



Web



Choose a Page



WEB &gt; GUIDES

# Enhancements in PSPDFKit for Web 2023.1

[PSPDFKit for Web 2023.1](#) introduces several larger changes. To determine if you need to take action, check your implementation and the information below.

## New OCR and Office conversion engines

PSPDFKit Server is now shipping with brand-new [OCR](#) and [Office conversion](#) engines, which are based on [GdPicture.NET](#).

Our previous OCR engine was based on the [Tesseract](#) open source project, and we used [LibreOffice](#) as the core of our Office conversion tools. This allowed us to produce quality results, but we found it lacking in certain aspects due to the nature of these two fundamental dependencies that were powering it. The main issue with our OCR engine was the performance, which was only acceptable at best. In the case of Office conversion, our main pain point was that we were unable to effectively improve the conversion quality itself.

Both new engines bring improved performance and accuracy, with documents being processed more quickly and accurately. The OCR performance gain is especially considerable: We measured improved performance of up to 7× when compared to the previous engine — all while delivering the same or sometimes even better accuracy.

Usage of these new engines requires a license update. If your license already includes OCR or Office conversion components, you qualify to get access to the updated engines for free. In that case, you'll need to [update your activated license](#) to enable them.

In case you encounter any changes in behavior or regressions that would break your workflow, you can revert back to the old engines via following Server's [configuration options](#):



- ✧ `OCR_ENGINE` — The OCR engine defaults to `gdpicture`. Use `core` to revert to the old engine.
- ✧ `CONVERSION_ENGINE` — The Office conversion engine defaults to `gdpicture`. Use `libreoffice` to revert to the old LibreOffice-based Office conversion engine.



Usage of old OCR and conversion engines is deprecated, and we'll drop support for them in a future version. Please submit any issues you encounter with new engines to [Support](#).

## Instant JSON changes

The Instant JSON schema has been updated to provide rich text support for [annotations](#) and [comments](#):

- ✧ The version of the annotations and comments JSON is now `2`. In other words, instead of using `"v": 1`, the annotations and comments JSON now uses `"v": 2`.
- ✧ The `text` property of text annotations (`pspdfkit/text`), note annotations (`pspdfkit/note`), and comments (`pspdfkit/comment`) is now an object with the following signature:

```
1 declare type Text = {
2   // Format of the text value. Defaults to 'plain'.
3   format?: "xhtml" | "plain",
4   // Value itself.
5   value: string
6 };
```



### Before

```
1 // Example of an Instant JSON schema for a text annotation:
2 {
3   "v": 1,
4   "pageIndex": 1,
5   "bbox": [150, 275, 120, 70],
6   "opacity": 1,
7   "pdfObjectId": 200,
8   "creatorName": "John Doe",
9   "createdAt": "2012-04-23T18:25:43.511Z",
10  "updatedAt": "2012-04-23T18:28:05.100Z",
11  "id": "01F46S31WM8Q46MP3T0BAJ0F85",
12  "name": "01F46S31WM8Q46MP3T0BAJ0F85",
```



```
13 "type": "pspdfkit/text",
14 "text": "Content for a text annotation",
15 "fontSize": 14,
16 "fontStyle": ["bold"],
17 "fontColor": "#000000",
18 "horizontalAlign": "left",
19 "verticalAlign": "center",
20 "rotation": 0
21 }
```

After

```
1 // Example of an Instant JSON schema for a text annotation:
2 {
3   "v": 2,
4   "pageIndex": 1,
5   "bbox": [150, 275, 120, 70],
6   "opacity": 1,
7   "pdfObjectId": 200,
8   "creatorName": "John Doe",
9   "createdAt": "2012-04-23T18:25:43.511Z",
10  "updatedAt": "2012-04-23T18:28:05.100Z",
11  "id": "01F46S31WM8Q46MP3T0BAJ0F85",
12  "name": "01F46S31WM8Q46MP3T0BAJ0F85",
13  "type": "pspdfkit/text",
14  "text": {"value": "Content for a text annotation", "format": "plain"},
15  "fontSize": 14,
16  "fontStyle": ["bold"],
17  "fontColor": "#000000",
18  "horizontalAlign": "left",
19  "verticalAlign": "center",
20  "rotation": 0
21 }
```

## Postgres 15 support

This version brings compatibility with PostgreSQL 15.

For upgrade instructions, please refer to either the official [PostgreSQL upgrade guide](#) if you're running the database on-premises, or your cloud provider's documentation.

## Database migrations

This release includes two database migrations. Always [upgrade PSPDFKit Server one node at a time](#), but this is especially important to do when the upgrade includes database migrations.

First, the migration adds a new table called `secrets` that's used as a data storage for the secrets rotation API. Adding the new table doesn't affect existing data, so you shouldn't experience performance issues because of it.

Second, the migration adds a new column to the `records` tables. This will temporarily block any reads and writes to these tables. Adding columns to existing tables is usually a quick process, so you shouldn't experience performance issues because of it. In our tests on tables with 1.5 million rows, the migration held a lock on the table for about 2.5 ms.

Each migration is run inside a transaction, so if it fails or you interrupt it, you can safely retry it.

If you experience any problems with migration during the upgrade, please [submit a support request](#).

## Other ohanges

This release includes many more improvements. For a full list of changes, check out the [Server changelog](#).

## Migrate PSPDFKit for Web

For more information, see [PSPDFKit for Web 2023.1 migration guide](#).

Was this helpful?

☒ YES

☐ NO

Questions? [Contact us](#)

