



УДК 004.912

© **Ольга Бриндзя**, студентка 4-го курсу, НТУУ «КПІ», 2011 р.
Науковий керівник: О. П. Кохановський, к.ф.-м.н., доцент,
НТУУ «КПІ»

ВИКОРИСТАННЯ ФОРМАТУ PDF ТА DJVU ДЛЯ СТВОРЕННЯ ЕЛЕКТРОННИХ ДОКУМЕНТІВ

*The issues of using formats PDF and DjVu for creating
electronic documents were considered.*

Нині створення електронних версій друкованих видань стало дуже актуальним і набуло широкого поши-

рення. З кожним роком відсоток електронних видань впевнено збільшується. Існує багато способів представлення електронних документів — HTML, PDF, DjVu та ін. Тому актуальності набуває проблема вибору електронного формату для представлення попередньо оцифрованої друкованої інформації. Це питання потрібно детально досліджувати, щоб навести достовірні аргументи вибору того чи іншого формату.

У результаті проведеного інтернет-дослідження виявилось, що 7 осіб із 10 надають перевагу формату PDF (створений в 1991 році корпорацією Adobe [<http://www.adobe.com/>]) перед DjVu ([<http://djvu.org/>] створений у 1996 році компанією AT&T [<http://www.att.com/>]), аргументуючи свій вибір популярністю та поширеністю формату.

Проте при виборі оптимального формату для якісного відтворення інформації на екрані необхідно враховувати інші критерії. Обидва формати були створені з дещо різними цілями і для вирішення не зовсім однакових завдань.

1. Різниця у властивостях. Формат PDF можна умовно поділити на 2 підвиди: векторний і растровий, а формат DjVu — виключно растровий. Тому порівнювати потрібно лише растровий PDF і DjVu.

2. Різниця у призначенні. У форматів PDF і DjVu різна ідеологія. Формат PDF створювався для друкованого відтворення документа, а формат DjVu — для електронного.

Для визначення недоліків і переваг форматів PDF і DjVu був проведений їх детальний аналіз і винесено основні особливості.

На відміну від формату DjVu, PDF стандартизований міжнародним комітетом по стандартам ISO. Існує багато програм по роботі з форматом (більшість з них платні і дорого коштують). Формат PDF досить складний за своєю внутрішньою структурою. Опис його специфікації займає більше 1000 сторінок, тоді як опис DjVu — біля 20



сторінок. PDF з самого початку розроблявся для роботи з електронними, а не відсканованими документами, а DjVu — для роботи зі сканованими зображеннями.

Формат DjVu використовує 2-шарову сегментацію, що забезпечує оптимальне стиснення і має ряд властивостей (які відсутні в PDF): дозволяє переглядати окремі сторінки файлу, який ще не повністю завантажився і робити вибірковий перегляд сторінок через індексний файл; забезпечує прогресивне відтворення на екрані.

Формат DjVu потребує мінімальних ресурсів, має малий розмір файлу. У DjVu-файл можна впровадити текстовий шар, що дає можливість користуватись навігацією по файлу та індексацією за ключовими словами.

У DjVu існує можливість помилки типу «проблема інь» (офіційна назва — Transposed Letters). Деякі літери в DjVu-файлі помилково замінюються на інші, зовні схожі з ними (наприклад, підміна букви «и» на «н»). Проблему можна побороти скануючи паперовий документ з роздільністю не менше 600 dpi.

Отже, врахувавши всі особливості форматів, виявилось, що їх не можна вважати суперниками. Вони призначені для вирішення різних задач, а тому не змагаються, а доповнюють один одного. Обидва формати, і DjVu, і PDF, широко використовуються при створенні електронних документів. Наприклад, якщо потрібно зробити переносну копію електронного документа із Microsoft Word, таблицю з Excel — то, звісно, краще використовувати векторний PDF. А коли мова йде про оцифрування паперового документа — то в цьому випадку доцільним буде використання DjVu.

