

# Spreadsheet Auditing and Change Analysis

**John Nash**

([jcnash@uottawa.ca](mailto:jcnash@uottawa.ca))

**Neil Smith**

([neil.f.smith@sympatico.ca](mailto:neil.f.smith@sympatico.ca))

**Andy Adler**

([adler@site.uottawa.ca](mailto:adler@site.uottawa.ca))

# Initial Motivations

- Course mark management:
  - Multiple teaching assistants submit marks as spreadsheet files after each assignment
  - Merge into master
  - Re-submissions after the merge
- Easier methods to review and check spreadsheet models for “bugs”

# Spreadsheet Auditing

## - a 2-part task

- A tool and technique that presents the status and changes in a spreadsheet file
- A Server-based system to ensures overall integrity and continuous record of changes in all relevant files

# SSScan – the audit tool

- Tool for auditing spreadsheet files
  - Change tracking + content analysis
- Reads OpenOffice.org **Calc** file format
  - Underlying XML structure
  - *Calc* Reads/writes Excel files (up to 97/2000)
  - Open Source taking hold in various industry sectors
  - Java stand-alone program
    - Multi-platform operation (Windows/Mac/Unix)
    - Auditor / user not required to have *Calc*

# Change display panel

C:\WINDOWS\Desktop\TestFiles\CashFlow\_v1.xlsx

Changes Filters Content Report

187 Change Records between: 2003-03-28 and 2003-03-28

Change	Sheet	Address	Author	Date	Time	Status	Change Details
Cell content	Cash Flow ...	K22	Neil Smith	2003-03-...	21:51:18	(U...pending <empty> -> =K8-K18-K20	{ \$5,150 (currency)}
Cell content	Cash Flow ...	L22	Neil Smith	2003-03-...	21:51:18	(U...pending <empty> -> =L8-L18-L20	{ \$7,650 (currency)}
Cell content	Cash Flow ...	M22	Neil Smith	2003-03-...	21:51:18	(U...pending <empty> -> =M8-M18-M20	{ -\$139,850 (currency)}
Cell content	Cash Flow ...	N22	Neil Smith	2003-03-...	21:51:35	(U...pending =SUM(B22:M22)	{ 0 (float) -> =N8-N18-N20 { \$202,602 (curr...
Insertion	Cash Flow ...	17	Neil Smith	2003-03-...	21:52:03	(U...pending 1 row at row 17	
Cell content	Cash Flow ...	A17	Neil Smith	2003-03-...	21:52:07	(U...pending <empty> -> Travel	(string)
Cell content	Cash Flow ...	N17	Neil Smith	2003-03-...	21:52:30	(U...pending <empty> -> =SUM(B17:M17)	{ \$3,600 (currency)}
Cell content	Cash Flow ...	B18	Neil Smith	2003-03-...	21:56:57	(U...pending =SUM(B11:B16)	{ 0 (float) -> =SUM(B11:B17) { \$25,300 (curr...
Cell content	Cash Flow ...	C18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(C11:C16)	{ 0 (float) -> =SUM(C11:C17) { \$11,100 (curr...
Cell content	Cash Flow ...	D18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(D11:D16)	{ 0 (float) -> =SUM(D11:D17) { \$11,100 (curr...
Cell content	Cash Flow ...	E18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(E11:E16)	{ 0 (float) -> =SUM(E11:E17) { \$11,100 (curr...
Cell content	Cash Flow ...	E18	Neil Smith	2003-03-...	21:50:46	(U...pending <empty> -> =SUM(E11:E16)	{ 0 (float)}
Cell content	Cash Flow ...	F18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(F11:F16)	{ 0 (float) -> =SUM(F11:F17) { \$37,350 (curr...
Cell content	Cash Flow ...	F18	Neil Smith	2003-03-...	21:50:46	(U...pending <empty> -> =SUM(F11:F16)	{ 0 (float)}
Cell content	Cash Flow ...	G18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(G11:G16)	{ 0 (float) -> =SUM(G11:G17) { \$37,350 (curr...
Cell content	Cash Flow ...	H18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(H11:H16)	{ 0 (float) -> =SUM(H11:H17) { \$33,600 (curr...
Cell content	Cash Flow ...	I18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(I11:I16)	{ 0 (float) -> =SUM(I11:I17) { \$33,600 (currency)}
Cell content	Cash Flow ...	J18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(J11:J16)	{ 0 (float) -> =SUM(J11:J17) { \$44,850 (curr...
Cell content	Cash Flow ...	K18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(K11:K16)	{ 0 (float) -> =SUM(K11:K17) { \$59,850 (curr...
Cell content	Cash Flow ...	L18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(L11:L16)	{ 0 (float) -> =SUM(L11:L17) { \$67,350 (curr...
Cell content	Cash Flow ...	M18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(M11:M16)	{ 0 (float) -> =SUM(M11:M17) { \$74,850 (curr...
Cell content	Cash Flow ...	N18	Neil Smith	2003-03-...	21:57:03	(U...pending =SUM(N11:N16)	{ 0 (float) -> =SUM(N11:N17) { \$447,398 (cu...
Cell content	Cash Flow ...	B5	Neil Smith	2003-03-...	22:00:44	(U...pending <empty> -> \$100,000	(currency)

# of Change Types: 2  
182 Cell content

# of Authors: 1  
187 by "Neil Smith"

# of Dates: 1  
187 on "2003-03-28"

# Filter panel

Change Record Filter Settings

Author Contents Date Address

Action	Before	After
exclude	<empty>	don't care
exclude	static value	static value
disable	don't care	don't care
disable	don't care	don't care
disable	don't care	don't care
disable	don't care	don't care

Active Filter Summary:  
Exclude: Contents= "<empty>" -> "don't care", "static value" -> "static value"

# Spreadsheet content panel

C:\WINDOWS\Desktop\testFiles\CashFlow\_v1.xlsx

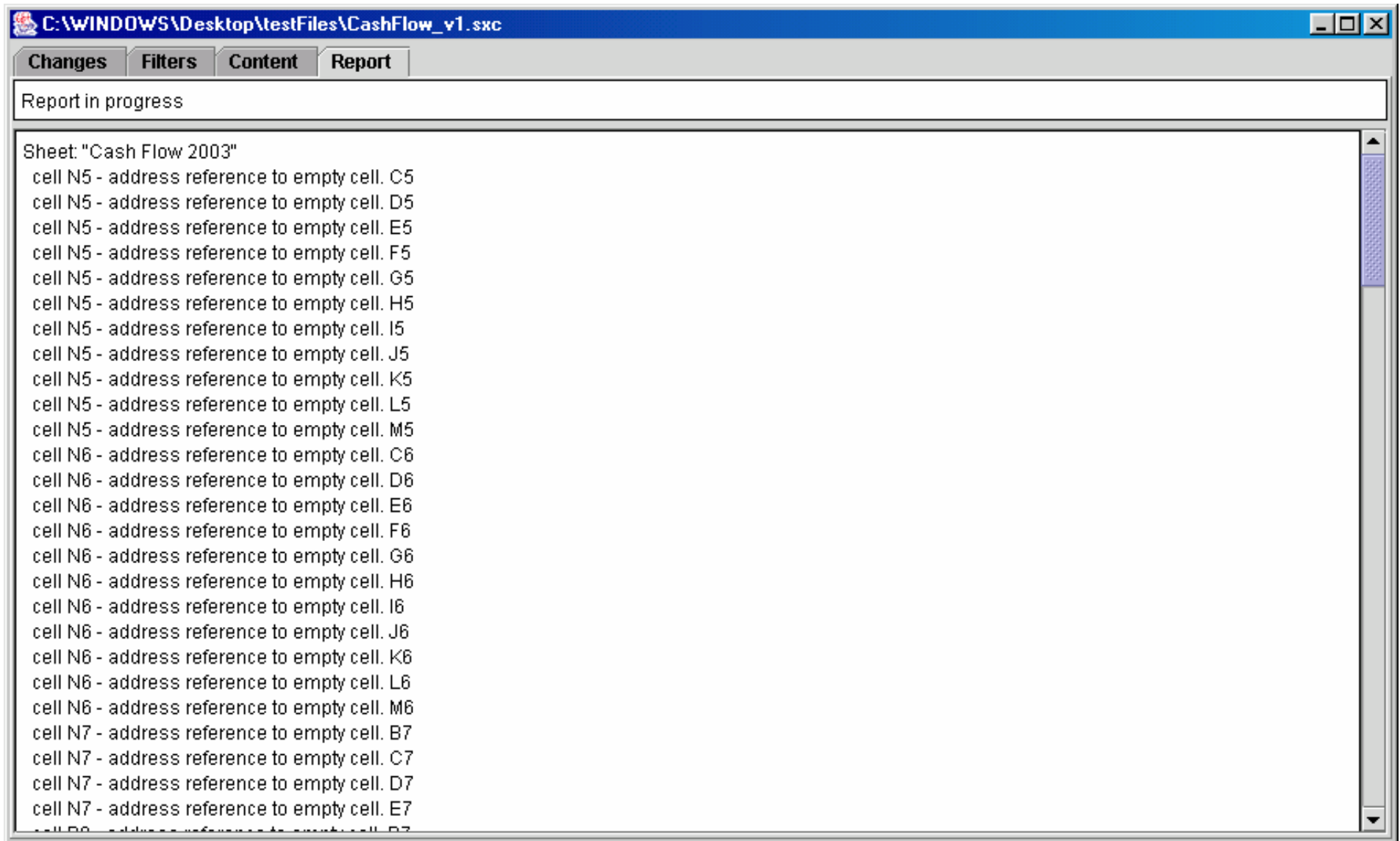
Changes Filters **Content** Report Tree

	A	B	C	D	E	F	G	H	I	J	K
1	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
2	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
3	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
4	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
5	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
6	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
7	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
8	(u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU
9	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
10	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
11	(u/p)	(u/p)	PM (u/p)	PM (u/p)	PM (u/p)	PM (u/p)	PM (u/p)	PM (u/p)	PM (u/p)	PM (u/p)	PM
12	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
13	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
14	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
15	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
16	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
17	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
18	(u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU (u/p)	SU
19	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
20	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
21	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	
22	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	(u/p)	

Cash Flow 2003 Sheet2 Sheet3

Cell E11: Contents: =-PMT(6%/12;60;1B6;0)  
 Value: 5,800 (currency)  
 Cell Protection Setting: protected

# Spreadsheet Report panel





# Server-based system

- Need to protect the file(s), control access
- Many issues
  - Efficient operation
  - Minimal fuss and bother for user
  - Maximal security and freedom from errors
  - Version history and checkpointing
  - Cross-platform usage

# Server architecture

- Graphics in VNC (virtual networked computing)
- Apache server
- Highly customized configuration of oocalc
- Perl scripts to configure and glue together components
- Web based interface to allow cross-platform use

# Server architecture ( cont.)

- Runs own password/user base
- CVS back-end to provide for versioning and checkpointing
- DB\_File::Lock file locking
- Many “fixes” needed to components
- Advantages and obstacles with each choice made in the implementation...

SSHEET User Login - Mozilla 1.0.1

File Edit View Go Bookmarks Tools Window Help

http://localhost/cgi-bin/ssheet/login.pl

Search



# SSHEET User Login

## Enter your identification information

Userid?

Password?

Login

---

Note: Your browser must allow a cookie to be set to permit the login.

---

**SSHEET file access to auditable spreadsheet files** is Copyright © 2003 by [Prof. J. C. Nash](#) and others,  
[School of Management, University of Ottawa](#).

Note: SSHEET uses a **cookie** to temporarily store login information. This expires after a few minutes any you will then have to log in again.

Document: Done (0.599 secs)

Launch Shell No 2 - Kon... SSHEET User L... SSHEET User L... Screen Capture 1 2 3 4 02:34 pm



## SSheet -- Auditable Spreadsheet Access (**john**)

**Choose file to open (those in RED are in use by others)**

**Files:**

- [Complex1-1.sxc](#)
- [CashFlow v1.sxc](#)
- [test.sxc](#)
- [FormulaTest.sxc](#)
- **portfolio.sxc** in use by roland
- [RowInsertTest.sxc](#)

---

[Update Status](#)

[Logout](#)

---

*SSHEET file access to auditable spreadsheet files* is Copyright © 2003 by [Prof. J. C. Nash](#) and others, [School of Management, University of Ottawa](#).

Note: SSHEET uses a **cookie** to temporarily store login information. This expires after a few minutes any you will then have to log in again.

USER: john, FILE: test.sxc, After File->Exit from Spreadsheet [click here to exit](#)

File Edit View Insert Format Tools Data Window Help

Tahoma 10 B i U

A7 Date

	A	C	E	F	G	H
1	Internet Institute - Ottawa					
2	<b>Instructor Schedule</b>					
3	<b>John Nash</b>					
4	18 Nov 02					
5						
6						
7	<u>Date</u>	<u>Appointment</u>	<u>Status</u>	<u>Start Time</u>	<u>End Time</u>	<u>Facilities</u>
8	Nash, John					
9	03/18/2003	Practical Forecasting for Managers	Pending	9:00 AM	4:30 PM	Classroom A, Internet Ins
10	06/11/2003	Practical Forecasting for Managers	Pending	9:00 AM	4:30 PM	Classroom A, Internet Ins
11						
12						
13						
14						
15						
16						

Schedule Sheet2 Sheet3



# files/

Current directory: [\[ssheet\]](#) / files

File	Rev.	Age	Last log entry
<a href="#">Parent Directory</a>			
<a href="#">CVS/</a>			
<a href="#">john/</a>			
<a href="#">quon/</a>			
<a href="#">roland/</a>			
<a href="#">CashFlow_v1.sxc</a>	<a href="#">1.1</a>	47 hours	Initial Version: setup
<a href="#">Complex1-1.sxc</a>	<a href="#">1.1</a>	47 hours	Initial Version: setup
<a href="#">FormulaTest.sxc</a>	<a href="#">1.1</a>	47 hours	Initial Version: setup
<a href="#">RowInsertTest.sxc</a>	<a href="#">1.5</a>	46 hours	Modified by quon
<a href="#">portfolio.sxc</a>	<a href="#">1.1</a>	47 hours	Initial Version: setup
<a href="#">test.sxc</a>	<a href="#">1.1</a>	47 hours	Initial Version: setup

Show only files with tag:  Module path or alias:

# Status and prognosis

- SSScan is working and quite stable
  - But filter customization is important to make audit efficient for particular situations
- Server configuration and setup is tricky
  - Large number of disparate components
  - Many customizations for efficiency
  - Interaction with OS and network
  - Some weaknesses, but with known “fixes”



# Status (cont.)

- Need a viable business model
- Current view:
  - “Package” customization of SSScan and configured server hardware
  - Release each customization as Open Source, but delay installer until experience is gained

# Structure: Apache – SSL

- Config
  - SSL
  - Directories
  - httpd.conf
- Runs as www-data, member of group ssheet
- Cookies
- Login --> database of users/passwords & files/permissions

# Apache-SSL cont.

- File locking & versioning
  - » --> ssheet userdb {users & shuids & files}
  - » ssheet-userdb.lock
  - Copy file to indir
  - Grow user name
  - (select new random password)
  - Create web page
    - » Active content for vnc
    - » Custom vncviewer.jar
  - Check in file in outdir (CVS manages “real” changes)

# SHUIDs: a “farm” of pseudo-users

- Runs as `ssheet-uid****` (`start-ssheet.sh` opens in `sshuid's`)
- Member of `www-data`
- Start `vncserver` --> `xstartup` (perl script)
- Wait on `indir` (group read)
- Copy to `procdir` (shuid only)
- Edit user
- (insert new password for vnc)
- `Oocalc` initiated with special profile
- (mechanism to kill `oocalc`)
- Move file to `outdir`
- (kill password)

# CVS Server

- /var/www/ssheet/repository
- INIT:
  - » Take each file and check in
- CHECKIN by shuids
  - » Comment string from www-data
- Web interface to versioning
  - » Modified CVSWeb accesses repository

# Admin Interface (to do)

- Add new files to repository
  - » Verify change recording is “on”
- Sophisticated import
  - » Set user name
  - » Set previous history
- Handling of “duplicate” files
- Handling of different permissions for different users
- User creation, modification, deletion