

DjVu (произносится как дежавю) - это формат компьютерного файла, в основном используемый для хранения отсканированных документов. Особенности этого формата включают слоистость изображений, постепенную загрузку, арифметическое кодирование и потерянное сжатие бинарных изображений, что позволяет хранить высококачественные читаемые изображения на меньшем пространстве и ранее широко использовался для форматов хранения отсканированных книг.

DjVu может иметь добавленный текстовый слой оптического распознавания символов, что позволяет выполнять операции копирования и вставки.

Технология DjVu была изначально разработана Яном ЛеКуном (Yann LeCun), Леоном Боттоу (Léon Bottou), Патриком Хаффнером (Patrick Haffner) и Полом Г. Ховардом (Paul G. Howard) в лабораториях AT&T в 1996 году. DjVu - это открытый формат файла, спецификации формата файла и исходный код ссылочной библиотеки публикуются открыто. Права на коммерческую разработку на протяжении многих лет передавались различным компаниям, включая AT&T и LizardTech. Оригинальные авторы поддерживают реализацию GPL под названием DjVuLibre.

В 2002 году DjVu, TIFF и PDF были выбраны проектом Million Book Project Интернет-архива в качестве форматов файлов для отсканированных книг общественного достояния для публикации в сети. Постепенная загрузка делает DjVu подходящим для использования в Интернете. Для отсканированных документов размер файла и скорость открытия DjVu обычно лучше, чем у PDF.

Однако, поскольку большинство программ редактирования не поддерживает создание формата DjVu, а редакторы, поддерживающие формат DjVu, не так популярны, как для PDF.

Пространство для хранения и скорость сети уже не являются препятствиями. Книги, ориентированные на текст, чаще сохраняются в формате ePub (подходит для чтения на маленьких мобильных устройствах, с возможностью настройки макета) и PDF, по различным причинам формат DjVu больше не является распространенным для новых личных отсканированных документов.

