# Solid Framework 10.0.12246 Quality Improvements

Reconstruction and API Improvements since Build 11962

Release Date – 9th August 2021

# Main Improvements

- Improved consistency of conversion, regardless of Operating System used
- Improved detection of page orientation
- Improved detection of All Caps and Small Caps
- Improved handling of Thai combining characters
- NSE improvements

# Each slide has an icon that indicates how the case is tested

   Tested automatically at every build

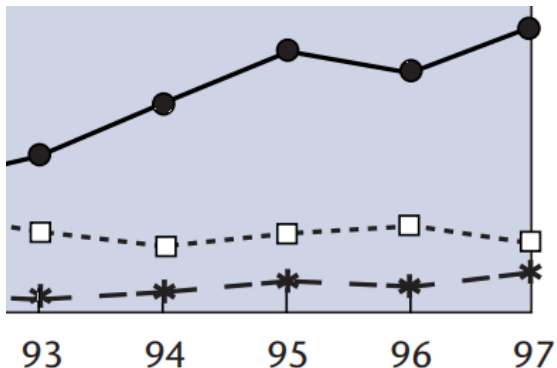   Tested manually. This may be because there is no obvious way to test, or that testing is too slow

Where there are multiple slides for a given issue, these generally represent a small subset of the many files that have been improved
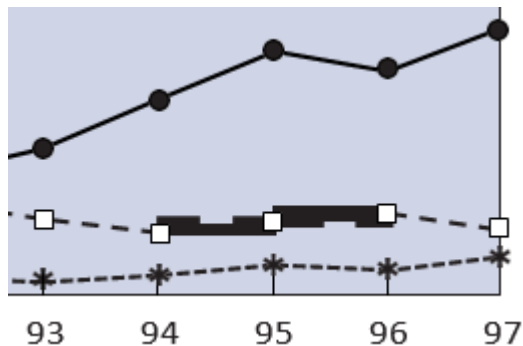
# Reconstruction as Word
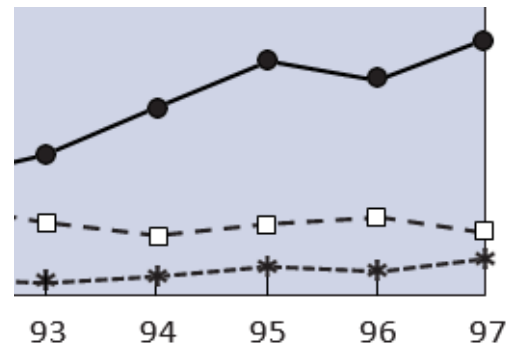
# Chart Lines Now Detected Correctly

PDF

11784

12246

Horizontal segments of
lines are very thick

Conversion is now
correct

| Original name | |
|---|---|
| SHA1 | a594de8b533cae8dcb5b91a711be273a78e6aed6 |
| FB | 26212 |

TESTED

# Correct Detection of All Caps Text

PDF

11962

12246

FEATURES

Vegas Experience

FEATuRES

Vegas Experience

FEATURES

Vegas Experience

One of the characters is not recovered as upper case

Conversion is now correct

| Original name | |
|---|---|
| SHA1 | 5b96b9a3f35805d9ff344519bb47e4330e1fb280 |
| FB | 25924 |

TESTED

# Correct Detection of Small Caps Text

PDF

THE ASTROPHYSICAL JOURNAL,
© 2007. The American Astronomi

D J. W. BARNES[4]
6 December 20

11962

THE AsTRoPHYsIcal JoURNaI,
© 2007. The American Astronomical

D J. W. BaRNEs[4]
6 December 20

Some of the characters are not recovered as Small Cap

12246

THE ASTROPHYSICAL JOURNAL,
© 2007. The American Astronomical

D J. W. BARNES[4]
6 December 20

Conversion is now correct

| Original name | |
|---|---|
| SHA1 | a994773616639976416ad255f657974f8723aa24 |
| FB | 26193 |

TESTED

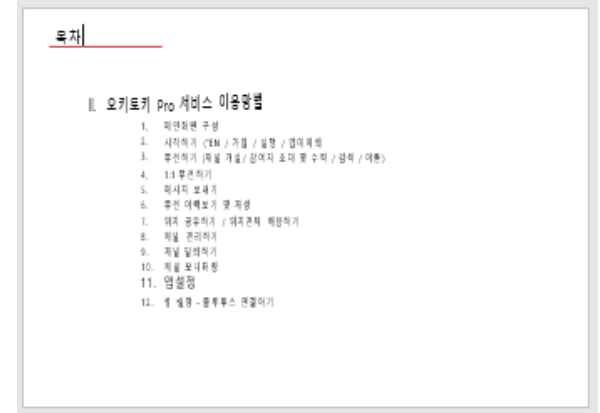# Correct Autorotation for Scanned Korean Document


PDF


11962


12246

**Document incorrectly considered to be upside down. This ruins OCR**

**Orientation correctly detected. OCR gives good result**

| Original name | |
|---|---|
| SHA1 | 9f30f130c7a8532c28cc9dc5ee2b9e4da055c0f1 |
| FB | 27065 |

TESTED

# Better Handling of Thai Combining Characters

**PDF**

**11962**

**12246**

๑. หลักการและเหตุผล
   สำนักงานสถิติแห่งช
คิดเห็นของประชาชนเกี่ยว

ด. หลักการและเหตุผล
   สํฺฺอํฺนฺฺฺกฺฺงอํฺนฺฺฺสถฺิตฺิแฺฺ
วฺฺฺิเครฺฺฺฺฺอําฺะหฺฺฺฺฺ์และประสฺฺฺฺอํฺนฺฺฺฺฺก

ด. หลักการและเหตุผล
   สํานักงานสถิติแห่งช
คิดเห็นของประชาชนเกี่ยว

The diacritics do not correctly combine, which results in a broken layout

Conversion is now much improved

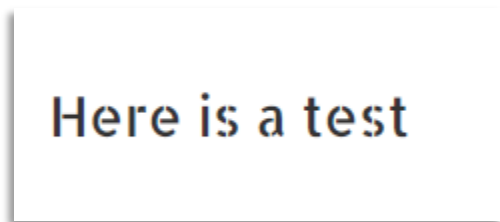| Original name | |
|---|---|
| SHA1 | 7ff948835ef653570c5c94561ed63ffa73daa853 |
| FB | 27293 |

# NSE Improvements

We have continued to improve conversion of PDFs that contain Non Standard Encoding.

The following slides are just a few examples of the huge number of improvements.
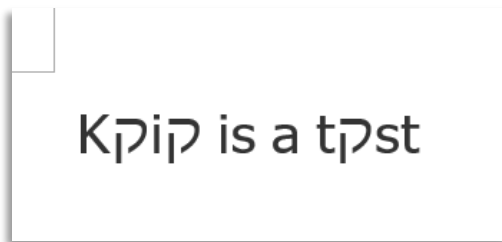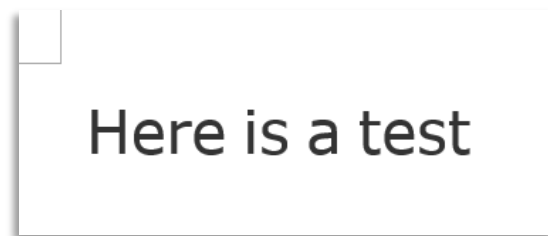
# Improved NSE for Exotic Font

## PDF

Here is a test

The characters in the font are made of up several disconnected parts

## 11962

Kpip is a tpst

The characters in the font have been mis-recognised

## 12246

Here is a test

Conversion is now correct

| Original name | |
|---|---|
| SHA1 | 3188fee0e60f3ae83233ddfadde34ebeb268b7e3 |
| FB | 27233 |

TESTED

# Improved NSE for Exotic Font

**PDF**

The Turtle Bay Pivot: How
Council Is Reshaping Nava
Rogue States, and Pirates

Reconciling Quasi-States v
Under the ICC Statute .....

The text is encoded in a non-standard way

**11962**

The Turtle Bay 3ivot: How
Council Is Reshaping Nav
Rogue States, and 3irates .

Reconciling 4uasi-States v
Under the ICC Statute .....

Some characters in the font have been mis-recognised

**12246**

The Turtle Bay Pivot: How
Council Is Reshaping Nava
Rogue States, and Pirates ..

Reconciling Quasi-States v
Under the ICC Statute ......

Conversion is now correct

| Original name | |
|---|---|
| SHA1 | b9999763659e64202abb84d1f0eba3d7f35de68b |
| FB | 21373 |

TESTED

# Improved NSE for Romanian Text

**PDF**



**11784**



Some characters in the font have been mis-recognised

**12246**



Conversion is now correct

| Original name | |
|---|---|
| SHA1 | 67d495364633500e8cfd5048713e2f39523d54b0 |
| FB | 18537 |

TESTED

# Recovery of Rupee Character from NSE text

**PDF**

oplication is, thus, disposed

ill be a cost of ₹ 2000.00 to b

itioner by the West Beng

15 days from date. The res

**11962**

oplication is, thus, disposed o

ill be a cost of ` 2000.00 to be

itioner by the West Beng

15 days from date. The res

The Rupee character, which is incorrectly encoded, is not recovered

**12246**

oplication is, thus, disposed o

ill be a cost of ₹ 2000.00 to be

itioner by the West Beng

15 days from date. The res

Conversion is now correct

| Original name | |
|---|---|
| SHA1 | 8b4161e227aa6be18319ba65300a252b22be0375 |
| FB | 21868 |

TESTED

# Reconstruction as Excel

# Conversion to Excel

**PDF**

| |
|---|
| Rs. 131.57 |
| Rs. 10,356.00 |
| Rs.    863.00 |

**11784**



Cell is split over 5 separate columns, with text split into separate cells making editing difficult

**12246**

| |
|---|
| Rs. 131.57 |
| Rs. 10,356.00 |
| Rs.    863.00 |

Cell is merged

| Original name | |
|---|---|
| SHA1 | c8aa4aa5163bae47df1d53c68ddc60be5f3fc8d1 |
| FB | 26599 |

TESTED

# Improved Consistency of Conversion on Different Platforms

Solid Framework is supported on Windows, macOS and Linux

We aim to have the same results regardless of which Operating System is used. In this release we have identified and solved some cases where the result was different on the different platforms

# Consistent behaviour on Windows, macOS and Linux

PDF

Connessione fr
portato all'ester
$\Psi_e = 0{,}77$ W/K

11962

Windows
macOS

Connessione fr
portato all'ester
$\Psi_e = 0{,}77$ W/K

12246

Connessione fra
portato all'ester
$\Psi_e = 0{,}77$ W/K

Linux

Connessione fra
portato all'estern
$\ddot{\circlearrowright}_e = 0{,}77$ W/K

Connessione fra
portato all'estern
$\Psi_e = 0{,}77$ W/K

On Linux "Ψ" is incorrect

Conversion correct
on all platforms

TESTED

| Original name | |
|---|---|
| SHA1 | 5b96b9a3f35805d9ff344519bb47e4330e1fb280 |
| FB | 25924 |

# Improved Layout on Linux

PDF



11962



Conversion on Windows was correct. On Linux some characters are not detected which then causes very poor layout.
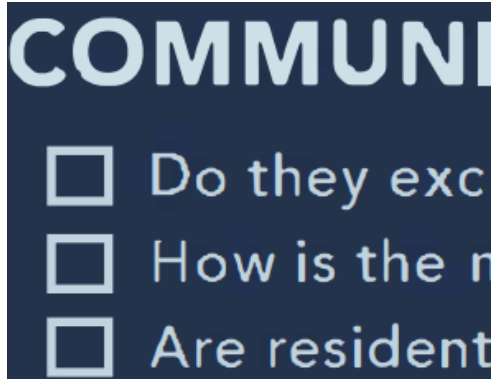
12246



Conversion correct on all platforms

| Original name | |
|---|---|
| SHA1 | 403249a03d48477ef6757bd287d63fab806365e1 |
| FB | 27125 |

TESTED

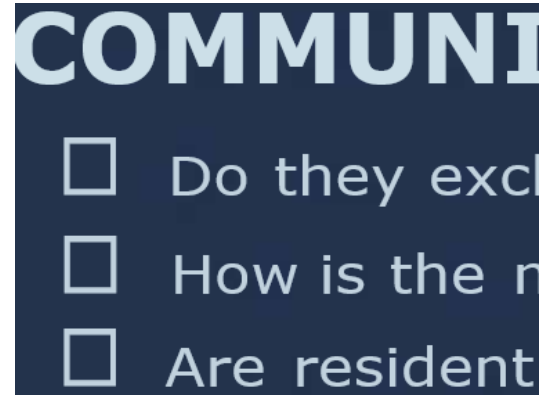# Improved Layout on Linux

PDF



11962



Check boxes reconstructed as "D" on Linux

12246



Conversion correct on all platforms

| Original name | |
|---|---|
| SHA1 | d62e674db8b4e497013954163a98f5bbe3ff81a8 |
| FB | 27504 |

TESTED

# Improved Layout on Linux - Other Examples

**PDF**

**11962**

**12246**



regra 8.4     Po

✓ *começar a*: O

✓ *continuar a*: O

**SEDE LEGALE:** VIA
BARBERINO DEL
**SEDE OPERATIVA:**

www.evansdavis.com
†*Licensed in OK & TX*

***Fiche travail Séan***

FB27202

FB27173

FB27203

FB27175

# Conversion to UTF7 Text on Linux





656c06bc6774e1450051bda5517d20b1d9a937f5_11962.2.txt - N...

File   Edit   Format   View   Help



656c06bc6774e1450051bda5517d20b1d9a937f5_12246.2.txt - N...

File   Edit   Format   View   Help

Brewprint

The Newsletter of the Boston Wort Processors
Volume IX, Number 12                    December 1996


WORTstation
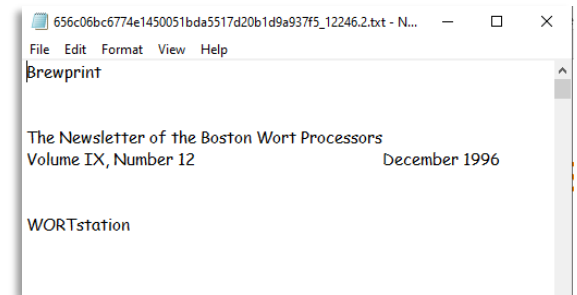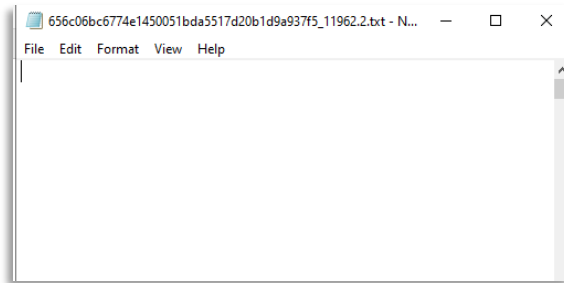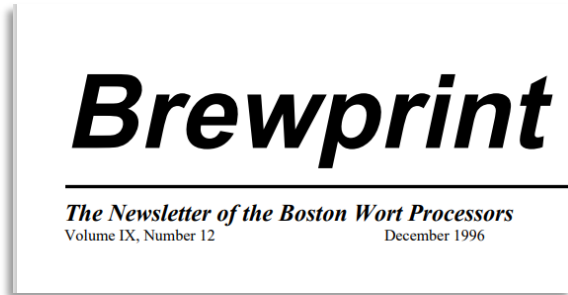
**While working on Windows, conversion to UTF7 text on Linux resulted in an empty file**

**Text is now extracted on Linux, Windows and macOS**

| Original name | |
|---|---|
| SHA1 | 656c06bc6774e1450051bda5517d20b1d9a937f5 |
| FB | 27248 |

TESTED

# Other Improvements

FB27236, 27238 and 24673: Slight modifications to the docx format, to avoid validation errors while retaining correct appearance

FB22725: Retention of hyperlinks attached to images

FB 26448: Improved support for hyperlinks with unusual characters

# Number of FogBugz Cases with Automated Tests

| | 30th Nov 2020 | 13th May 2021 | 12th Aug 2021 |
|---|---|---|---|
| Regression cases - Active | c. 210 | 268 | 311 |
| Regression cases - Closed | 1155 | 1740 | 1991 |
| All cases - Active | c. 1000 | 1218 | 1301 |
| All cases - Closed | 2590 | 3575 | 4118 |
| Total number of tests | c. 3600 | 4793 | 5419 |

A huge effort has gone into creating automated tests to ensure that things that are fixed stay fixed.

In addition there are more than 80,000 other tests which are run on multiple machines and OSs on every build.