

## ТЕХНОЛОГІЯ СТИСНЕННЯ DJ VU

**Некоз Д.І., гр. ХТ-42**

Наукові керівники: канд. фіз.-мат. наук, доц. **Рибін О.М.,**

ст. викл. **Сибірякова І.Г.**

Харківський державний університет харчування та торгівлі

Незважаючи на зростаючу важливість Інтернету, велика частина знань у вигляді культурного та освітнього матеріалу, вже існуючого сьогодні, все ще доступна лише в паперовій формі. Впровадження цього інформаційного багатства в цифрову область в схожому з оригіналом вигляді, з легким доступом і пошуком є основним кроком до створення Всесвітньої Інтернет Бібліотеки.

DjVu (дежа вю) – це технологія стиснення, файловий формат та універсальна платформа, спеціально спроектована для створення цифрових бібліотек на основі друкованого, сканованого або цифрового матеріалу. Так, наприклад, для сканованого документа розмір DjVu файлу зазвичай в 3-10 разів менше, ніж чорно-білий TIFF або PDF, і в 5-10, ніж кольоровий JPEG.

Звичайна сторінка з книги, журналу або стародавнього рукопису, що відсканована в кольорі з роздільною здатністю 300dpi, містить близько 8 мільйонів пікселів і займає 24МВ в стислому вигляді. У форматі DjVu текст і контрастні малюнки зберігаються з роздільною здатністю 300dpi, все інше вважається фоном і зберігається зі зменшеною роздільною здатністю. Це дозволяє добре стиснути електронний документ без втрати його читабельності. У DjVu розмір файлу відсканованої книги виявляється в межах декількох мегабайт, що цілком прийнятно. DjVu як формат складається з декількох технологій стиснення: DjVuPhoto (IW44) – стиснення з втратою якості; DjVuBitonal (JB2) – метод стиснення, в якому увага акцентується на монохромних елементах (наприклад, жорсткі лінії, китайські ієрогліфи, індексовані зображення), що дозволяє отримувати документи достатньо високої якості; DjVuDocument – документ поділяється на передній та задній шари: передній зберігається як JB2, задній – IW44.

Особливе значення цей формат набуває для переносу до мережі математичної і взагалі технічної літератури, де велика кількість схем і формул робить розпізнавання і переклад в текстовий формат практично нездійсненним. В даний час формат DjVu стає фактичним стандартом для електронних бібліотек технічної та наукової літератури.