XXtoCSV.jar

Утилита конвертации отчетов RS-Bank в CSV формат.

Описание.

XXtoCsV.jar - маленькая универсальная утилита командной строки для конвертации отчетов сформированных AБC RS-Bank в CSV формат, для возможности последующего просмотра в табличном редакторе. Т.к. утилита написана на JAVA, для удобства ее использования применяется пакетный командный файл **RUN_XXtoCsV.cmd**

История версий.

XXtoCSV.jar v 1.4 — добавлена возможность поиска значений, на определенную глубину вперед.

XXtoCSV.jar v 1.3 — добавлена возможность заполнять столбцы историческими записями.

XXtoCSV.jar v 1.2 — добавлено принудительное указание числовых и строковых столбцов.

XXtoCSV.jar v 1.1 — добавлена возможность использования синонимов поисковых строк.

XXtoCSV.jar v 1.0 — первый релиз.

Примеры использования.

Утилита требует 2 или 3 параметра командной строки.

```
1-ый параметр — (обязательный) исходный файл отчета.
```

2-ый параметр — (обязательный) файл конфигурации.

3-ый параметр — (**не обязательный**) выходной файл, в данном случае имя выходного файла формируется из имени входного с добавлением суффикса «-out» и расширения «.csv».

```
java.exe -jar XXtoCSV.jar ReportF8.102 F8.csv
java.exe -jar XXtoCSV.jar ReportF8.102 F8.csv F8tabOUT.csv
```

Для удобства использования утилиты применяется пакетный командный файл RUN_XXtoCSV.cmd, который автоматически находит и отдает на обработку все предопределенные файлы отчетов.

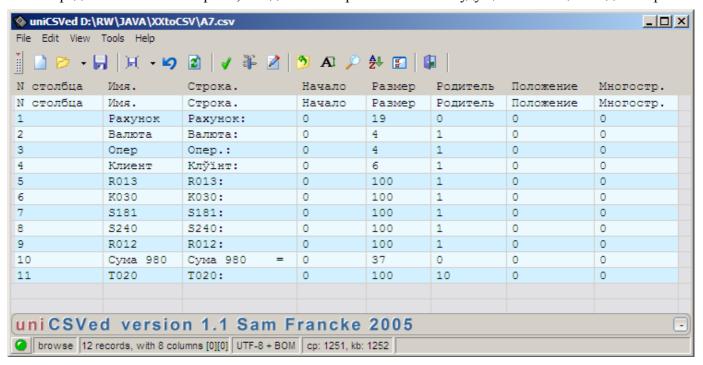
Настройка.

Для хранения настроек преобразования применяются конфигурационные файлы, в которых указаны параметры столбцов выходного файла. Посредствам файла настроек, возможно определить критерии преобразования исходного файла отчета.

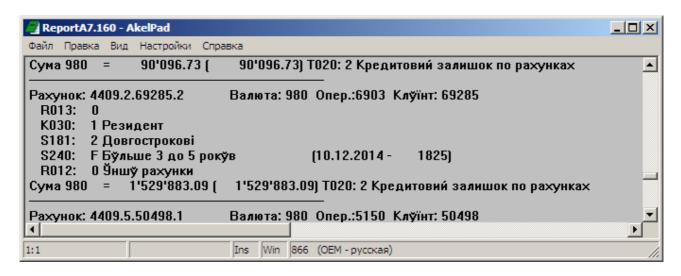
Структура конфигурационного файла.

ВАЖНО. Кодировка конфигурационного файла должна быть ЮНИКОД (UTF-8).

Файл представляет собой строки, каждая из которых описывает будущий столбец выходного файла.



Для примера рассмотрим файл конфигурации для преобразования A7 отчета. Структура отчета файла отчета AБС RS-Bank такова -

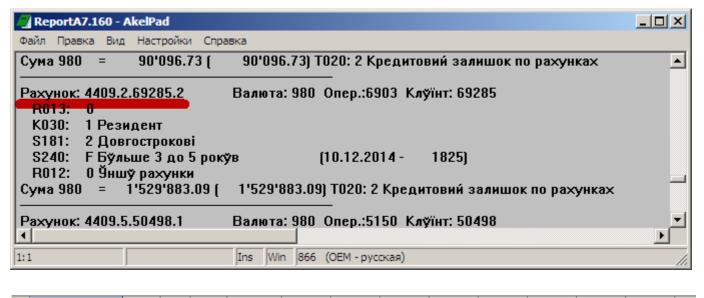


Что мы хотим увидеть на выходе?

	А	В	С	D	Е	F	G	Н		J	K	L	M	
1	Рахунок	Валюта	Опер	Клиент	R013	K030	S181	S240	R012	Сума 980	T020			
2	1410.8.10250.6/4	980	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	10'000'000	1 Дебетов	ий залишо	к по рахуні	ках
3	1410.7.10250.3/840	840	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	20'043'805	1 Дебетов	ий залишо	к по рахуні	ках
4	1415.3.10250.6/4	980	1200	10250	7 Переоцўн	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	7'255.72 (2 Кредито	вий залиш	ок по рахун	нках
5	1416.4.10250.6/4	980	1200	10250	7 Неаморт.,	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	227'944.28	2 Кредито	вий залиш	ок по рахун	нках
6	1416.3.10250.3/840	840	1200	10250	7 Неаморт.,	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	582'037.69	2 Кредито	вий залиш	ок по рахун	нках
7	1418.5.10250.3/840	840	1200	10250	1 Нарахова	1 Резиден	1 Коротко	6 Від 32 д	4 Рахунки	530'120.17	1 Дебетов	ий залишо	к по рахуні	ках
8	1420.5.10250.1	980	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	600.00 (1 Дебетов	ий залишо	к по рахуні	ках
9	1420.6.10250.1/840	840	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	15'646'999	1 Дебетов	ий залишо	к по рахуні	ках
11	1/1/26 1 10/250 1	980	1200	10250	7 Незмолт	1 Резилен	1 Κοροτνοι	1 На вимо	3 Рэуушин	197 25 /	2 Крелито	вий запини	ov no navvi	uvav

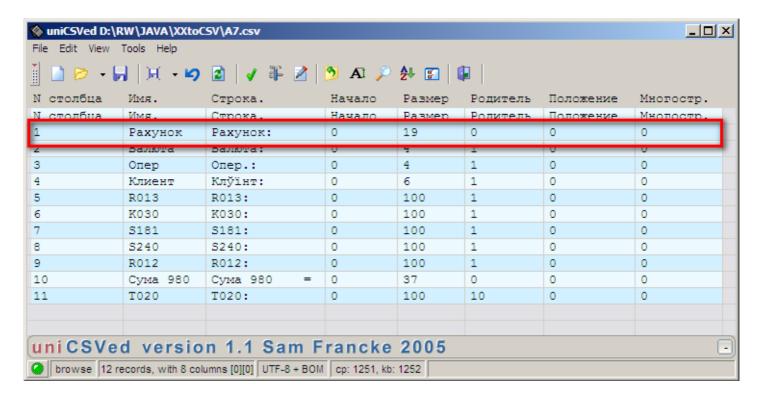
Т.е. Необходимо значение строк записать в ячейки соответствующих столбцов. Что утилита **XxtoCSV.jar** и сделает, если ее научить.

Первое, что надо сделать, определить самую первую корневую запись, от которой будет начинается отсчет. Наиболее подходит строка **Рахунок:**, т.е. по нахождению слова "**Рахунок:**" будет начинаться формирование новой записи. Корневую запись надо выбирать таким образом, чтобы она присутствовала всегда.



	Α	В	С	D	E	F	G	H	1	J	K	L	M	
1	Рахунок	Валюта	Опер	Клиент	R013	K030	S181	S240	R012	Сума 980	T020			
2	1410.8.10250.6/4	980	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	10'000'000	1 Дебетов	ий залишо	к по рахунк	(ax
3	1410.7.10250.3/840	840	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	20'043'805	1 Дебетов	ий залишо	к по рахунк	(ax
4	1415.3.10250.6/4	980	1200	10250	7 Переоцўн	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	7'255.72 (2 Кредито	вий залиш	ок по рахун	нках
5	1416.4.10250.6/4	980	1200	10250	7 Неаморт.,	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	227'944.28	2 Кредито	вий залиш	ок по рахун	нках
6	1416.3.10250.3/840	840	1200	10250	7 Неаморт.,	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	582'037.69	2 Кредито	вий залиш	ок по рахун	ках
7	1418.5.10250.3/840	840	1200	10250	1 Нарахова	1 Резиден	1 Коротко	6 Від 32 д	4 Рахунки	530'120.17	1 Дебетов	ий залишо	к по рахунк	(ax
8	1420.5.10250.1	980	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	600.00 (1 Дебетов	ий залишо	к по рахунк	(ax
9	1420.6.10250.1/840	840	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	15'646'999	1 Дебетов	ий залишо	к по рахунк	(ax
10	1/126 1 10250 1	980	1200	10250	7 Неаморт	1 Резилен	1 Κοροτνοί	1 На вимо	3 Рауушии	197 25 /	2 Крелито	оий ээпини	OV DO DOVIJE	JVOV

"Рахунок" - будет первый столбец, а значит надо сформировать запись конфигурационного файла.



Поля, которые надо заполнить, таковы –

N столбца;Положение;Имя.;Многостр.;Строка.;Формат;Начало;Заполнять;Размер;Искать до.

Родитель;

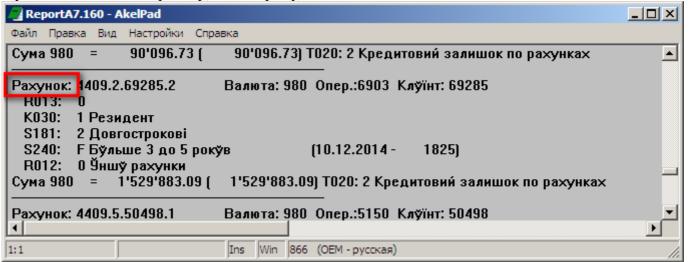
Что они означают?

N столбца — просто порядковый номер строки, программой не читается, и служит для нашего удобства, при указании родительских полей, о родительских полях далее.

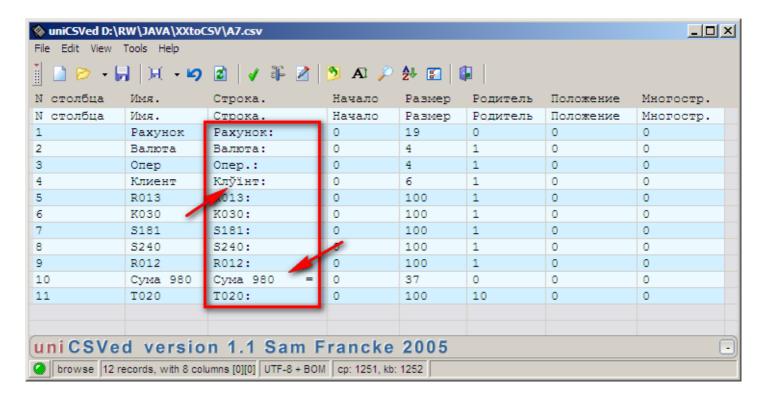
Имя. - желаемое имя столбца, отображаемое в заголовке выходного файла.

		А	R	С.	D	F	F	G	H			K	_ L	M	
	1	Рахунок	Валюта	Опер	Клиент	R013	K030	S181	S240	R012	Сума 980	T020			
	2	14 10.0. 10250.6/4	960	1200	10250	9 унше	і Резиден	і Короткос	і на вимо	3 Рахунки	10 000 000	і Дебетови	и залишон	спо рахунка	ax
	3	1410.7.10250.3/840	840	1200		9 Ўнше						1 Дебетови			
	4	1415.3.10250.6/4	980	1200	10250	7 Переоцўн	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	7'255.72 (2 Кредитов	ий залишо	ок по рахун	ках
	5	1416.4.10250.6/4	980	1200	10250	7 Неаморт.,	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	227'944.28	2 Кредитов	ий залишо	к по рахун	ках
	6	1416.3.10250.3/840	840	1200	10250	7 Неаморт.,	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	582'037.69	2 Кредитов	ий залишо	ок по рахун	ках
	7	1418.5.10250.3/840	840	1200	10250	1 Нарахова	1 Резиден	1 Коротко	6 Від 32 д	4 Рахунки	530'120.17	1 Дебетови	й залишон	с по рахунка	ax
	8	1420.5.10250.1	980	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	600.00 (1 Дебетови	й залишон	с по рахунка	ax
	9	1420.6.10250.1/840	840	1200	10250	9 Ўнше	1 Резиден	1 Коротко	1 На вимо	3 Рахунки	15'646'999	1 Дебетови	й залишон	с по рахунка	ax
,	10	1/106 1 10060 1	920	1200	10250	7 Незморт	1 Резилен	1 Konotvor	1 На вимо	3 Davinini	197 25 /	2 Кралитов	ий ээпини	שעעפת חם ער	vev

Строка. - поисковый параметр, по которому, мы находим значение.

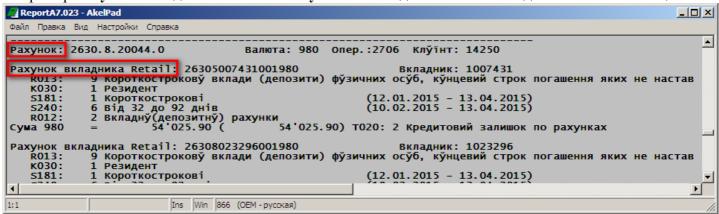


ВАЖНО - данный параметр необходимо указывать явно со всеми пробелами и непечатными символами.



Начиная с **XXtoCsV.jar v 1.1** поисковый параметр может иметь синонимы. Что это значит? Значит мы можем применять несколько поисковых слов, результаты которых будут однозначно заносится в указанный столбец.

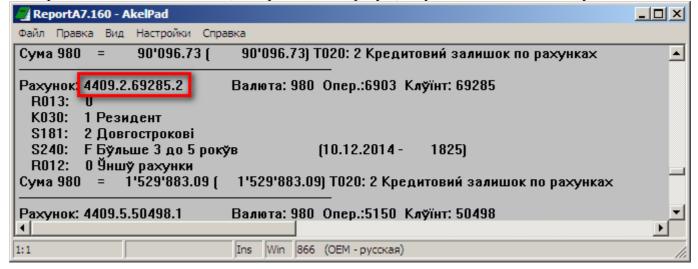
Например **Рахунок вкладника Retail:** и **Рахунок:** необходимо заносить в один и тот же столбец.



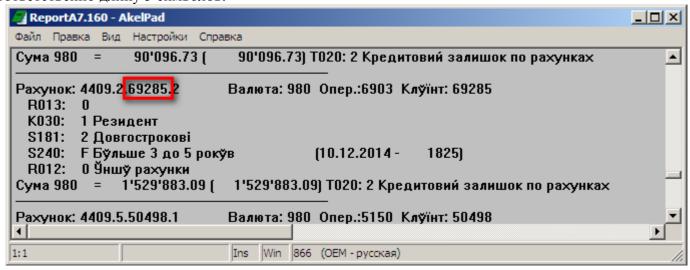
Синонимы разделяются в строке 3-мя символами <@>

	:\RW\JAVA\XXtoCSV\A7.csv			
File Edit View	/ Tools Help			
1 🗋 🗁 🕶	□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □	🖹 📝 🤼 Al 🔑 🛃 🖫		
N столбца	Имя.	Строка.	Начало	Pa
N столбца	Имя.	Строка.	Начало	Pa
1	Рахунок	Paxyнoк:<@>Paxyнoк вкладника Retail:	0	19
2	Бал.Рах.	Рахуно	0	5
3	Валюта	Валюка:	0	4
4	Опер	Onep.:	0	4
5	Клиент/Вкладник	Клўїнт:<@>Вкладник:	0	7
6	R013	R013:	0	10
7	K030	K030:	0	10
8	S181	S181.	0	10
9	5240	S240:	0	10
10	R012	R012:	0	10
11	Залишок в ОДБ	Залишок в ОДБ:	0	16
12	Залишок в Retail заг	в Retail :	0	16
13	Залишок в Retail	в Retail :	18	16
14	розбіжність	розбужнусть	0	5
15	Сума 980 заг.	Сума 980 =	0	17
16	Сума 980	Сума 980 =	19	17
17	T020	T020:	0	10

Размер — количество символов, которое нас интересует, и будет вносится в таблицу.



Начало — количество символов, пропускаемое после ключевого слова, т.е. можно вырезать из строки, что нас интересует. Например, можно вырезать код клиента указав начало с 8-го символа, и соответственно длину 5 символов.



Родитель — если названия полей повторяются, то надо указывать кто их родитель, т.е. Для дочерней строки номер ее родительской строки, а родительскую строку помечать значением 0.

0 — в поле родитель значит, что поле будет иметь таковое дочерние.

🄷 uniCSVed D:\F	RW\JAVA\XXtoC	SV\A7.csv					_ 🗆 ×			
File Edit View	Tools Help									
📗 🗋 📂 🗣] H • 10	2 🗸 🕸	🐧 Al 🔑	Å ↓ 😰						
N столбца	MMM.	Строка.	Начало	Размер	Родитель	Положение	Многостр.			
N столбца	Имя.	Строка.	Начало	Размер	Родитель	Положение	Многостр.			
1	Рахунок	Рахунок:	0	19	0	0	0			
2	Валюта	Валюта:	0	4	1	0	0			
3	Опер	Опер.:	0	4	1	0	0			
4	Клиент	Клўінт:	0	6	1	0	0			
5	R013	R013:	0	100	1	0	0			
6	K030	K030:	0	100	1	0	0			
7	S181	S181:	0	100	1	0	0			
8	S240	S240:	0	100	1	0	0			
9	R012	R012:	0	100	1	0	0			
10	Сума 980	Сума 980 =	0	37	0	0	0			
11	T020	T020:	0	100	10	0	0			
uniCSVe	d versio	n 1.1 Sam F	rancke	2005			-			
browse 12 records, with 8 columns [0][0] UTF-8 + BOM cp: 1251, kb: 1252										

Примечание.

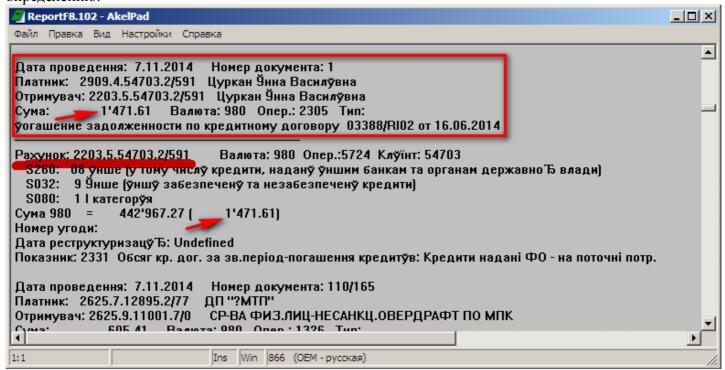
Родительские поля не могут идти в вперемешку. Т.е. после очередного «0» должны идти только его дочерние поля.

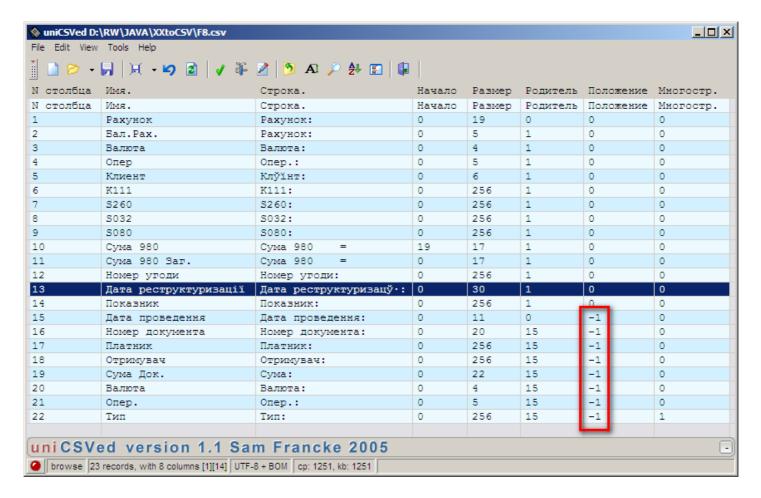
1 или любое число более — в поле родитель значит, что поле дочерние, и указанное число говорит о соответствующем номере родительского.

	同一月・ウ	2 / 3 2	Al 🎉	2 ₩ 🔛 1	u l		
І столбца	MMA.	Строка.	Начало	Размер	Родитель	Положение	Многостр.
V столбца	Имя.	Строка.	Начало	Размер	Родитель	Положение	MHOPOCTP.
	Рахунок	Рахунок:	0	19	0	0	0
100	Вапюта	Валюта:	0	4	1	0	0
3	Опер	Omen.:	0	4	1	0	0
1	Клиент	Клуїнт:	0	6	1	0	0
5	R013	R013:	0	100	1	0	0
5	K030	K030:	0	100	1	0	0
7	S181	S181:	0	100	1	0	0
3	5240	S240:	0	100	1	0	0
)	R012	R012:	0	100	1	0	0
10	Сума 980	Сума 980 =	: 0	37	0	0	0
11	T020	T020:	0	100	10	0	0

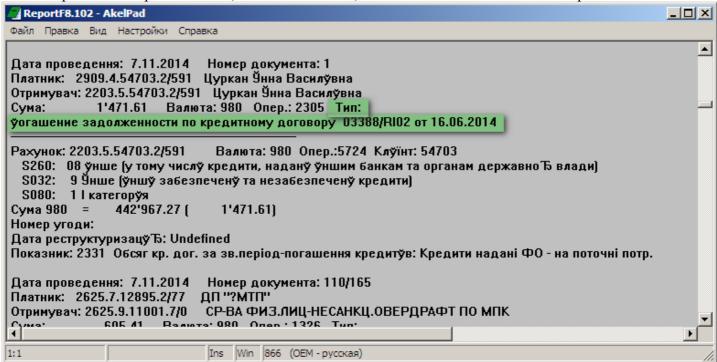
Положение — поле позволяет использовать значение полей которые идут до корневой записи, которую мы выбрали. Значения 0 или -1. Минус говорит что поле идет перед корневой записью, в отчете А7 такого нет, но в F8 встречается.

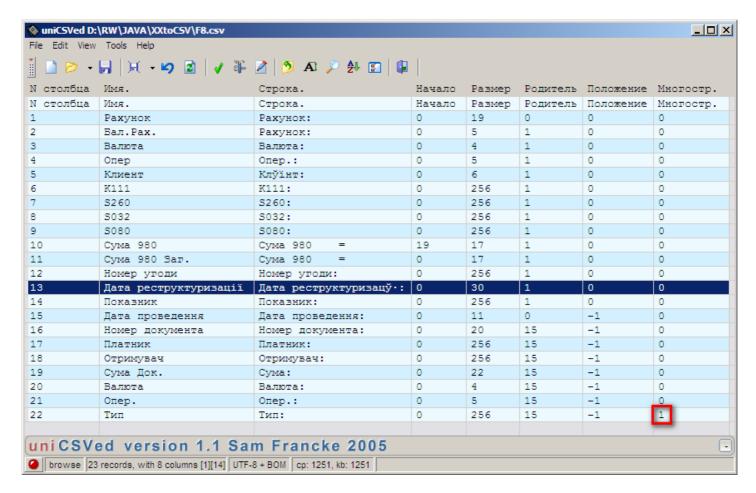
Здесь информация о документе предшествует выбранной корневой записи "**Рахунок:**" Почему мы не выбрали другое корневое поле, просто потому что документа может и не быть, а "**Рахунок:**" есть по определению.



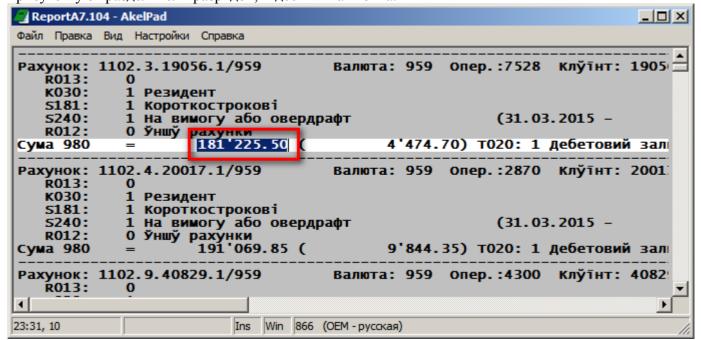


Многостр. - многострочные поля, значения 0 или 1, 1 — означает что поле многострочное.





Формат – указывает формат данных содержащихся в столбце. Принимает значение NUMBER и STRING. При указании значении NUMBER, программа удалит все символы кроме 0-9 и знака минус, при том десятичным разделителем будет запятая. Сделано это потому, что в отчетах системы RS-Bank присутствуют разделители разрядов, и десятичная точка.



В таблицу заносится корректное значение.

	0	Р	
СТ	Сума 980 заг.	Сума 980	T(
	176750,8	176750,8	1
	181225,5	4474,7	1
	101060 95	0011 25	4

При указании значении STRING, программа добавит в конец строки непечатный символ, для предотвращения распознания программой EXCEL строк содержащих только цифры как чисел, и их округления.

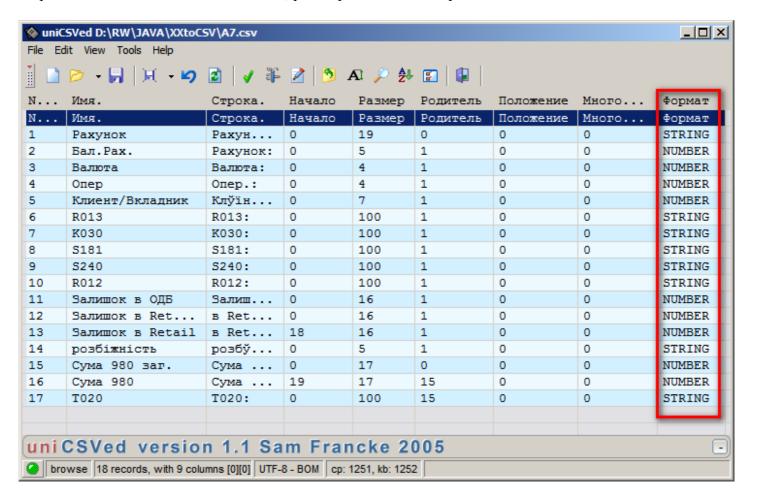
Было так -

A2	8896	•		₱ 26281020252002900								
			Α		В	С						
1	Рахунок				Бал.Рах.	Валюта						
28894				2,6282E+16	-	-						
28895				2,6284E+16	-	-						
28896				2,6281E+16	-	-						
28897				2,6289E+16	-	-						
28898				2,6287E+16	-	-						
22200				2 6288E±16								

После доработки -

A2	8896	~	fx	262	2810202520	002978
		Α			В	С
1	Рахунок				Бал.Рах.	Валюта
28892	2628702	1170002	2978		-	-
28893	2628602	0548001	1978		-	-
28894	2628200	9647002	2978		-	-
28895	2628402	1583001	1978		-	-
28896	2628102	0252002	2978		-	-
28897	2628902	0589001	1978		-	-
20000	2620702	1461004	1070			

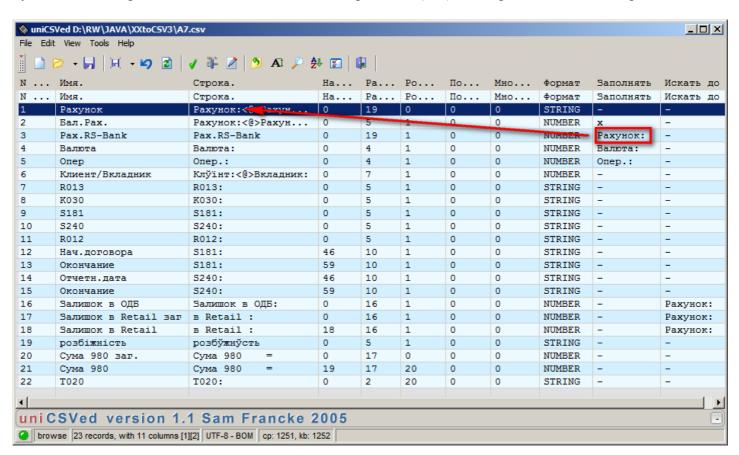
Формат должен иметь каждый столбец, регистр символов не принципиален.



Заполнять – указывает каким историческим значением заполнять пустое поле, хранить ли или запретить хранить исторические записи. При помощи данной функции можно реализовать различные комбинации связи полей.

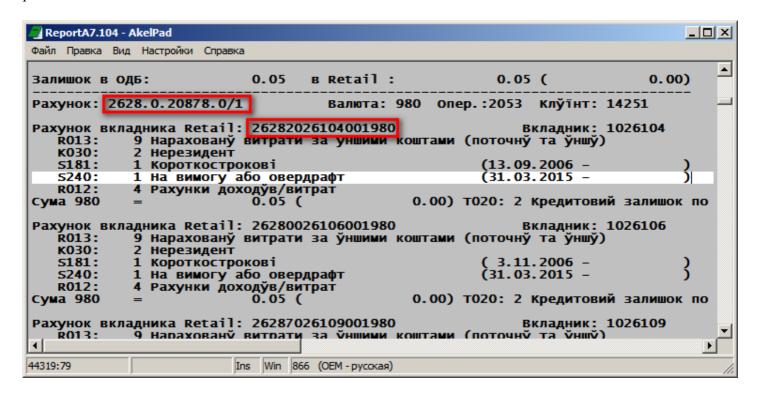
A3	37167	•	fx	2628.6.6	9285.0/978										
		Α		В	С		D	Е	F	G	Н	1	J	K	L
1 /	Рахуно	K		Бал.Рах	Pax.RS-Ban	k	Валюта	Опер	Клиент/Вкл	R013	K030	S181	S240	R012	Нач.договор (
37167	2628.6.	69285.0	/978 -	2028	2628.6.6928	5.0/978	978	6925	14250	-	-	-	-	-	
37168	262890	3224100	3978	2028	2628.6.6928	5.0/978	978	6925	103224	9	1	1	1	4	15.08.2005
37169	262890	3053400	2978	2028	2628.6.6928	5.0/978	978	6925	103053	9	1	1	1	4	2.09.2005
37170	262810	3617200	2978	2028	2628.6.6928	5.0/978	978	6925	103617	9	1	1	1	4	28.09.2005
37171	262870	3089300	2978	2628	2628.6.6928	5.0/978	978	6925	103089	9	1	1	1	4	29.12.2005
37172	262870	3388000	1978	2628	2628.6.6928	5.0/978	978	6925	103388	9	1	1	1	4	8.02.2006
07470	000000			0000	0000 0 0000		070	0005	400004	_					11.00.0000

Принимает значение Строки – поискового параметра, по которому, мы находим значение, в таком случае, поле будет заполнятся его историческим значением. "-" (прочерк) – говорит, что не заполняем пустое поле исторической записью, но ведем историю. "х" (икс) – говорит, что вести историю не надо.



Зачем нужен параметр "x" (икс)? Если имеются поля, которые используют одинаковые поисковые ключи, но вырезают разные значения, во избежание перезаписи истории необходимо запретить запись истории для поля с дублирующим ключом.

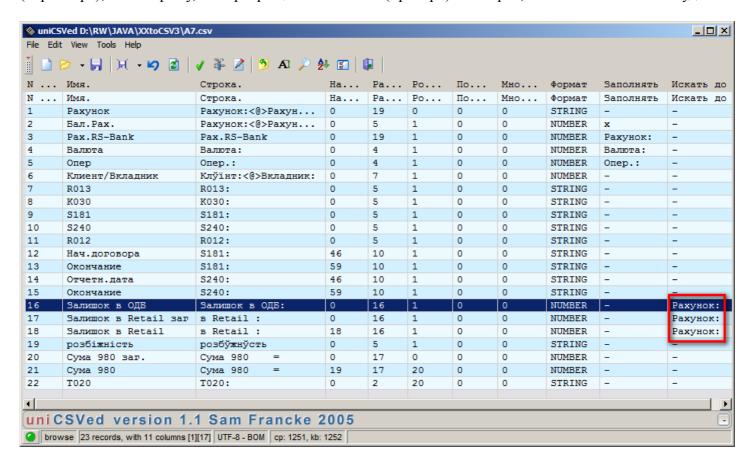
Например при обработке отчета файла A7, при помощи заполнения историческим значением, реализована связь счета RS-BANK и счета Retail.



Т.е. просто введен столбец Pax.RS-Bank который заполняется историческими записями столбца Paxyнок, и в результате можно видеть связь счета.

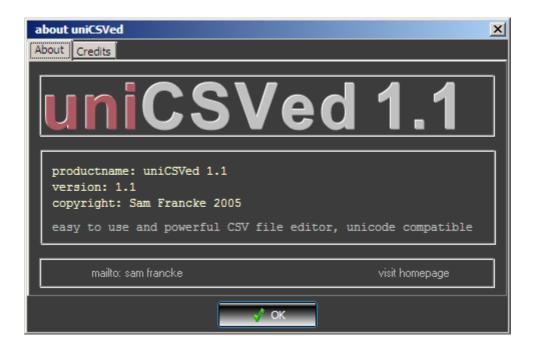
	Α	В	С	D	
1	Рахунок	Бал.Рах.	Pax.RS-Bank	Валют	O
541	2628.0.20878.0/1	2628	2628.0.20878.0/1	980	2
542	26282026104001980	2628	2628.0.20878.0/1	980	2
543	26280026106001980	2628	2628 0 20878 0/1	980	2

Искать до. – указывает, что надо искать значение поля дальше, не взирая ни на что, пока не встретится строка содержащая указанный ключ. Соответственно принимает значение поискового ключа (параметра), по которому, мы прекращаем поиск. "-" (прочерк) – говорит, что искать ничего не будем.



Редактирование файлов конфигурации.

Предлагаю использовать маленькую бесплатную программу uniCSVed, собственно ее интерфейс и представлен выше.



Действия по работе с файлами конфигураций стандартны.

Дополнительно.

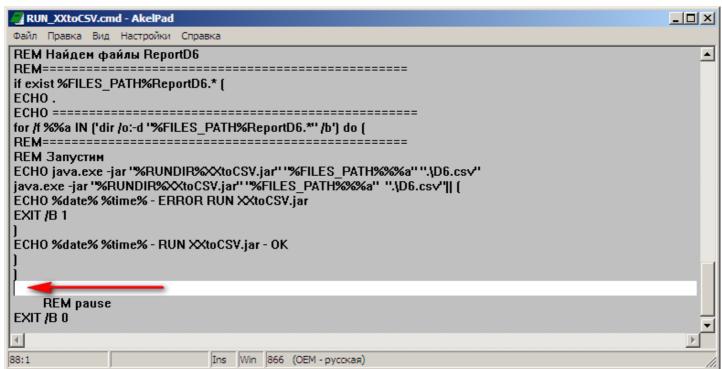
Для удобства работы применяется пакетный командный файл RUN_XXtoCSV.cmd, и если мы хотим конвертировать новые отчеты, необходимо немного изменить и RUN_XXtoCSV.cmd, а именно добавить новый блок с именем нового отчета и его конфигурации.

Где.

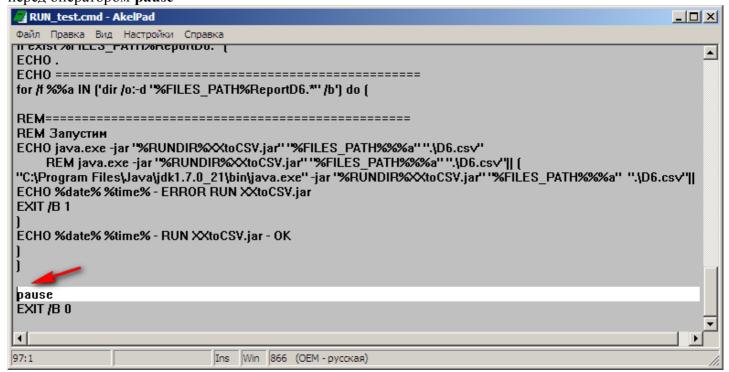
ReportD6.* - поисковый шаблон для файлов отчетов.

D6.csv — конфигурационный файл.

Добавлять блок необходимо в конец пакетного командного файла RUN_XXtoCSV.cmd, перед командой EXIT /B 0



Если нужно контролировать сообщения выдаваемые утилитой при роботе, можно убрать команду **REM** перед оператором **pause**



Что даст консольному окну оставаться открытым.