

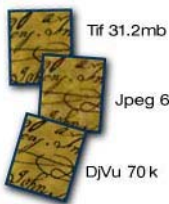
## Что такое DjVu

### и откуда у него такие возможности?



Более 90 процентов информации в мире по прежнему остаются на бумаге. Многие из этих документов содержат цветные иллюстрации и/или фотографии, которые имеют большую ценность. И почти ничего из этого богатого материала нет в интернете.

Это потому, что сканирование таких документов и размещение их на веб сайте в лучшем случае было большой проблемой. При сканировании текста с высоким разрешением и сохранении качества изображений размер файлов становится очень громоздким для быстрого скачивания. Уменьшение разрешения для повышения скорости скачивания ведет к потере качества и разборчивости. Такие распространенные веб форматы как JPEG, GIF и PNG при приличном разрешении создают невозможно большие изображения. Как следствие, разработчики веб сайтов практически не могли предоставить существующие бумажные материалы.



DjVu (читается "де-жа-вю") – это новая технология сжатия изображений разработанная в 1996г. AT&T Labs именно для этой цели. DjVu позволяет распространение по интернету изображений документов очень высокого качества, цифровых документов и фотографий. DjVu позволяет редакторам сканировать в высоком разрешении цветные страницы книг, журналы, каталоги, руководства, газеты, исторические и древние документы, и сделать их доступными на вебе.

Информация, которая ранее была сохранена в твердом виде, теперь может быть представлена широкой публике.

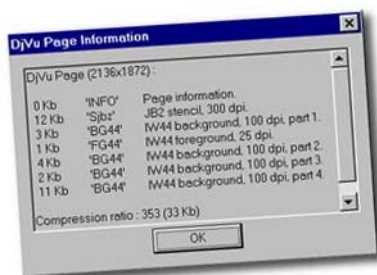
### Преимущества и новые возможности

Исследовательские институты, библиотеки и государственные учреждения могут предоставить доступ к своим архивам. Компании могут распространять внутренние документы по своим каналам.

Продвижением DjVu занимается расположенная в Сизтле [LizardTech Inc.](http://LizardTech.com) совместно с AT&T Labs. DjVu – открытый стандарт. Доступны как описание формата файла (спецификация), так и исходный код декодера (и части кодировщика).

### Технология Сжатия нового поколения

- DjVu combines six technology break-throughs from AT&T Labs, including:
- Two-stage decompression algorithm for on-the-fly, on-demand image decompression
  - Foreground-background separation technique
  - W44 wavelet-based progressive image compression technology
  - JB2 bi-level image compression technology
  - ZP-coder binary adaptive entropy coder
  - Projection-image masking technique



Обычно DjVu сжимает в 5-10 раз лучше, чем существующие методы вроде JPEG и GIF для цветных документов, и в 3-8 раз лучше TIFF для черно-белых документов. Цветные страницы, сканированные в полном цвете с разрешением 300 DPI могут быть сжаты с 25МБ до 30-100КБ. Черно-белые документы 300 DPI обычно сжимаются до 5-30КБ. Это ставит размер качественно сканированных страниц в один ряд с обычными HTML страницами (занимающими обычно 50КБ)

Для цветных документов содержащих одновременно текст и картинки файлы DjVu обычно в 5-10 раз меньше JPEG файлов того же качества. Для черно-белых страниц DjVu файлы обычно в 10-20 раз меньше JPEG и в 5 раз меньше GIF. Также DjVu файлы в 3-8 раз меньше черно-белых PDF файлов, полученных из сканированных документов (сканировать цветные докуементы в PDF непрактично).

Помимо сканирования документов DjVu можно применять к документам, созданным другими программами, например Adobe PostScript или PDF. В этом случае размер файла варьируется от 15 до 20КБ за страницу 300 DPI.



DjVu плагин доступен для стандартных браузеров (IE и Netscape) для различных платформ. DjVu плагин позволяет легко перемещаться и увеличивать страницы документа. Уникальная технология распаковки "на лету" позволяет изображениям размером 25МБ занимать только 2МБ ОЗУ.

Обычные программы просмотра изображений перед показом распаковывают изображение целиком. Это непрактично для образов документов высокого качества, поскольку они исчерпывают оперативную память многих ПК, вызывая чрезмерный обмен с жестким диском. Напротив, DjVu никогда не распаковывает образ целиком,



он хранит его в памяти в компактном виде и распаковывает фрагмент, видимый на экране именно тогда, когда его открывает пользователь. Крупные изображения как например 2500 x 3300 точек (обычная страница 300 DPI) могут быть загружены и просмотрены на очень слабых ПК.

Формат DjVu распаковывается поэтапно (прогрессивно). Первый набросок изображения пользователи получают очень быстро, с поступлением данных

качество изображения постепенно увеличивается. Например на модеме 56K текст обычного журнала появится через три секунды. Через одну-две секунды появятся наброски картинок и фона. Затем, через несколько секунд формируется окончательная качественная версия страницы.

Технология DjVu разделяет изображение на задний слой (например текстура бумаги и фон) и передний слой (текст и рисунки). Традиционные методы сжатия отлично работают с простыми фотографиями, но они значительно теряют четкость цветовых переходов на насыщенных контрастных областях, посему и получают такие нечеткие результаты. Отделяя текст от фона DjVu хранит текст с высоким разрешением (таким образом сохраняя четкие края и максимальную читаемость), в то же время сжимая фон и изображения с более низким разрешением используя технологию сжатия вейвлет (wavelet).

Сегодня Djvu используется многими коммерческими и некоммерческими веб сайтами (в США/ прим. перев.).

Технические детали DjVu: [раздел документации \(англ\)](#), и [раздел технических статей \(англ\)](#)

Информация по использованию DjVu: раздел [обучения \(англ\)](#).

---

Перевод выполнен fuser. Последняя редакция 2.06.2002

Оригинал статьи находится на сайте [djvu zone](#).