

iArchives, Inc.

Creating efficiency in the legal world

History of iArchives

Newspapers from microfilm

The Dallas Morning News

Early American Newspapers

University Libraries

Harvard, University of Florida, University of Utah

Mormon Church

Family History Library, Genealogy Program

Optical Word Recognition=*OWR*

Optical Character Recognition=*OCR*

“OWR”
Patented
Trademarked

Technology of iArchives, Inc.

Focus on words

Multiple OCR Engines

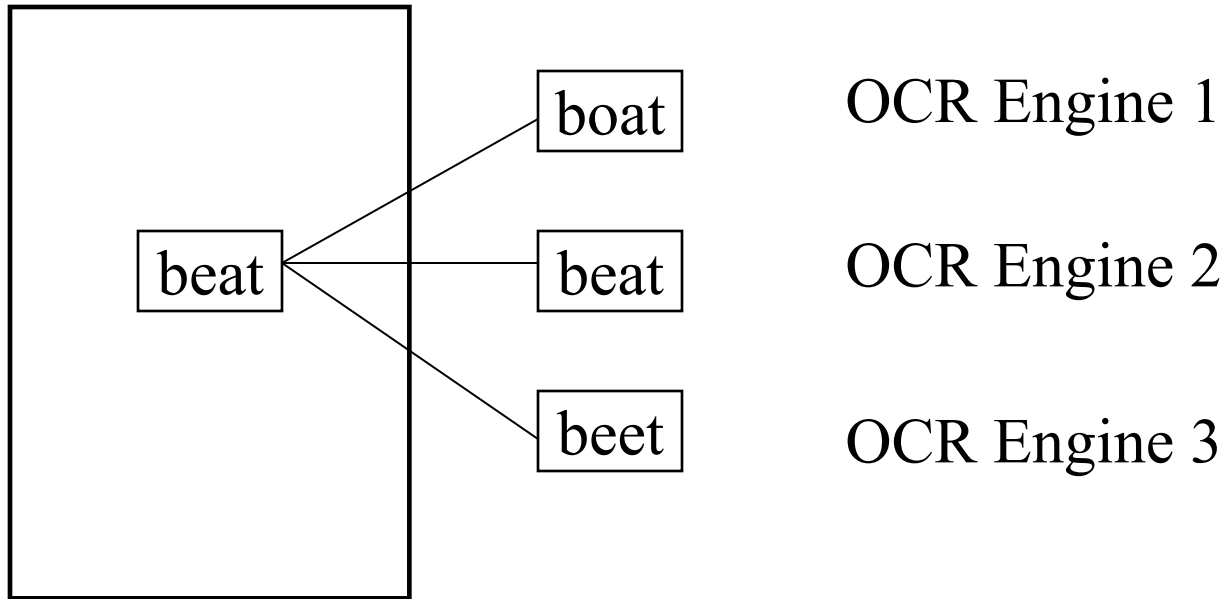
2,000,000 word dictionary

Specialized dictionaries

Customized dictionaries

iArchives' Advantage

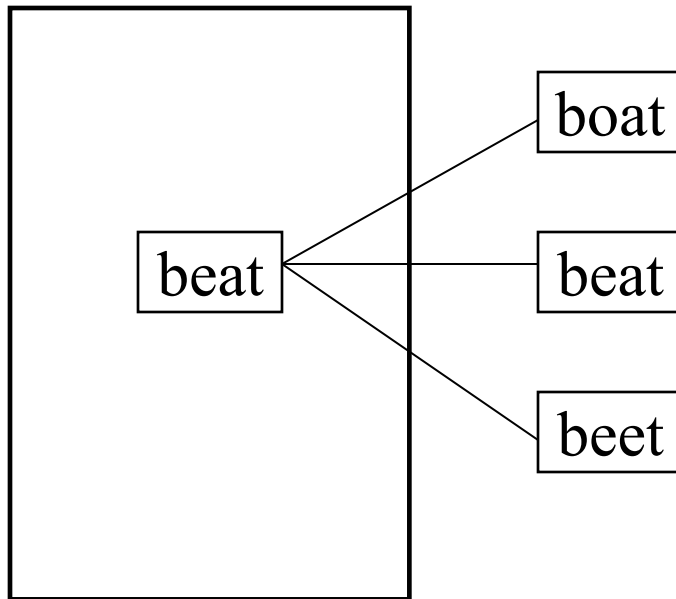
Text image



All OCR results are preserved
And OWR is created!

Voting OCR – Who will win?

Text image



OCR Engine 1 votes for boat

OCR Engine 2 votes for beat

OCR Engine 3 votes for beet

Voting produces one answer.

If the incorrect character(s) gets the vote, the correct answer will never be found when searched.

iArchives' Advantages

- Fast turn-a-round time
- Proven “Highest accuracy on the market”
- Reduced costs
- Highlighted hits
- Per page/document accuracy
- Tiered-pricing option
- Process 1,000,000 pages per week
- New products coming !!!
 - *Names Mentioned*

OWR vs OCR

	iArchives OWR <u>accuracy</u>	Commercial OCR <u>accuracy</u>	OWR additional <u>words per page</u>
Clean Images	99.21%	97.15%	8
Dirty Images			
Legal documents	92.20%	78.79%	54
Book	91.16%	81.60	38
Almanac	98.07%	77.95%	80
Bad Images	63.74%	38.53%	101

In all tests, the iArchives' OWR out-performed the commercial OCR engine.

An 8 1/2 x 11 page contains 400 words. Even on the Clean Images, the iArchives' OWR located an average of eight more words per page than the other OCR engine. The difference between the results becomes more defined as iArchives' OWR located 54 more words than the other OCR engine on Dirty Images by being 13.41% more accurate and 101 more words were found by iArchives' OWR on Bad Images by being 25.21% more accurate.

Changing the Legal World