF, ODF, OOXML, WCAG, ATAG, and UUAG (and other Microsoft roles before that)
- Author, editor, publisher
- Want to hear about my dogs?

What is PDF Forensics

- Involves the examination of PDF files or groups of PDF files to detect possible anomalies used for deceptive purposes
- Files to examine may include invoices, drivers licenses, passports, receipts, reports, or any other kind of document that might be shared from one person or entity to another
- In other arenas it may also include examining PDF files for software security risks
- Common uses for forensics includes detecting fraudulent ID by border agencies or other government agencies, finding altered invoices, sales receipts, or other document for use in insurance claims, banking applications, or court cases, or tracing provenance of files used in almost any situation

Just a few types of documents for forensic analysis



amazon.com

Details for Order #114-3581055-0621813
Print this page for your records.

Order Placed: December 29, 2022
Amazon.com order number: 114-3581055-0621813
Order Total: $0.00

Not Yet Shipped

Items Ordered
1 of: LEGO Marvel Black Panther Dragon Flyer 76186 Building Kit Toy; Create The Final Battle Scene from Avengers: Endgame; New 2021 (202 Pieces)
Sold by: Mad Hatter Investments, Inc (seller profile)
Condition: New

Shipping Address:
PDF Forensics
Amazon Hub Locker - Ruban Map
3066 Issaquah Pine Lake Rd SE
at Rite Aid
Sammamish, WA 98075-7253
United States

Recipient for Pick-Up:
PDF Forensics

Shipping Speed:
FREE Prime Delivery

Payment information

Payment Method:
Gift Card

Item(s) Subtotal:
Shipping & Handling:

Total before tax:
Estimated tax to be collected:
Gift Card Amount:-

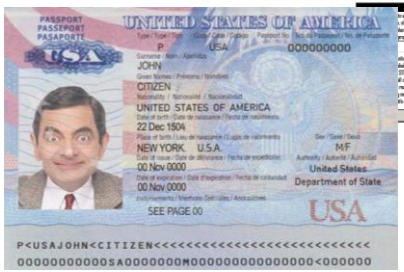
J.P. EXPRESS

Affix
PRO Label
Here

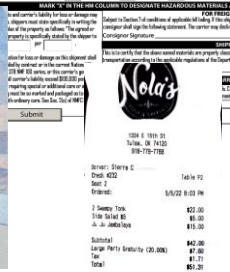
SHIPPER (from)
CONSIGNEE (to)
BILL FREIGHT CHARGES TO: (if different than above)
C. O. D.
AMOUNT
REMIT C.O.D. TO: (if different from shipper above)
DESCRIPTION OF ARTICLES, KIND OF PACKAGE, SPECIAL MARKS AND EXCEPTIONS
TOTAL HANDLING UNITS:

WA USA WASHINGTON DRIVER LICENSE
DONOR
4d LIC# SAMPJA185800
1 SAMPLE
2 JOHN A
3 DOB 01/06/1958
4b EXP 08/12/2020
4a ISS 01/06/2015
8 123 STREET ADDRESS
YOUR CITY WA 99999-0000
15 SEX M
16 HGT 5'-08"
17 WGT 165 lb
12 RESTRICTIONS NONE
18 EYES DIC
9a END NONE
Veteran
SAMPLE

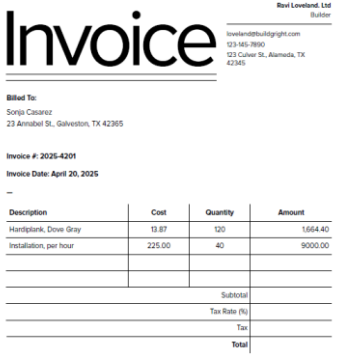
Instrument inspection report



UNITED STATES OF AMERICA
PASSPORT
PASSEPORT
PASAPORTE
USA
P
000000000
JOHN
CITIZEN
UNITED STATES OF AMERICA
22 Dec 1504
NEW YORK, U.S.A.
M/F
00 Nov 0000
00 Nov 0000
United States
Department of State
SEE PAGE 00
USA
P<USAJOHN<CITIZEN<<<<<<<<<<<<<<<<<<<<<<<<<<
0000000000USA0000000M0000000000000000<0000000



Nola's
Submit
Server: Sierra C
Check: 4232
Seat: 2
Table: P2
Ordered: 5/5/22 8:03 PM
2 Sweaty Tonk $22.00
Tribe Salad XS $5.00
Jambalaya $15.00
Subtotal $42.00
Large Party Gratuity (20.00%) $7.80
Tax $1.71
Total $51.51



Invoice

Ravi Loveland, Ltd
Builder
loveland@buildright.com
123-145-7890
123 Culver St., Alameda, TX 42345

Billed To:
Sonja Casarez
23 Annabel St., Galveston, TX 42365

Invoice #: 2025-4201
Invoice Date: April 20, 2025

Description	Cost	Quantity	Amount
Hardiplank, Dove Gray	13.87	120	1,664.40
Installation, per hour	225.00	40	9000.00
		Subtotal	
		Tax Rate (%)	
		Tax	
		Total	

Work order
Complete south-facing exterior of home at 23 Annabel St., Galveston, TX 42365

Condition: Good
Fines: $35

Saxophones
Video provides a powerful way to help you prove your point. When you click Online Video, you can paste in the embed code for the video you want to add.

You can also type a keyword to search online for the video that best fits your document. To make your document look professionally produced, Word provides header, footer, cover page, and text box designs that complement each other.

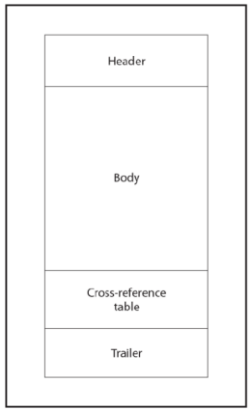
Alto
Video provides a powerful way to help you prove your point. When you click Online Video, you can paste

TRAVELERS Minnesota Automobile Insurance Identification Card
This card must be carried in the vehicle at all times as evidence of insurance.
Year 2007 Make VOLKS Model RABBIT
Vehicle identification number (VIN) WV
Policy number 99
Effective date 04/11/2014
Expiration date 10/11/2014
Insured
JAMES HOLM
Company: THE STANDARD FIRE INSURANCE COMPANY
For policy questions and changes
THE INSURANCE OFFICE
218-828-1118
To report a claim
or get roadside assistance
1-800-252-4633 (24 hours x 365 days a year)
Keep this card in the vehicle at all times. See reverse side.

What’s your intent?

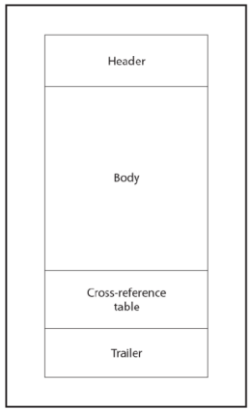
- Intent is one of the hardest things to identify and prove
- A fixed format should make it easy to see whether an alteration is fraudulent
- PDF is not a fixed format
- Did the person editing the file intend to simply edit a file they think of as a template for other documents?
- Or did the person editing the file intend something a little more sinister?
- Version history is one of the best ways to understand the intent of edits

Full save vs incremental save



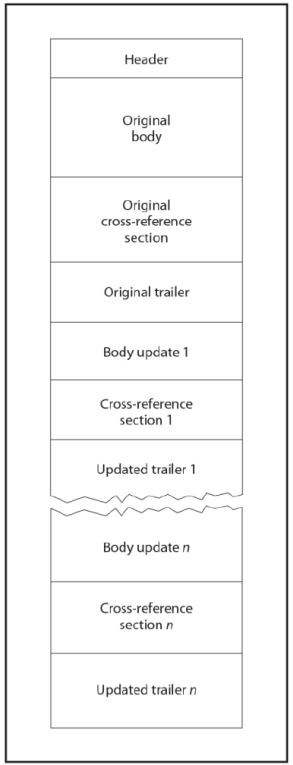
Full Save

Full save vs incremental save



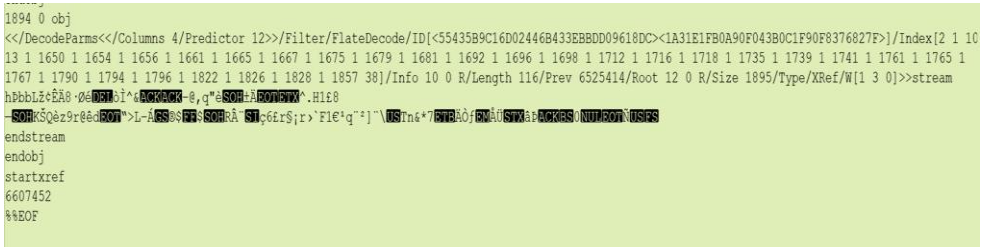
Full Save

Incremental Save



PDF file structure

- Well-defined in ISO 32000 section 7.5
- Cross-reference table vs. cross-reference stream
  - Tracks objects & generations
  - Shows changes, additions, replacements
  - Keeps pointers to changes available
- Additional trailer info
  - Prev key included when file is updated
  - Include DocumentID and incremented InstanceID



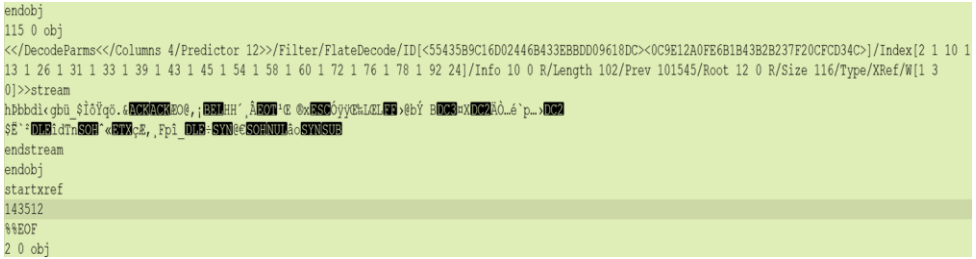
Cross reference table

- Traditional structure used in PDF versions prior to 1.5
- Stores byte offsets for each object in the file
- Enables quick access to objects by their location
- Located at the end of the file after the last object
- Used in conjunction with other information in the trailer

```
xref
0  6
0000000003 65535 f
0000000017 00000 n
0000000081 00000 n
0000000000 00007 f
0000000331 00000 n
0000000409 00000 n
```

Cross reference streams

- Introduced in PDF 1.5
- Stores cross-reference data in a stream object
- Better supports for object compression
- More efficient for large and complex PDFs
- Replaces traditional xref tables and trailer dictionaries



What about hybrid cross references?

- Please don’t
- No one else is
- 32000-2 7.5.8.4

Use cases for incremental saves

- Version History Tracking
- Collaboration Support
- Audit Trails
- Efficient Updates
- Preservation of historical data
- Signature Preservation
- Recovery Potential

Comparison of incremental vs full saves

Aspect	Incremental Saves	Full Saves
Version History	Preserved across saves	Overwritten with each save
File Size	Can grow significantly over time	Typically smaller and more efficient
Forensic Value	High – retains historical data	Low – previous data removed
Collaboration Support	Supports multi-stage edits	Requires manual versioning
Security Implications	Can expose deleted/redacted data	Safer – removes old content

How do I find incremental saves?

1. Look in the file trailer
2. Find the EOF%% markers
3. Look for the Prev key
4. Find the cross-reference table or stream
5. Walk the markers back through the file to new or changed objects

- Some tools to help
  - Text editors – Many files can be opened in a text editor and manually examined
  - Via an open source tool – many of the tools like qpdf/pikepdf allow you to crack open the file and examine the structures
  - Via a proprietary tools – Adobe Acrobat Pro, Windjack's PDFCanOpener, & others let you look into the file and find the markers

Example incremental save

```
1649 0 obj
<</DecodeParms<</Columns 4/Predictor 12>>/Filter/FlateDecode/ID[<55435B9C16D02446B433EBBDD09618DC><6B7901BB2E3E4042A07033014C20C1A7>]/Index[2 1 10 1 13 1 1510 1 1514 1 1516 1 1521 1 1525 1 1527 1 1535 1 1539 1 1541 1
1572 1 1576 1 1578 1 1595 1 1599 1 1601 1 1621 29]/Info 10 0 R/Length 110/Prev 6068199/Root 12 0 R/Size 1650/Type/XRef/W[1 3 0]>>stream
hÞbbŒ™ù‰qIn/ÓýENQSOLFFFF,Ñ@,;STH^‡STX &qRC4.#^Ë^ÉeÀ+‹ÎeßSTX2TSYNÄÔƒÉ°Ã"0„SOH f:F]CAN—)CANDèÁ¹ ÂEOT®ø‡Ì(BSw)ŒEBSvÜ
béÃ¹Y Öm€NULETXNULN}FFBf£
endstream
endobj
startxref
6126947
%%EOF
1660 0 obj
<</DecodeParms<</Columns 4/Predictor 12>>/Filter/FlateDecode/ID[<55435B9C16D02446B433EBBDD09618DC><E1A5EB6FA4BE9D4D860CC5C0AC07B8AC>]/Index[2 1 10 1 13 1 1650 11]/Info 10 0 R/Length 62/Prev 6126947/Root 12 0 R/Size 1
0]>>stream
hÞbbŒ-odb\Üþ‹éDELlÛe&ACKACKæ, Á8DC1H^3STX ÆACKYDLEWSTHHôé¸î ®8^NAKACKRÜSTXàÞACKBS0NULÛ9
d
endstream
endobj
startxref
6141236
%%EOF
1674 0 obj
<</DecodeParms<</Columns 4/Predictor 12>>/Filter/FlateDecode/ID[<55435B9C16D02446B433EBBDD09618DC><BFBD5C2EBE64B7488C41782F8E780A4A>]/Index[2 1 10 1 13 1 1650 1 1654 1 1656 1 1661 14]/Info 10 0 R/Length 76/Prev 61412
1675/Type/XRef/W[1 3 0]>>stream
hÞbbGÝ¶‰qqÀ3¦ý‡{Ÿ3100WSTX  Æ‰@B¢BELÄDC2BELDC2FF .»$^%VT$ô™€â YACKwDC2Æ0D.^¥BELç¶€X·SOHSTXFFNULaÓ
4
endstream
endobj
startxref
6162851
%%EOF
1691 0 obj
<</DecodeParms<</Columns 4/Predictor 12>>/Filter/FlateDecode/ID[<55435B9C16D02446B433EBBDD09618DC><3BF06307F0B69742AB7430CF09DACC72>]/Index[2 1 10 1 13 1 1650 1 1654 1 1656 1 1661 1 1665 1 1667 1 1675 17]/Info 10 0 R
6162851/Root 12 0 R/Size 1692/Type/XRef/W[1 3 0]>>stream
hÞbbGã6eb\ô;•éDELæPACKDC3ETXETXsBEL`œBS$Ä{@,qFS\v/ Á VT$ôa²FFî .H    CCAN^•  RFF'eEOTssA,=8·ENQÄ°
DLE·NUL+
¦
endstream
endobj
startxref
6192023
%%EOF
```

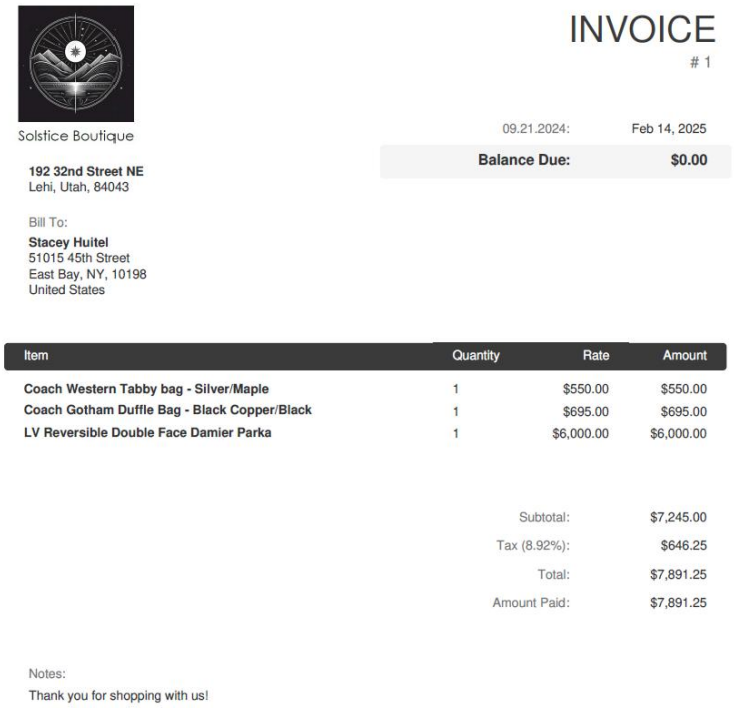

PDF association

Use tools to restore and display file history

- Proprietary - may or may not be available for licensing
  - Often created by companies that do volume forensic examination
  - Use limited to company or partner use
- Open source – tools such as PDF Resurrect
  - Good for quick and dirty examination
  - May have limitations with unnecessary hard coding or reliability
- Roll your own – 32000 gives the background to do this
  - Use one or more reliable libraries or SDKs to
    - Parse the file
    - Find the versions
    - Rebuild and present files showing previous versions

How can PDF forensics use incremental saves

- Look for changes to
  - Prices
  - Names
  - Dates
  - Account numbers
  - Descriptions
  - Insertion of signatures
  - Form fields
- Timing of changes
- Type of document
- Change in editors



Solstice Boutique

192 32nd Street NE
Lehi, Utah, 84043

Bill To:
Stacey Huitel
51015 45th Street
East Bay, NY, 10198
United States

INVOICE
1

09.21.2024: Feb 14, 2025

Balance Due: $0.00

Item	Quantity	Rate	Amount
Coach Western Tabby bag - Silver/Maple	1	$550.00	$550.00
Coach Gotham Duffle Bag - Black Copper/Black	1	$695.00	$695.00
LV Reversible Double Face Damier Parka	1	$6,000.00	$6,000.00

Subtotal:	$7,245.00
Tax (8.92%):	$646.25
Total:	$7,891.25
Amount Paid:	$7,891.25

Notes:
Thank you for shopping with us!

Workflow vs malicious edits

- **Were the changes just workflow?**
  - Adding linearization
  - Filling out forms
  - Using a file as a template

Workflow vs potentially fraudulent edits

- **Were the changes just workflow?**
  - Adding linearization
  - Filling out forms
  - Using a file as a template

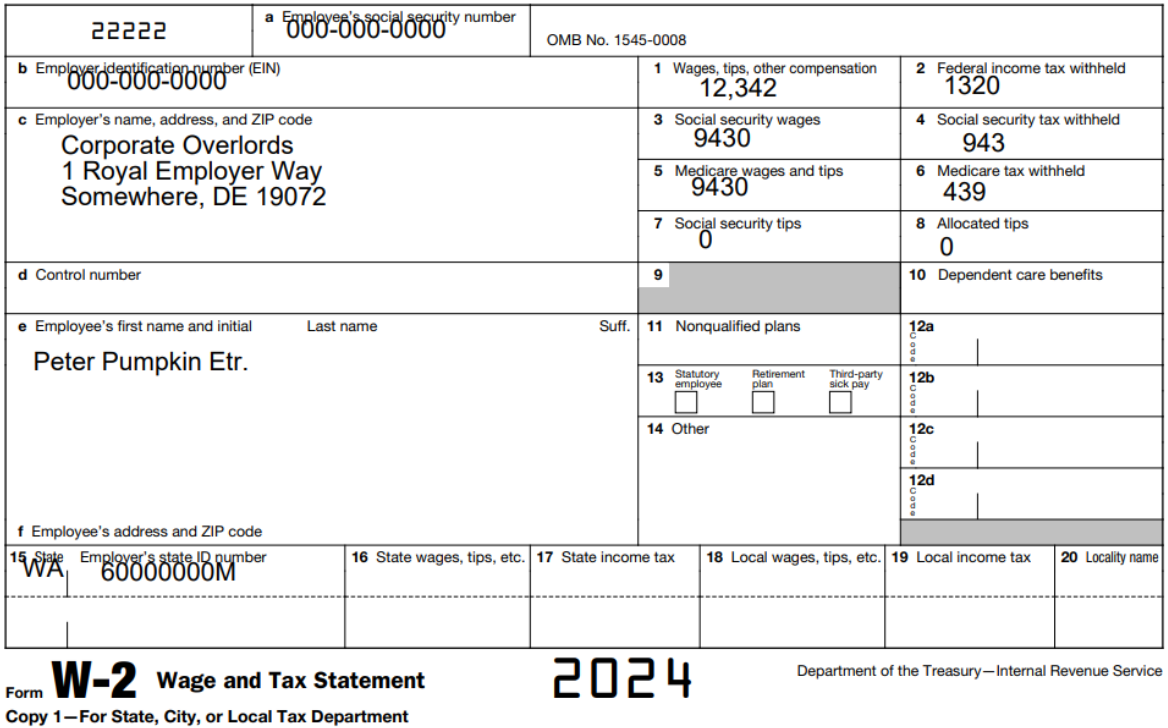
- **Or were they potentially fraudulent?**
  - Changing key information on forms
  - Changing prices or addresses or items on an invoice
  - Altering information on an ID

Linearization example


```
%PDF-1.6
%âãÏÓ
36 0 obj
<</Linearized 1/L 38267/O 38/E 29875/N 2/T 37956/H [ 487 199]>>
endobj
```

Linearization looks a lot like an incremental save. However, linearization is not a standard incremental save. It's more like a full save with some added sauce. See Annex F in ISO 32000-2.

Form example



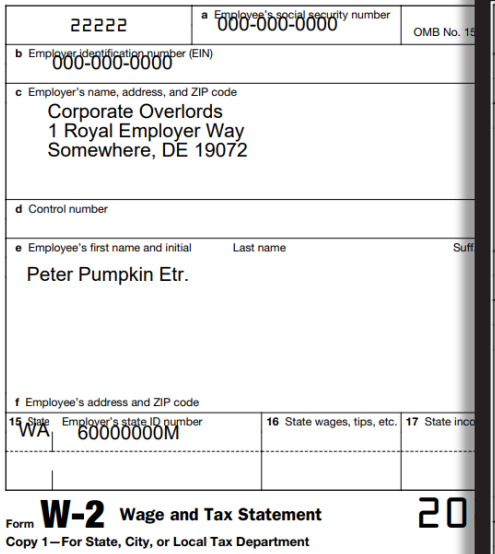
22222	a Employee's social security number 000-000-0000	OMB No. 1545-0008
b Employer identification number (EIN) 000-000-0000	1 Wages, tips, other compensation 12,342	2 Federal income tax withheld 1320
c Employer's name, address, and ZIP code Corporate Overlords 1 Royal Employer Way Somewhere, DE 19072	3 Social security wages 9430	4 Social security tax withheld 943
	5 Medicare wages and tips 9430	6 Medicare tax withheld 439
	7 Social security tips 0	8 Allocated tips 0
d Control number	9	10 Dependent care benefits
e Employee's first name and initial Last name Suff. Peter Pumpkin Etr.	11 Nonqualified plans	12a Code
	13 Statutory employee ☐ Retirement plan ☐ Third-party sick pay ☐	12b Code
	14 Other	12c Code
f Employee's address and ZIP code		12d Code

15 State	Employer's state ID number	16 State wages, tips, etc.	17 State income tax	18 Local wages, tips, etc.	19 Local income tax	20 Locality name
WA	60000000M					

Form **W-2** **Wage and Tax Statement** 2024 Department of the Treasury—Internal Revenue Service

Copy 1—For State, City, or Local Tax Department

Form example



22222

a Employee's social security number 000-000-0000

OMB No. 15

b Employer identification number (EIN) 000-000-0000

c Employer's name, address, and ZIP code
Corporate Overlords
1 Royal Employer Way
Somewhere, DE 19072

d Control number

e Employee's first name and initial Last name Suff
Peter Pumpkin Etr.

f Employee's address and ZIP code

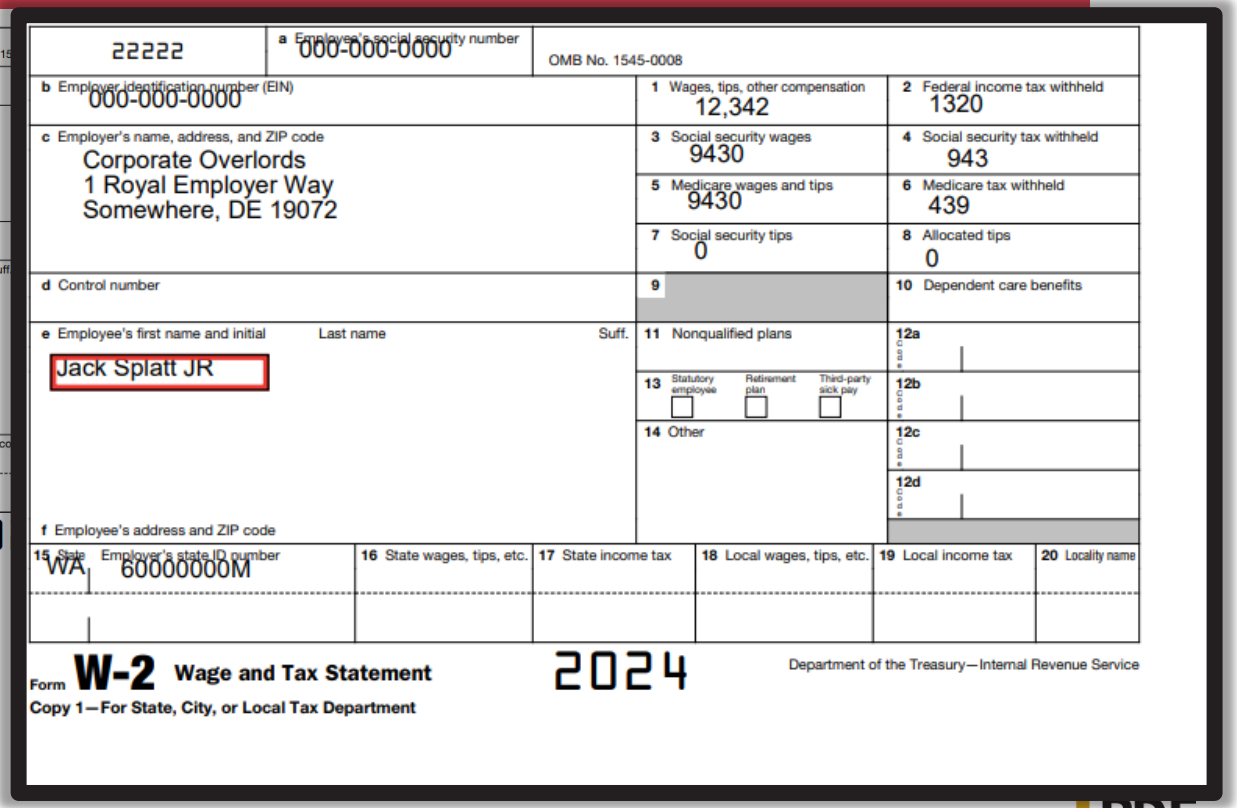
15 State Employer's state ID number
WA 60000000M

16 State wages, tips, etc.

17 State inco

Form **W-2** Wage and Tax Statement 20

Copy 1—For State, City, or Local Tax Department



22222

a Employee's social security number 000-000-0000

OMB No. 1545-0008

b Employer identification number (EIN) 000-000-0000

c Employer's name, address, and ZIP code
Corporate Overlords
1 Royal Employer Way
Somewhere, DE 19072

1 Wages, tips, other compensation 12,342

2 Federal income tax withheld 1320

3 Social security wages 9430

4 Social security tax withheld 943

5 Medicare wages and tips 9430

6 Medicare tax withheld 439

7 Social security tips 0

8 Allocated tips 0

d Control number

9

10 Dependent care benefits

e Employee's first name and initial Last name Suff.
Jack Splatt JR

11 Nonqualified plans

12a

13 Statutory employee Retirement plan Third-party sick pay

12b

14 Other

12c

12d

f Employee's address and ZIP code

15 State Employer's state ID number
WA 60000000M

16 State wages, tips, etc.

17 State income tax

18 Local wages, tips, etc.

19 Local income tax

20 Locality name

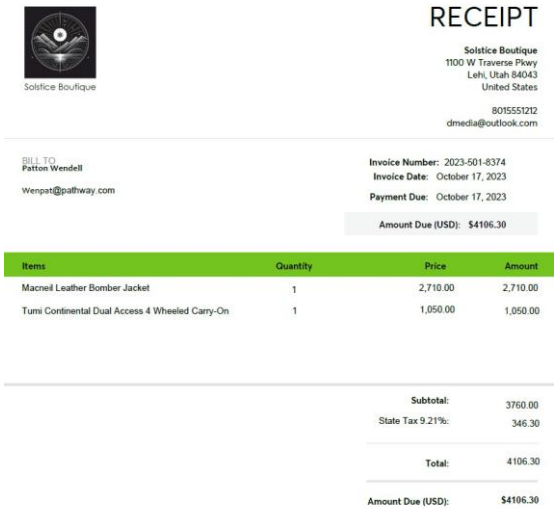
Form **W-2** Wage and Tax Statement 2024

Department of the Treasury—Internal Revenue Service

Copy 1—For State, City, or Local Tax Department

Hmm … Peter appears to have borrowed Jack's IRS form. Why?

Was this file edited to commit fraud?



Solstice Boutique

RECEIPT

Solstice Boutique
1100 W Traverse Pkwy
Lehi, Utah 84043
United States

8015551212
dmedia@outlook.com

BILL TO
Patton Wendell

Wenpat@pathway.com

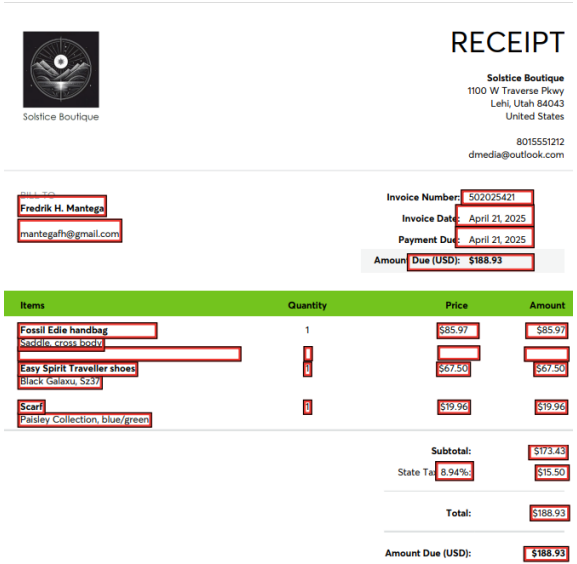
Invoice Number: 2023-501-8374
Invoice Date: October 17, 2023
Payment Due: October 17, 2023
Amount Due (USD): $4106.30

Items	Quantity	Price	Amount
Macneil Leather Bomber Jacket	1	2,710.00	2,710.00
Tumi Continental Dual Access 4 Wheeled Carry-On	1	1,050.00	1,050.00

Subtotal:	3760.00
State Tax 9.21%:	346.30
Total:	4106.30
Amount Due (USD):	$4106.30

On the left, the document as it currently appears

On the right, an earlier version of the file, with changes for customer, items, prices, etc.



Solstice Boutique

RECEIPT

Solstice Boutique
1100 W Traverse Pkwy
Lehi, Utah 84043
United States

8015551212
dmedia@outlook.com

BILL TO
Fredrik H. Mantega

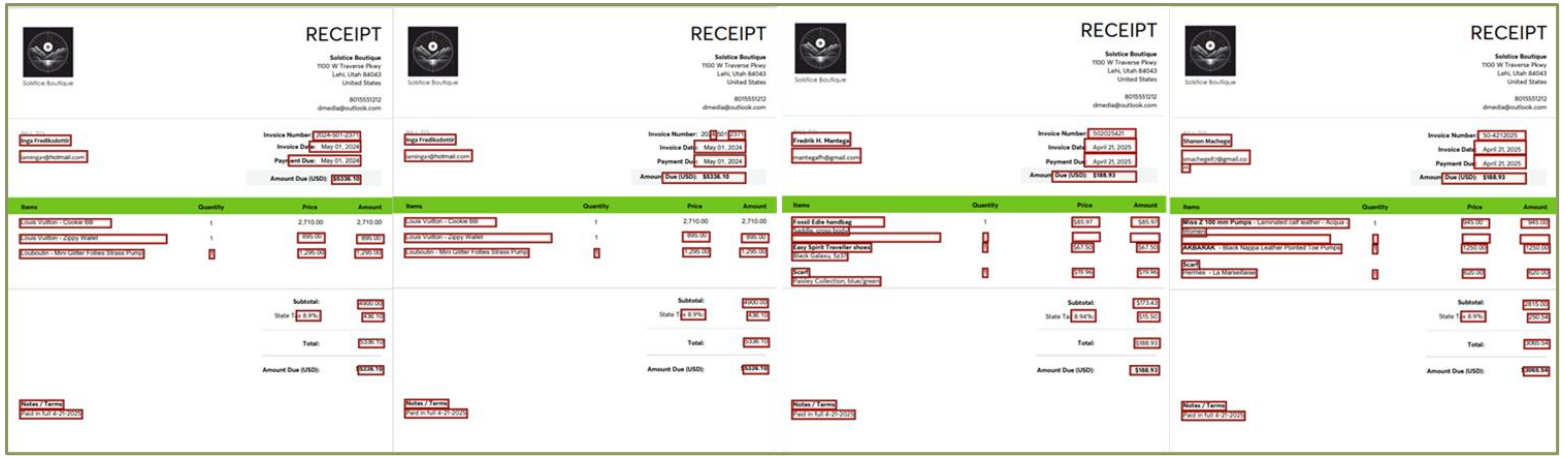
mantegafh@gmail.com

Invoice Number: 502025421
Invoice Date: April 21, 2025
Payment Due: April 21, 2025
Amount Due (USD): $188.93

Items	Quantity	Price	Amount
Fossil Edie handbag Saddle, cross body	1	$85.97	$85.97
Easy Spirit Traveller shoes Black Galaxy, Sz37	1	$67.50	$67.50
Scarf Paisley Collection, blue/green	1	$19.96	$19.96

Subtotal:	$173.43
State Tax 8.94%:	$15.50
Total:	$188.93
Amount Due (USD):	$188.93

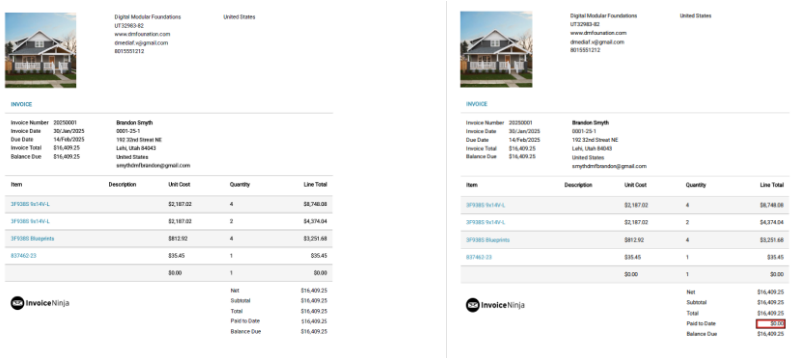
Or did the person editing use it as a template



We often see receipts (usually simpler than this one) where someone has used an older receipt as a “template” and overwritten the initial content in the file with specifics for the next customer. This *could be* someone generating invoices for multiple fraudulent claims, or it could be poor template use.

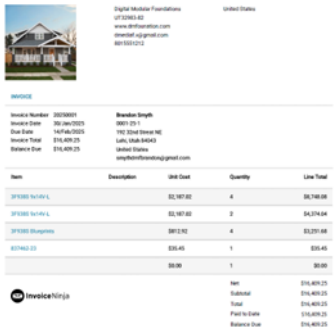
Secondary clues might be the editor tool(s) used and whether fonts are consistent.

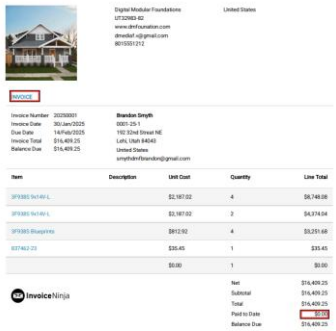
More template changes: invoice to receipt?

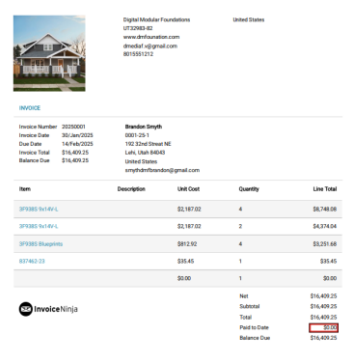


Is this invoice change suspicious?

Or is it benign?







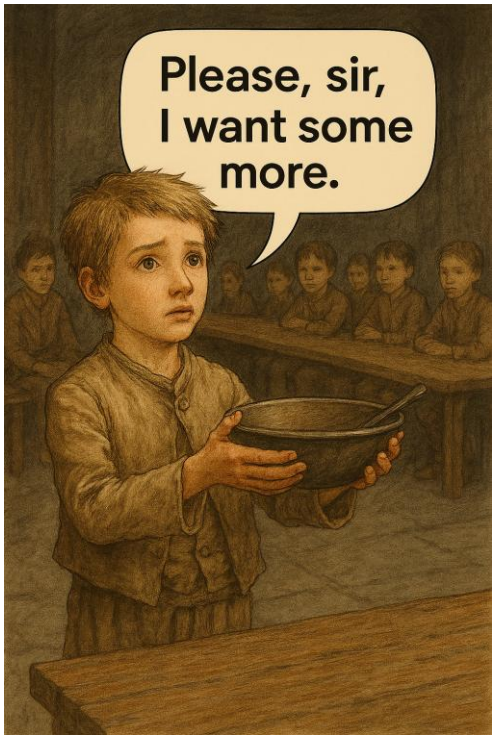
“Invisible” changes

- Many changes have no visual component
- Are non-visible changes less valid? No
- Examples of non-visible changes that may trigger generational updates in an incremental save
  - Changes to compression
  - Changes to some associated files or collections, including the attachment of C2PA manifests
  - Changes to objects or text on the page that do not have a visual representation
  - Editing of some fonts or color space information
  - Changes to optional content
  - Changes to encryption
  - Addition or changes to annotations

Thoughts on security and incremental saves

- Sensitive data may be recoverable even after redaction or deletion
  - Apply secure redaction: Remove sensitive content at the object level without erasing history.
- Lack of encryption on older revisions poses data leakage risks
  - Encrypt files with the most up-to-date encryption algorithms
- Malicious code may be embedded
  - Yes, and malicious code may be embedding in full saves as well. Avoid sloppy coding and use most current protocols to ensure the safest possible files

Final thoughts



- Incremental saves can be one of the most useful tools in PDF Forensics
- Being able to trace previous versions can illuminate workflow vs potentially fraudulent editing – intent can be more apparent
- Incremental saves can also be hugely helpful with document versioning and restoring previous versions in collaborative workflows
- Historical data can be preserved, studied, and better archived when incremental saves are used
- Not enough PDF producers and editors are allowing incremental saves
- Security issues can be mitigated

Verisk

Questions?

Cherie.Ekholm@verisk.com