

SeekEdgarSM

Rajendra P. Srivastava

PhD (Physics) & PhD (Accounting)
EY Professor of Accounting and Information Systems
School of Business, The University of Kansas
Lawrence, KS 66045

Email 1: rsrivastava@ku.edu

Email 2: rsrivastava@seekedgar.com

Cell Phone: 785-218-4409

SeekEdgar website: <https://www.seekedgar.com>

SeekΣdgarSM

2

Introduces

SeekîNFSM

Anytime, Anywhere, ...

SEC Filings and PCAOB Reports
at your finger tips in seconds

For a quick introduction, watch this 6 minutes video at
<https://www.youtube.com/watch?v=wR3TqoQpYQk&feature=youtu.be>



Example 12: Compare

Example 12a: Cosine Similarity Measure

&

Example 12b: Word Variations Year-to-Year

SeekiNFSM

Example 12: Compare: Cosine Similarity Measure and Word Variations

For illustration, we consider two Indian IT companies, Satyam Computer Services Ltd, CIK = 1106056, and WIPRO Ltd., CK = 1123799, both filed 20-F with the SEC. In January 2009, Satyam's Chairman Ramalinga Raju sent a confession letter to SEBI (Securities and Exchange Board of India) admitting that he inflated the financial results and cooked books of Satyam. We would like to compare the readability indices for Satyam, a fraudulent company, with a competing non-fraudulent company, WIPRO Ltd. Here is what we need to do to get the readability indices for 20-F for the period 2002 to 2008.

Step 1: Type “a|the” without the quotes in the slot of “With the exact phrase”. Or, you can use any search criteria including Boolean logic, if that is relevant for your search.

Step 2: Type either “Satyam Computer Services” and select option “Name” from the right or type the CIK code and select “CIK” option from the right. Select the time period 2002 – 2008.

Step 3: Select “Paragraph (All)”.

Step 4: Select 20-F and 20-F/A

Step 5: Submit

You will get the result in seconds. The next slide shows how you enter the above information and the following slide shows the [Compare](#) button at the top right of the display window. Click on this button. The system will open another window. Select the CIK that is already present (CIK of the company just searched) or type in another CIK that you want to get the compare data. Select the filing, in the present case 20-F, and the period 2002-2008 and then “Search”. The system will list all the filings for your compare data. Now you select the specific year filings for the compare data or select all the years. In the present example, I will do it two at a time keeping 2003 year common. This is because I want to get Cosine Similarity measure relative to 2003 Filing. See the subsequent slides for details.

STEP 1: Please enter Phrase(s)/Keyword(s) you wish to search:

With the exact phrase:

The following features are for Single Phrase/Word only: (In case of Multiple Phrases/Words, please use the [Request Form](#) to submit the request.)

Phrase/Word Counter

Display words before & after

Number of words before: Number of words After:

Proximity Search:

Please input within how many words:

With all of the words:

With at least one of the words:

Without the words:

STEP 2: Please select the options below (Default - ALL):

COMPANY:

From: To:

SIC (Optional):

STEP 3: Please select the search Resolution here:

Paragraph(All) Footnote SOX 404 Mgt Report MD&A

Table Audit Report SOX 404 Audit Report

STEP 4: Please select the Document(s) you want to search here:

All

Select (Please [CLICK HERE](#) for Filing Description): Un-Check All

Annual Reports Check All Annual Reports Un-Check All Annual Reports

10-K 10-KT 10-K405 10KT405 10KSB 10KSB40

10-K/A 10-KT/A 10-K405/A 10KT405/A 10KSB/A 10KSB40/A

Quarterly, Current & Foreign Annual Reports

10-Q 10-QT 10QSB 8-K 6-K 20-F 40-F

10-Q/A 10-QT/A 10QSB/A 8-K/A 6-K/A 20-F/A 40-F/A

Proxies & Registrations

DEF 14A DEFM14A S-1 S-4 15-12B 15-12G

DEFA14A DEFM14C S-1/A S-4/A 15-12B/A 15-12G/A

15-12D 15F-12B 15F-12G 15F-12D PRE 14A

15-12D/A 15F-12B/A 15F-12G/A 15F-12D/A

Ownership & Prospectuses

Form 3 Form 4 Form 5 424B1 424B3 424B5

Form 3/A Form 4/A Form 5/A 424B2 424B4 424B7

13F-HR 13F-NT SC 13D SC 13G 424B8

13F-HR/A 13F-NT/A SC 13D/A SC 13G/A

Other Filings

N-CSR N-CSRS NSAR-A NSAR-AT NSAR-B

N-CSR/A N-CSRS/A NSAR-A/A NSAR-AT/A NSAR-B/A

NSAR-BT NSAR-U N-Q SD CORRESP (Comment Letters)

NSAR-BT/A NSAR-U/A N-Q/A SD/A UPLOAD (Response Letters)

AAER**

10-K, 10-Q & 8-K Exhibits & Shareholders Letters

PRESS RELEASE SHAREHOLDERS MEETINGS CONFERENCE CALLS

OTHER 8-K EX SHAREHOLDERS LETTER* EXHIBIT 21 (10-K & 10-Q)

EXHIBIT 95 (10-K & 10-Q) EXHIBIT 10 (10-K) EXHIBIT 10 (10-Q)

EXHIBIT 99 (10-K) EXHIBIT 99 (10-Q) ITEM 1A (10-K)

Public Company Accounting Oversight Board

INSPECTION REPORTS** SETTLED DISCIPLINARY ORDERS**

ADJUDICATED DISCIPLINARY ACTIONS**

*Beta Version: We are working on improving the accuracy of the letters

**AAER & PCAOB do not have CIK's, they have File number and Firm number respectively

STEP 5: SUBMIT

SeekiNFSM

Example 12: Compare: Cosine Similarity Measure and Word Variability

1 - 8 of 8 < >
Download Results

- WIPRO LTD**
 CIK :1123799
 SIC :7371
 File Type :20-F
 File Date :05-30-2008
 Filename :0000950134-08-010518.txt
- WIPRO LTD**
 CIK :1123799
 SIC :7371
 File Type :20-F
 File Date :05-30-2007
 Filename :0000950134-07-012614.txt
- WIPRO LTD**
 CIK :1123799
 SIC :7371
 File Type :20-F
 File Date :06-22-2006
 Filename :0000950134-06-011989.txt
- WIPRO LTD**
 CIK :1123799
 SIC :7371
 File Type :20-F
 File Date :06-13-2005
 Filename :0000950134-05-011880.txt
- WIPRO LTD**
 CIK :1123799
 SIC :7371
 File Type :20-FAMEND
 File Date :05-19-2004
 Filename :0000950134-04-007907.txt
- WIPRO LTD**
 CIK :1123799
 SIC :7371
 File Type :20-F
 File Date :05-17-2004
 Filename :0000950134-04-007736.txt
- WIPRO LTD**
 CIK :1123799
 SIC :7371
 File Type :20-F
 File Date :05-17-2004
 Filename :0000950134-04-007736.txt

View File Download Raw File **Compare**

Class

None align="center" valign="top" nowrap>Name of Each Exchange on Which Registered

Not applicable ABL>

align="center" style="font-size: 10pt; margin-top: 12pt">>Securities registered pursuant to Section 12(g) of the Act:

align="center" style="font-size: 10pt; margin-top: 12pt">>American Depository Shares,

align="center" style="font-size: 10pt; margin-top: 0pt">>each represented by one Equity Share, par value Rs. 2 per share.

align="center" style="font-size: 10pt; margin-top: 12pt">>(Title of Class)

align="center" style="font-size: 10pt; margin-top: 12pt">>Securities for which there is a reporting obligation pursuant to Section 15(d) of

the Act:

align="center" style="font-size: 10pt; margin-top: 12pt">>Not Applicable

align="center" style="font-size: 10pt; margin-top: 12pt">>(Title of Class)

align="left" style="font-size: 10pt; margin-top: 6pt">> Indicate the number of outstanding shares of each of the issuer's classes of capital or common stock as of the close of the period covered by the annual report: 1,461,453,320 **Equity Shares**.

align="left" style="font-size: 10pt; margin-top: 6pt">> Indicate by check mark if the registrant is a well-known seasoned issuer, as defined in Rule 405 of the Securities Act Yes ace="Wingdings"> No ace="Wingdings">

align="left" style="font-size: 10pt; margin-top: 6pt">> If this report is an annual or transition report, indicate by check mark if the registrant is not required to file reports

pursuant to Section 13 or 15(d) of the Securities Exchange Act, 1934 Yes ace="Wingdings"> No ace="Wingdings">

align="left" style="font-size: 10pt; margin-top: 6pt">> Indicate by check mark whether the registrant (1) has filed all reports required to be filed by Section 13 or 15(d) of the Securities Exchange Act of 1934 during the preceding 12 months (or for such shorter period that the registrant was required to file such reports), and (2) has been subject to such filing requirements for the past 90 days. Yes ace="Wingdings"> No ace="Wingdings">

align="left" style="font-size: 10pt; margin-top: 6pt">> Indicate by check mark whether the registrant is a large accelerated filer, an accelerated filer, or a non-accelerated filer. See definition of "accelerated filer and large accelerated filer" in Rule 12b-2 of the Exchange Act.

align="center" style="font-size: 10pt">> ABL>

" valign="top" style="font-size: 10pt; margin-top: 12pt">> ace="Wingdings"> align="center" valign="top" style="font-size: 10pt; margin-top: 12pt">>International Financial reporting Standards as used by the International Accounting Standards Board align="center" valign="top" style="font-size: 10pt; margin-top: 12pt">> ace="Wingdings"> align="center" valign="top" style="font-size: 10pt; margin-top: 12pt">>Others align="center" valign="top" style="font-size: 10pt; margin-top: 12pt">> ace="Wingdings"> ABL>

align="left" style="font-size: 10pt; margin-top: 6pt">> If this is an annual report, indicate by check mark whether the registrant is a shell company (As defined in Rule 12b-2 of the Securities Exchange Act of 1934. Yes ace="Wingdings"> No ace="Wingdings">

Word Distribution:

Download Distribution

FILE ANALYSIS	
Total Word Count	93966
Total Word Count without numerics	85584
Sentence Count	3624
Gunning-Fog Index	19.2
Smog Index	17.059
Flesch Reading Ease	28.928
Flesch-Kincaid Grade Level	15.113
Automated Readability Index	15.082
Coleman-Liau Index	14.04

CIK :

[Change CIK](#)

File Type:

Year : to

[search](#)

Cosine Similarity

Word Variation

The list of 20-F 's filed for the selected years

Filing	Description	Filing Date	Word Distribution
<input checked="" type="checkbox"/> 20-F	Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001145549-08-001441 (34 Act) Size: 1 MB	2008-08-08	Download Distribution
<input type="checkbox"/> 20-F	Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001145549-07-000670 (34 Act) Size: 1 MB	2007-04-30	Download Distribution
<input type="checkbox"/> 20-F	Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001145549-06-000579 (34 Act) Size: 1 MB	2006-04-28	Download Distribution
<input type="checkbox"/> 20-F	Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001145549-05-000706 (34 Act) Size: 765 KB	2005-04-28	Download Distribution
<input type="checkbox"/> 20-F/A	[Amend] Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001145549-04-000961 Size: 45 KB	2004-07-26	Download Distribution
<input type="checkbox"/> 20-F	Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001145549-04-000850 Size: 1 MB	2004-06-29	Download Distribution
<input checked="" type="checkbox"/> 20-F	Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001145549-03-000795 Size: 1 MB	2003-06-30	Download Distribution
<input type="checkbox"/> 20-F	Annual and transition report of foreign private issuers [Sections 13 or 15(d)] Acc-no: 0001021408-02-009039 Size: 1 MB	2002-07-01	Download Distribution

SeekiNFSM

Example 12a: Cosine Similarity Measure Data for Satyam and WIPRO

Once you click on the [Compare](#) button after selecting the two years filings, the system will yield the result in a CSV file which you can open and save in Excel. Repeating this process for each pair, (2003, 2008), (2003, 2007), (2003, 2006), (2003, 2005), (2003, 2005), and (2003, 2004), I obtained the following data.

Cosine Similarity Scores 20F		File 1 VS File 2					
Satyam	CIK	2008 VS 2003	2007 VS 2003	2006 VS 2003	2005 VS 2003	2004 20F/A VS 2003	2004 VS 2003
	1106056	0.986219031	0.986518731	0.986687789	0.983219142	0.884167909	0.995569951
WIPRO	CIK	2008 VS 2003	2007 VS 2003	2006 VS 2003	2005 VS 2003	2004/A VS 2003	2004 VS 2003
	1123799	0.992107956	0.993143381	0.993785122	0.992941528	0.802915992	0.996115982

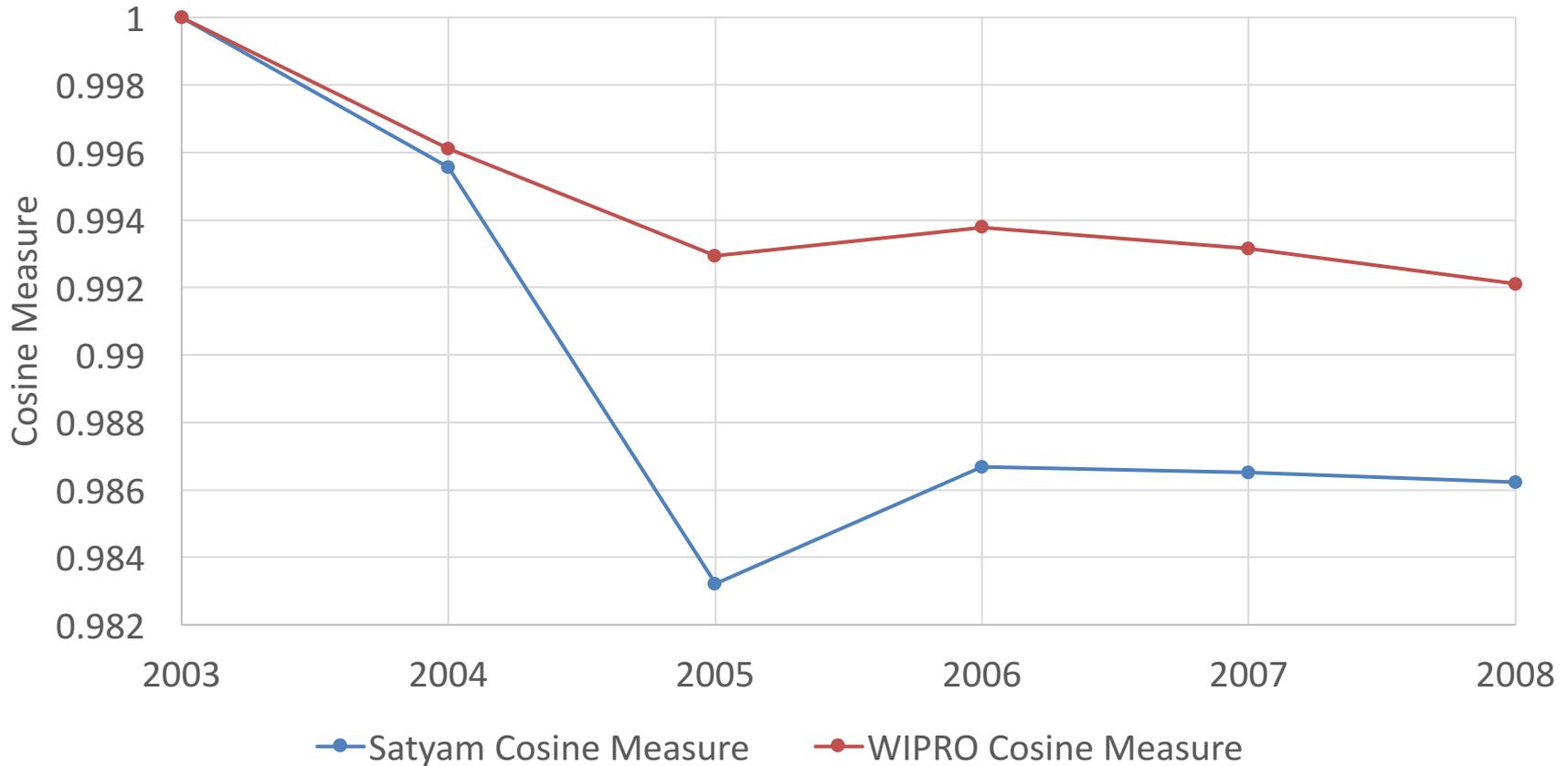
The following data was obtained by rearranging the above data. See the next slide for the plots

Year	Satyam Cosine Measure	WIPRO Cosine Measure
2008	0.986219031	0.992107956
2007	0.986518731	0.993143381
2006	0.986687789	0.993785122
2005	0.983219142	0.992941528
2004	0.995569951	0.996115982
2003	1.0	1.0



Example 12a: Graph of Cosine Similarity Data for Satyam and WIPRO

Cosine Measure of Similarity for Satyam with WIPRO in relation to 2003 20-F





Example 12: Compare - Cosine Similarity Measure and Word Variability

Example 12b:

Compare:

Word Variation Year-to-Year

SeekiNFSM

Example 12b: Word Variations Year-to-Year

- Let us consider the example of Satyam, again. In Example 12a, we selected “Cosine Similarity”. For this example, after selecting two years 20-F, we select “Word Variation”. The system will yield the data in a csv file that you can open and save in Excel.
- Next slide shows the un-sorted data for the years 2007 and 2008.
- The second slide shows, after sorting, the words that are added in 20-F for 2008 that were not present in 20-F of 2007.
- The third slide shows the words (after sorting again) that are removed from 20-F of 2008, but were present in 20-F of 2007.
- One can sort the file and obtain changes in certain key words from year to year.

SeekiNFSM

WORD VARIATIONS FOR 1106056

WORDS	2008	2007
begin	1	1
privacy	3	3
enhanced	6	6
message	2	2
proc	1	1
type	14	15
2001,mic	1	1
clear	1	1
originator	2	2
name	16	20
webmaster	1	1
www.sec.gov	2	2
key	37	36
asymmetric	1	1
mfgwgcgyevqgbaqicaf8d	1	1
wsm7vrzladbmyqaiolv	1	1
o	55	59
widaqab	1	1
mic	1	1
info	1	1
sa	1	1

Example 12b: Word Variations Year-to-Year

un-sorted data for Satyam for the years 2007 and 2008

SeekiNFSM

Example 12b: New Words in 20-F of 2008 that were not present in 20-F of 2007 for Satyam

WORD VARIATIONS FOR 1106056																				
	https://www.sec.gov/Archives/edgar/data/1106056/0001145549-07-000670-index.html																			
WORDS	2008	2007																		
nitor	22	0																		
afm	13	0																		
bridge	10	0																		
euronext	9	0																		
belts	6	0																		
deconsolidation	6	0																		
lieu	6	0																		
pharma	6	0																		
hetzel	5	0																		
caterpillar	4	0																		
chicago	4	0																		
conditional	4	0																		
contributed	4	0																		
fbt	4	0																		
fmsa	4	0																		
infocity	4	0																		

Words in 20-F of 2008 that are not present
in 20-F of 2007 for Satyam

SeekiNFSM

Thanks!

Email: rsrivastava@seekedgar.com