

# Paper Records Storage vs. Scanning

## A Cost Comparison

March, 2010

I created this cost comparison because I've been confronted many times with the assertion "it's cheaper to scan [inactive] paper records than it is to store them." I've always answered it with "not necessarily so, and, in most cases, probably not." But I've never had the numbers to back me up; I've increasingly felt the need for them, and so here they are.

This cost comparison is narrowly-focussed on the argument "scanning inactive records is cheaper than storing paper". It assumes a massive backfile conversion of inactive or semi-active records. It does not address "scan-on-demand". It does not consider the many good reasons for conversion.

In various ways, it is biased in favour of scanning. This is because, inevitably, someone is going to challenge the foundation data, and I'd rather say "the cost comparison is biased in favour of scanning."

The assumptions listed in the worksheets titled "Paper Storage" and "Scanning" are essential to understanding how the cost comparison is structured.

- For example, as stated in the assumptions under "Scanning", I did not include the costs of acquiring, operating, maintaining and (eventually) migrating a content/document management system, or on-going storage of paper when required, or content-based retrieval.
- In these worksheets, you can change the basic numbers in the cells tinted in light green (everything else in the workbook is Protected).

The Excel workbook contains 7 worksheets. These are:

- Cover Page
- Summary (this worksheet summarises, year by year, the comparative costs under three different scenarios, where the cost of storage in boldface shows the year in which the cost of storage exceeds the cost of scanning)
- Paper Storage (this worksheet includes the baseline numbers for paper storage)
- Scanning (this worksheet includes the baseline numbers for scanning)
- Optimistic Scenario (this worksheet shows a detailed breakdown of comparative costs under the scenario most favourable to scanning)
- Realistic Scenario (this worksheet shows a detailed breakdown of comparative costs under what might be considered a typical or average scenario)
- Pessimistic Scenario (this worksheet shows a detailed breakdown of comparative costs under a scenario more favourable to paper storage)

Have fun, and comments are most welcome.

Fred

=====  
Frederic J. Grevin  
Email: fgrevin@records.nyc.gov  
Land phone: 212-788-8615

# Comparison of Paper Storage vs. Scanning Costs - Summary

## Assumptions

- |                          |             |  |     |
|--------------------------|-------------|--|-----|
| 1. Total number of boxes | 100,000     | 3. Total annual retrieval & re-filing, as a % of whole | 10% |
| 2. Total number of pages | 120,000,000 | 4. Annual cost increase of paper storage               | 5%  |

Year	Optimistic Scenario		Realistic Scenario		Pessimistic Scenario	
	Storage Costs (cumulative)	Scanning	Storage Costs (cumulative)	Scanning	Storage Costs (cumulative)	Scanning
1	\$4,655,000	<b>\$18,031,333</b>	\$3,410,000	<b>\$30,606,000</b>	\$2,165,000	<b>\$70,200,000</b>
2	\$5,264,000	\$0	\$3,893,000	\$0	\$2,522,000	\$0
3	\$5,903,450	\$0	\$4,400,150	\$0	\$2,896,850	\$0
4	\$6,574,873	\$0	\$4,932,658	\$0	\$3,290,443	\$0
5	\$7,279,866	\$0	\$5,491,790	\$0	\$3,703,715	\$0
6	\$8,020,109	\$0	\$6,078,880	\$0	\$4,137,650	\$0
7	\$8,797,365	\$0	\$6,695,324	\$0	\$4,593,283	\$0
8	\$9,613,483	\$0	\$7,342,590	\$0	\$5,071,697	\$0
9	\$10,470,407	\$0	\$8,022,220	\$0	\$5,574,032	\$0
10	\$11,370,178	\$0	\$8,735,831	\$0	\$6,101,483	\$0
11	\$12,314,937	\$0	\$9,485,122	\$0	\$6,655,308	\$0
12	\$13,306,933	\$0	\$10,271,878	\$0	\$7,236,823	\$0
13	\$14,348,530	\$0	\$11,097,972	\$0	\$7,847,414	\$0
14	\$15,442,207	\$0	\$11,965,371	\$0	\$8,488,535	\$0
15	\$16,590,567	\$0	\$12,876,139	\$0	\$9,161,712	\$0
16	\$17,796,345	\$0	\$13,832,446	\$0	\$9,868,547	\$0
17	<b>\$19,062,412</b>	\$0	\$14,836,569	\$0	\$10,610,725	\$0
18	<b>\$20,391,783</b>	\$0	\$15,890,897	\$0	\$11,390,011	\$0
19	<b>\$21,787,622</b>	\$0	\$16,997,942	\$0	\$12,208,261	\$0
20	<b>\$23,253,253</b>	\$0	\$18,160,339	\$0	\$13,067,424	\$0
21	<b>\$24,792,166</b>	\$0	\$19,380,856	\$0	\$13,969,546	\$0
22	<b>\$26,408,024</b>	\$0	\$20,662,399	\$0	\$14,916,773	\$0
23	<b>\$28,104,676</b>	\$0	\$22,008,019	\$0	\$15,911,362	\$0
24	<b>\$29,886,159</b>	\$0	\$23,420,919	\$0	\$16,955,680	\$0
25	<b>\$31,756,717</b>	\$0	\$24,904,465	\$0	\$18,052,214	\$0
26	<b>\$33,720,803</b>	\$0	\$26,462,189	\$0	\$19,203,574	\$0
27	<b>\$35,783,093</b>	\$0	\$28,097,798	\$0	\$20,412,503	\$0
28	<b>\$37,948,498</b>	\$0	\$29,815,188	\$0	\$21,681,878	\$0
29	<b>\$40,222,173</b>	\$0	<b>\$31,618,447</b>	\$0	\$23,014,722	\$0
30	<b>\$42,609,532</b>	\$0	<b>\$33,511,870</b>	\$0	\$24,414,208	\$0
31	<b>\$45,116,258</b>	\$0	<b>\$35,499,963</b>	\$0	\$25,883,669	\$0
32	<b>\$47,748,321</b>	\$0	<b>\$37,587,462</b>	\$0	\$27,426,602	\$0
33	<b>\$50,511,987</b>	\$0	<b>\$39,779,335</b>	\$0	\$29,046,682	\$0
34	<b>\$53,413,836</b>	\$0	<b>\$42,080,801</b>	\$0	\$30,747,766	\$0
35	<b>\$56,460,778</b>	\$0	<b>\$44,497,341</b>	\$0	\$32,533,904	\$0

## Cost of Storing Paper Records

### Assumptions

1. Storage is outsourced to a Commercial Records Center (CRC).
2. "Inventory and boxing" includes box assembly, moving files from file cabinet to box, inventorying contents of box, labelling box, and staging box for pickup by storage vendor.
3. Pickup and data entry charge does not include transportation charge (\$15-\$20 per occurrence)
4. Storage charges will vary with quantity of boxes to be stored (more boxes = lower unit cost)
5. Retrieval and re-filing costs do not include transportation charge (\$15-\$20 per occurrence)
6. Box purchase and storage prices based on 1.2 cubic-foot box
7. Retrieval and re-filing rates are 10% of total number of boxes in storage, are equal (1 retrieval = 1 re-filing), and remain constant over total period of storage.
8. Total number of boxes in storage 100,000
9. Total number of pages in boxes 120,000,000
10. Total annual retrieval & re-filing, as a % of whole 10%
11. Annual cost increase 5%

	Optimistic	Realistic	Pessimistic
1. Purchase of box (one-time cost)	\$1.25	\$1.50	\$1.75
2. Inventory and boxing (one-time cost)	\$15.00	\$25.00	\$35.00
3. Pickup and data entry (one-time cost)	\$2.00	\$3.00	\$4.00
4. Storage (per box, per year - recurring cost)	\$3.00	\$4.00	\$5.00
5. Retrieval (recurring cost)	\$2.00	\$3.00	\$4.00
6. Re-filing (recurring cost)	\$2.00	\$3.00	\$4.00

	Optimistic	Realistic	Pessimistic
Total one-time cost (per box) (1+2+3 above)	\$18.25	\$29.50	\$40.75
Total one-time cost (all boxes)	\$1,825,000	\$2,950,000	\$4,075,000
Total annual storage cost (all boxes)	\$300,000	\$400,000	\$500,000
Total annual retrieval & re-filing costs	\$40,000	\$60,000	\$80,000

## Cost of scanning records

### Assumptions

1. The scanning work is outsourced to a conversion service bureau. The costs of boxing, inventorying, and transporting the records to the service bureau are not included. These costs may be substantial.
2. Document preparation is limited to removal of documents from boxes and folders, removal of paper clips and staples, unfolding of folded pages, and otherwise making pages scanner-ready; NO sorting or checking files for completeness. The number of sheets of paper in a 1.2 cu ft box will vary according to how the records are divided: I've seen at least one instance where every Notice of Violation had its own folder, and most folders had no more than 3 sheets of paper. So, while the number 1,200 sheets of paper in a 1.2 cu ft box might seem wasteful, it's not an outrageous number to use as an average. In addition, while it's probably fair to say that at least 1/3 of the sheets in a box are two-sided, it simplifies the calculations—at no great cost to the validity of the analysis—to assume all of the sheets are single-sided.
3. Scanning rates assume automatic document feed scanner; all pages are single-sided; no bound volumes; scanner set up time included.
4. Image quality control is based on two-phase method from ANSI/AIIM MS44 "Recommended Practice for Quality Control of Image Scanners", and a sampling method chosen from ANSI/AIIM TR34 "Sampling procedures for inspection by attributes of images in electronic image management (EIM) and micrographics systems". The QC methods consist of screen inspection AND match proof print (screen inspection alone does NOT constitute quality control). In some cases, involving highly-valuable records where the original (paper) records are destroyed after scanning, screen inspection of every image may be required. Four images on a screen does not constitute acceptable image QC. MS44 QC involves two different sampling rates, a higher one for the screen inspection, and a lower one for match proof print (see TR34). Neither the scanning operator nor the indexing data entry operator are allowed to QC the images—there has to be a separate QC workstation and operator, 100% of the time.
5. Index data entry based on average of 3 pages per document, average of 6 fields per document, average of 10 characters per field, total number of characters = 60; time estimate includes error detection and correction. It is assumed that at least 1/3 of the fields have information that is repeated from one entry to the next. There's a trade-off between a rich index (lots of searchable fields) and a thin index (few searchable fields). In my estimation, the latter mandates key verification, which doubles the total number of characters keyed.
6. Document destruction assumes pages will be packed in boxes and destroyed following scanning and quality control.

<b>Production rates (per hour)</b>	<b>Optimistic</b>	<b>Realistic</b>	<b>Pessimistic</b>
Document preparation (pages/hour)	1,200	800	600
Scanning (pages/hour)	2,500	2,000	1,500
Image quality control (pages/hour)	200	150	100
Index data entry (documents/hour)	60	40	20
Document destruction, certified secure (pages/hour)	2,400	2,000	1,500

<b>Labor rate (per hour)</b>	<b>\$12.50</b>	<b>\$15.00</b>	<b>\$20.00</b>

<b>Scanning cost per page</b>	<b>Optimistic</b>	<b>Realistic</b>	<b>Pessimistic</b>
Document preparation	\$0.010	\$0.019	\$0.033
Scanning	\$0.005	\$0.008	\$0.013
Image quality control	\$0.063	\$0.100	\$0.200
Index data entry	\$0.069	\$0.125	\$0.333
Document destruction, certified secure	\$0.003	\$0.004	\$0.005
<b>Total per page</b>	<b>\$0.150</b>	<b>\$0.255</b>	<b>\$0.585</b>

<b>Total per box (1,200 pages)</b>	<b>\$180.31</b>	<b>\$306.06</b>	<b>\$702.00</b>

## Comparison of Paper Storage vs. Scanning Costs - Optimistic Scenario

### Assumptions

1. Total number of boxes	100,000
2. Total number of pages	120,000,000
3. Amount all storage costs increase per year, beginning in year 2	5%

### MOST OPTIMISTIC Comparison (lowest conversion costs vs. highest paper storage costs)

Year	Storage Costs (cumulative)	Storage Costs (annual)	Annual Increase in Storage Costs	Scanning
1	\$4,655,000	\$4,655,000	\$0	<b>\$18,031,333</b>
2	\$5,264,000	\$609,000	\$29,000	\$0
3	\$5,903,450	\$639,450	\$30,450	\$0
4	\$6,574,873	\$671,423	\$31,973	\$0
5	\$7,279,866	\$704,994	\$33,571	\$0
6	\$8,020,109	\$740,243	\$35,250	\$0
7	\$8,797,365	\$777,255	\$37,012	\$0
8	\$9,613,483	\$816,118	\$38,863	\$0
9	\$10,470,407	\$856,924	\$40,806	\$0
10	\$11,370,178	\$899,770	\$42,846	\$0
11	\$12,314,937	\$944,759	\$44,989	\$0
12	\$13,306,933	\$991,997	\$47,238	\$0
13	\$14,348,530	\$1,041,597	\$49,600	\$0
14	\$15,442,207	\$1,093,677	\$52,080	\$0
15	\$16,590,567	\$1,148,360	\$54,684	\$0
16	\$17,796,345	\$1,205,778	\$57,418	\$0
17	<b>\$19,062,412</b>	\$1,266,067	\$60,289	\$0
18	<b>\$20,391,783</b>	\$1,329,371	\$63,303	\$0
19	<b>\$21,787,622</b>	\$1,395,839	\$66,469	\$0
20	<b>\$23,253,253</b>	\$1,465,631	\$69,792	\$0
21	<b>\$24,792,166</b>	\$1,538,913	\$73,282	\$0
22	<b>\$26,408,024</b>	\$1,615,858	\$76,946	\$0
23	<b>\$28,104,676</b>	\$1,696,651	\$80,793	\$0
24	<b>\$29,886,159</b>	\$1,781,484	\$84,833	\$0
25	<b>\$31,756,717</b>	\$1,870,558	\$89,074	\$0
26	<b>\$33,720,803</b>	\$1,964,086	\$93,528	\$0
27	<b>\$35,783,093</b>	\$2,062,290	\$98,204	\$0
28	<b>\$37,948,498</b>	\$2,165,405	\$103,115	\$0
29	<b>\$40,222,173</b>	\$2,273,675	\$108,270	\$0
30	<b>\$42,609,532</b>	\$2,387,359	\$113,684	\$0
31	<b>\$45,116,258</b>	\$2,506,727	\$119,368	\$0
32	<b>\$47,748,321</b>	\$2,632,063	\$125,336	\$0
33	<b>\$50,511,987</b>	\$2,763,666	\$131,603	\$0
34	<b>\$53,413,836</b>	\$2,901,849	\$138,183	\$0
35	<b>\$56,460,778</b>	\$3,046,942	\$145,092	\$0

## Comparison of Paper Storage vs. Scanning Costs - Realistic Scenario

### Assumptions

1. Total number of boxes	100,000
2. Total number of pages	120,000,000
3. Amount all storage costs increase per year, beginning in year 2	5%

### MOST REALISTIC Comparison (mid-range conversion costs vs. mid-range paper storage costs)

Year	Storage Costs (cumulative)	Storage Costs (annual)	Annual Increase in Storage Costs	Scanning
1	\$3,410,000	\$3,410,000	\$0	<b>\$30,606,000</b>
2	\$3,893,000	\$483,000	\$23,000	\$0
3	\$4,400,150	\$507,150	\$24,150	\$0
4	\$4,932,658	\$532,508	\$25,358	\$0
5	\$5,491,790	\$559,133	\$26,625	\$0
6	\$6,078,880	\$587,090	\$27,957	\$0
7	\$6,695,324	\$616,444	\$29,354	\$0
8	\$7,342,590	\$647,266	\$30,822	\$0
9	\$8,022,220	\$679,630	\$32,363	\$0
10	\$8,735,831	\$713,611	\$33,981	\$0
11	\$9,485,122	\$749,292	\$35,681	\$0
12	\$10,271,878	\$786,756	\$37,465	\$0
13	\$11,097,972	\$826,094	\$39,338	\$0
14	\$11,965,371	\$867,399	\$41,305	\$0
15	\$12,876,139	\$910,769	\$43,370	\$0
16	\$13,832,446	\$956,307	\$45,538	\$0
17	\$14,836,569	\$1,004,122	\$47,815	\$0
18	\$15,890,897	\$1,054,328	\$50,206	\$0
19	\$16,997,942	\$1,107,045	\$52,716	\$0
20	\$18,160,339	\$1,162,397	\$55,352	\$0
21	\$19,380,856	\$1,220,517	\$58,120	\$0
22	\$20,662,399	\$1,281,543	\$61,026	\$0
23	\$22,008,019	\$1,345,620	\$64,077	\$0
24	\$23,420,919	\$1,412,901	\$67,281	\$0
25	\$24,904,465	\$1,483,546	\$70,645	\$0
26	\$26,462,189	\$1,557,723	\$74,177	\$0
27	\$28,097,798	\$1,635,609	\$77,886	\$0
28	\$29,815,188	\$1,717,390	\$81,780	\$0
29	<b>\$31,618,447</b>	\$1,803,259	\$85,869	\$0
30	<b>\$33,511,870</b>	\$1,893,422	\$90,163	\$0
31	<b>\$35,499,963</b>	\$1,988,093	\$94,671	\$0
32	<b>\$37,587,462</b>	\$2,087,498	\$99,405	\$0
33	<b>\$39,779,335</b>	\$2,191,873	\$104,375	\$0
34	<b>\$42,080,801</b>	\$2,301,467	\$109,594	\$0
35	<b>\$44,497,341</b>	\$2,416,540	\$115,073	\$0

## Comparison of Paper Storage vs. Scanning Costs - Pessimistic Scenario

### Assumptions

1. Total number of boxes	100,000
2. Total number of pages	120,000,000
3. Amount all storage costs increase per year, beginning in year 2	5%

### MOST PESSIMISTIC Comparison (highest conversion costs vs. lowest paper storage costs)

Year	Storage Costs (cumulative)	Storage Costs (annual)	Annual Increase in Storage Costs	Scanning
1	\$2,165,000	\$2,165,000	\$0	<b>\$70,200,000</b>
2	\$2,522,000	\$357,000	\$17,000	\$0
3	\$2,896,850	\$374,850	\$17,850	\$0
4	\$3,290,443	\$393,593	\$18,743	\$0
5	\$3,703,715	\$413,272	\$19,680	\$0
6	\$4,137,650	\$433,936	\$20,664	\$0
7	\$4,593,283	\$455,633	\$21,697	\$0
8	\$5,071,697	\$478,414	\$22,782	\$0
9	\$5,574,032	\$502,335	\$23,921	\$0
10	\$6,101,483	\$527,452	\$25,117	\$0
11	\$6,655,308	\$553,824	\$26,373	\$0
12	\$7,236,823	\$581,515	\$27,691	\$0
13	\$7,847,414	\$610,591	\$29,076	\$0
14	\$8,488,535	\$641,121	\$30,530	\$0
15	\$9,161,712	\$673,177	\$32,056	\$0
16	\$9,868,547	\$706,836	\$33,659	\$0
17	\$10,610,725	\$742,177	\$35,342	\$0
18	\$11,390,011	\$779,286	\$37,109	\$0
19	\$12,208,261	\$818,251	\$38,964	\$0
20	\$13,067,424	\$859,163	\$40,913	\$0
21	\$13,969,546	\$902,121	\$42,958	\$0
22	\$14,916,773	\$947,227	\$45,106	\$0
23	\$15,911,362	\$994,589	\$47,361	\$0
24	\$16,955,680	\$1,044,318	\$49,729	\$0
25	\$18,052,214	\$1,096,534	\$52,216	\$0
26	\$19,203,574	\$1,151,361	\$54,827	\$0
27	\$20,412,503	\$1,208,929	\$57,568	\$0
28	\$21,681,878	\$1,269,375	\$60,446	\$0
29	\$23,014,722	\$1,332,844	\$63,469	\$0
30	\$24,414,208	\$1,399,486	\$66,642	\$0
31	\$25,883,669	\$1,469,460	\$69,974	\$0
32	\$27,426,602	\$1,542,933	\$73,473	\$0
33	\$29,046,682	\$1,620,080	\$77,147	\$0
34	\$30,747,766	\$1,701,084	\$81,004	\$0
35	\$32,533,904	\$1,786,138	\$85,054	\$0