How to scrape data column which consists of multi columns in the table

Instruction step by step using "Data Scraping" Wizard

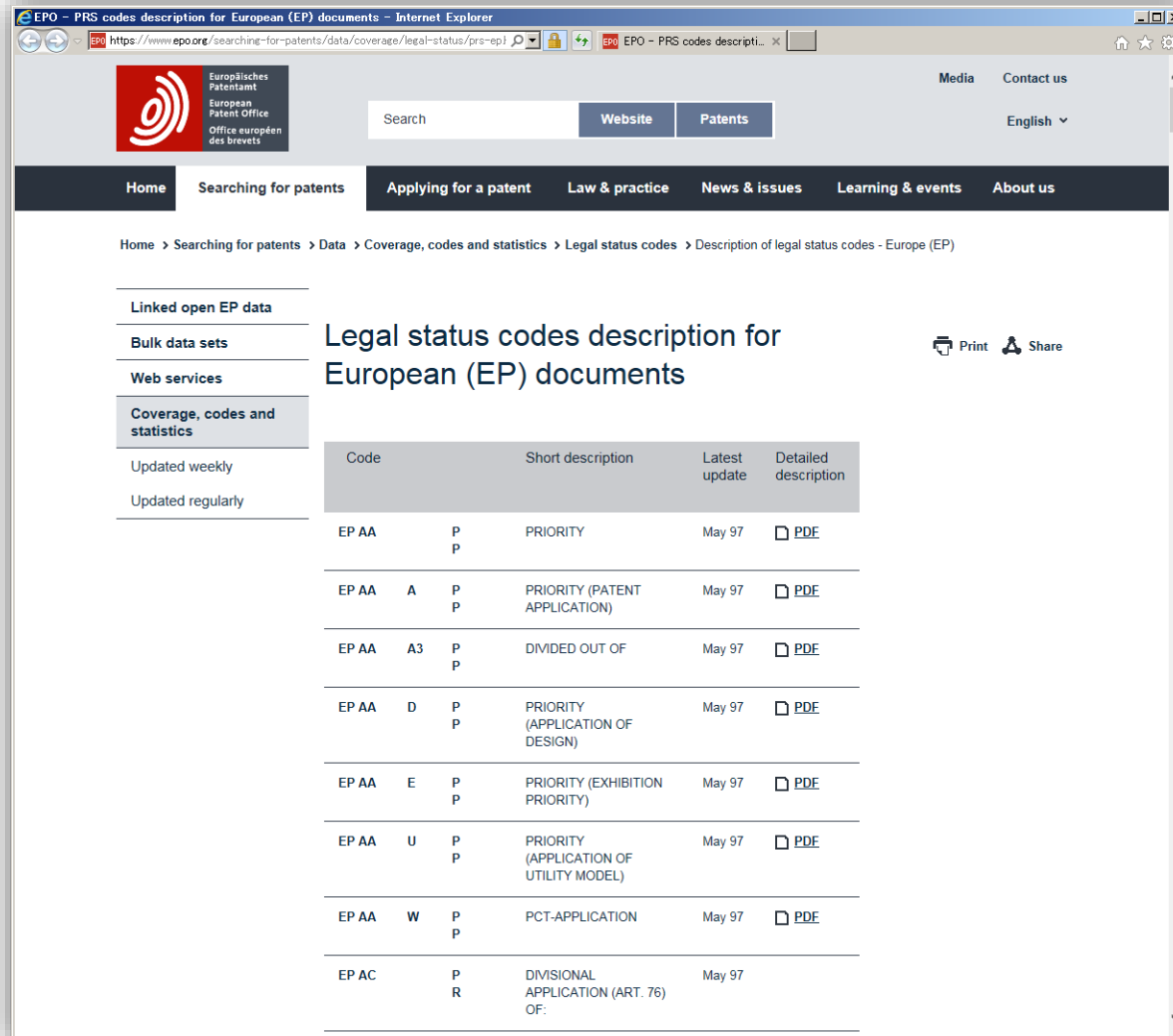# Column which consists of multi columns in the table

- URL : https://www.epo.org/searching-for-patents/data/coverage/legal-status/prs-ep.html

Step 1 : Open Browser at the Page

Step2 : Click "Data Scraping"

Step3 : Click the "Next" on the Wizard

Ui Extract Wizard

Select Element

1. Open your browser, application or document and navigate to where you want to extract data from.

2. Press Next in this dialog and hover the mouse cursor over a data source field.

3. Click that field.

| Help | Cancel | < Back | Next |
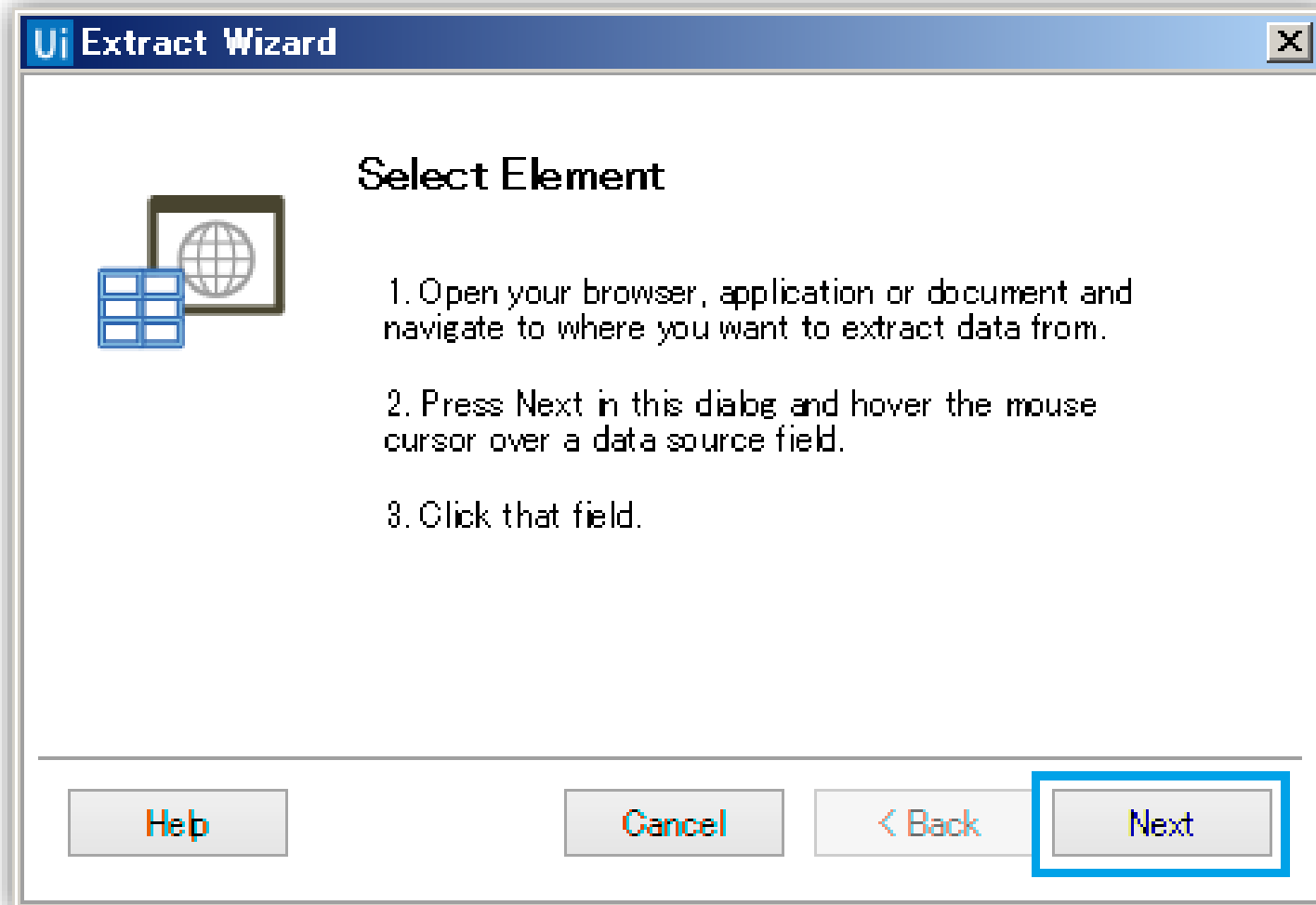
Step4 : Select the cell that colored by Wizard

Step5 : Click the "No" on the dialog
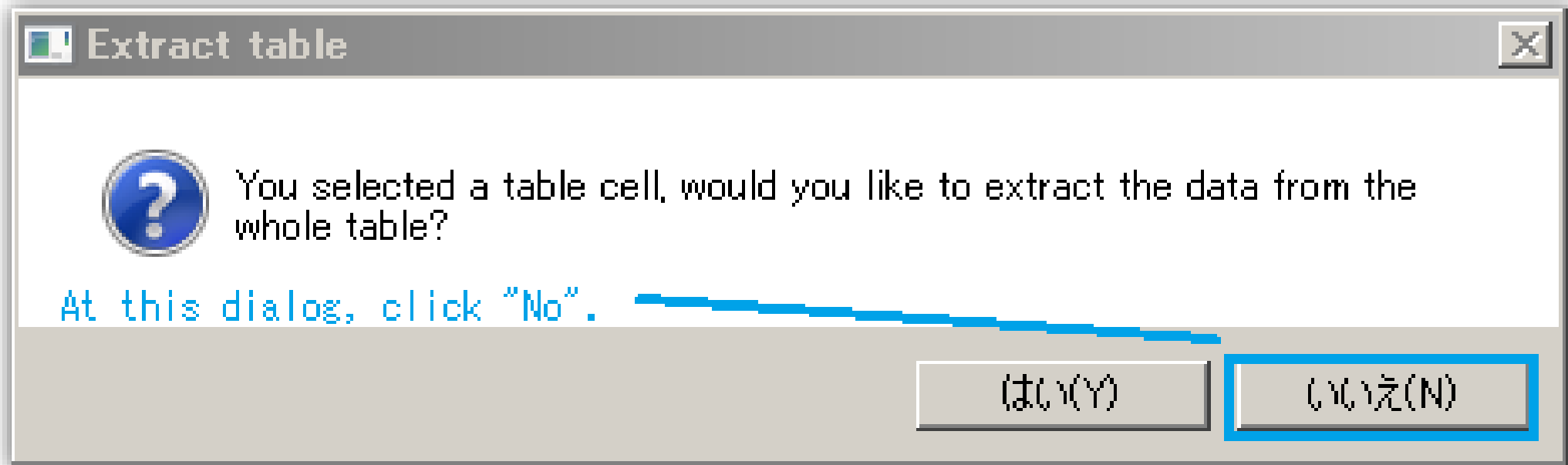
Step6 : Click the "Next" on "Select Second Element"
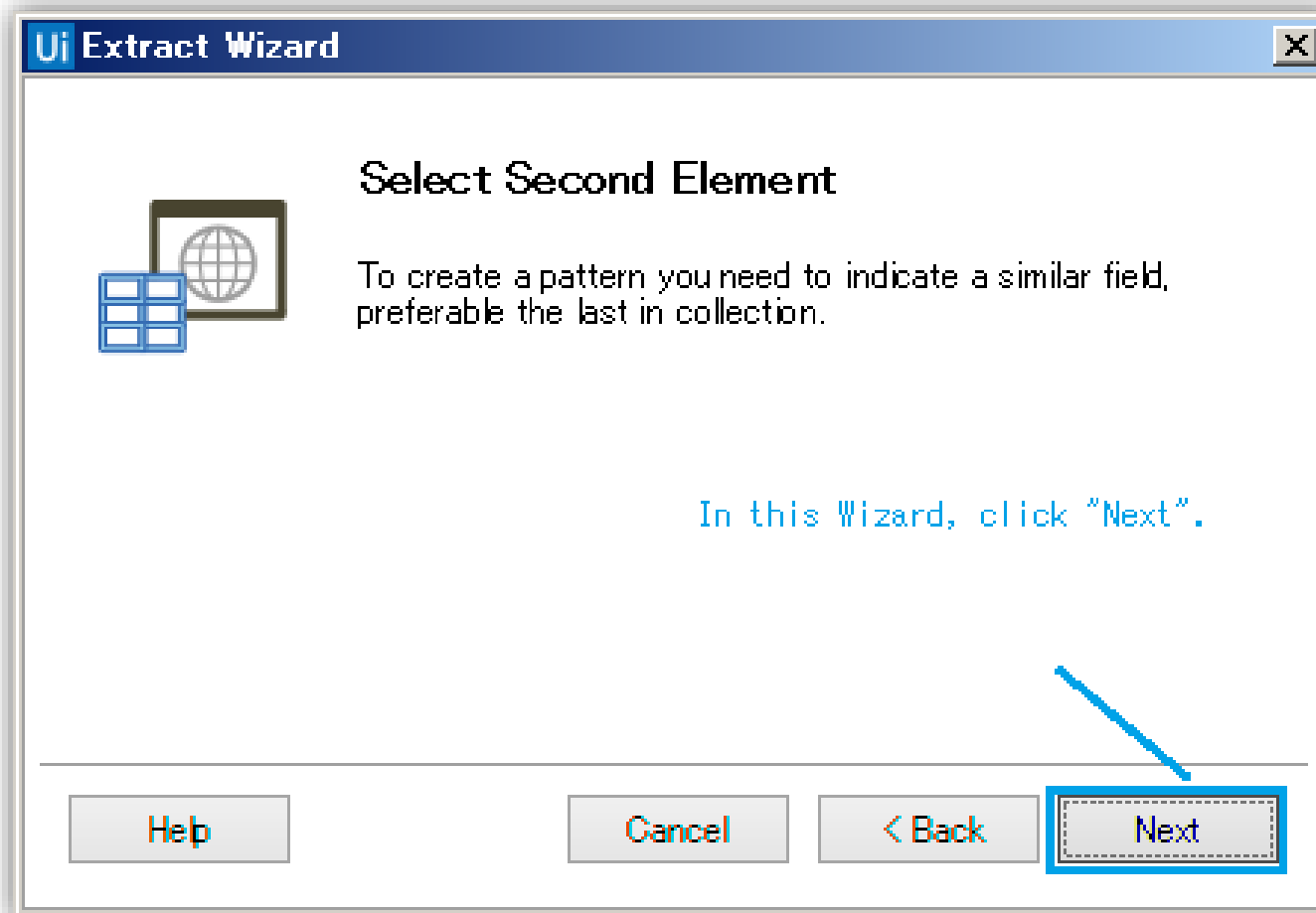
Step7 : Click below cell that colored by Wizard

Legal status codes description for European (EP) documents

Click below cell, and wait for few seconds, and then display a next Wizard.

Code			Short description	Latest update	Detailed description
EP AA		P P	PRIORITY	May 97	📄 PDF
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97	📄 PDF
EP AA	A3	P P	DIVIDED OUT OF	May 97	📄 PDF
EP AA	D	P P	PRIORITY (APPLICATION OF DESIGN)	May 97	📄 PDF
EP AA	E	P P	PRIORITY (EXHIBITION PRIORITY)	May 97	📄 PDF
EP AA	U	P P	PRIORITY (APPLICATION OF UTILITY MODEL)	May 97	📄 PDF
EP AA	W	P P	PCT-APPLICATION	May 97	📄 PDF
EP AC		P R	DIVISIONAL APPLICATION (ART. 76) OF:	May 97	

Step8 : Change the Column Name."Column1" to "Code1"
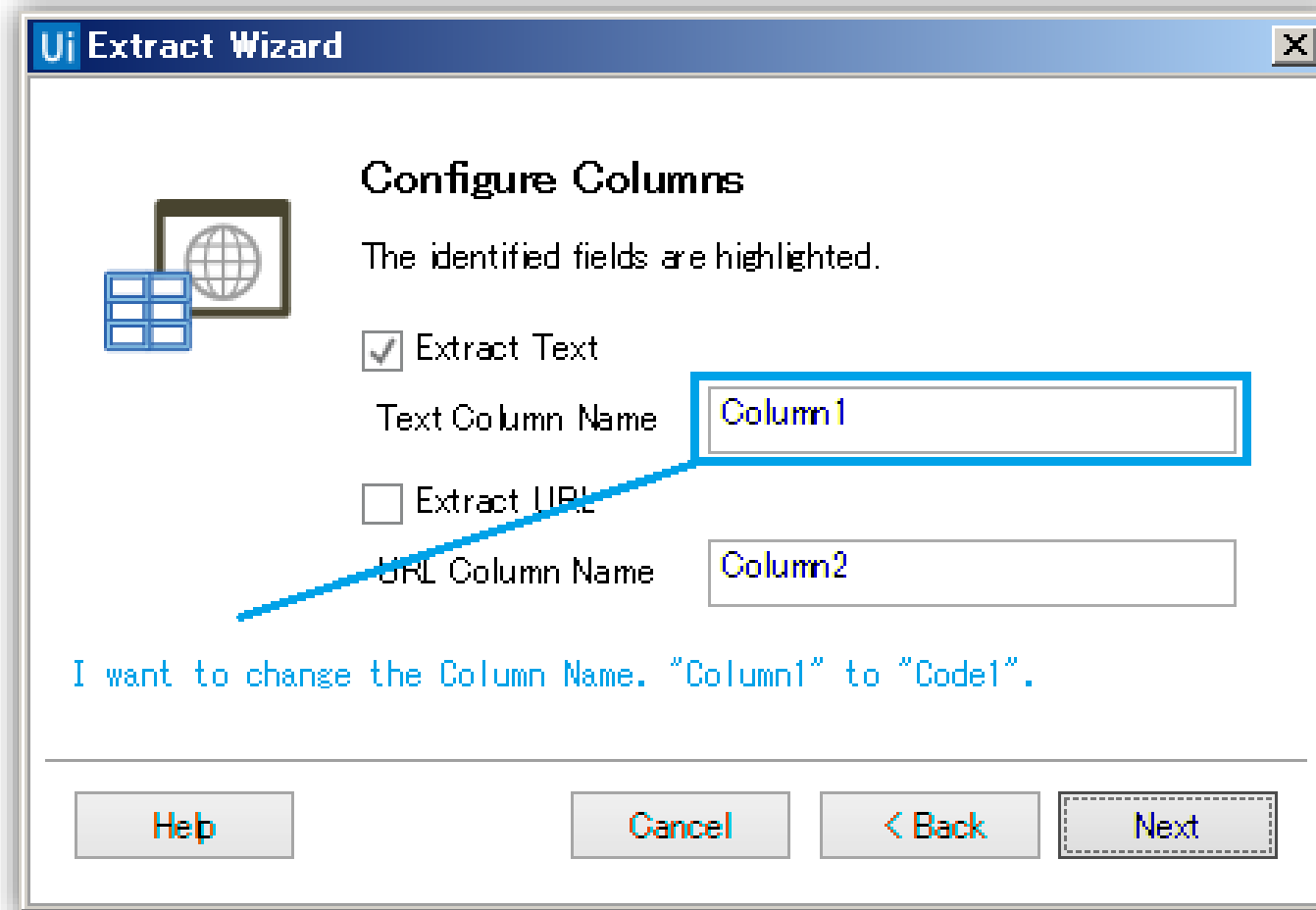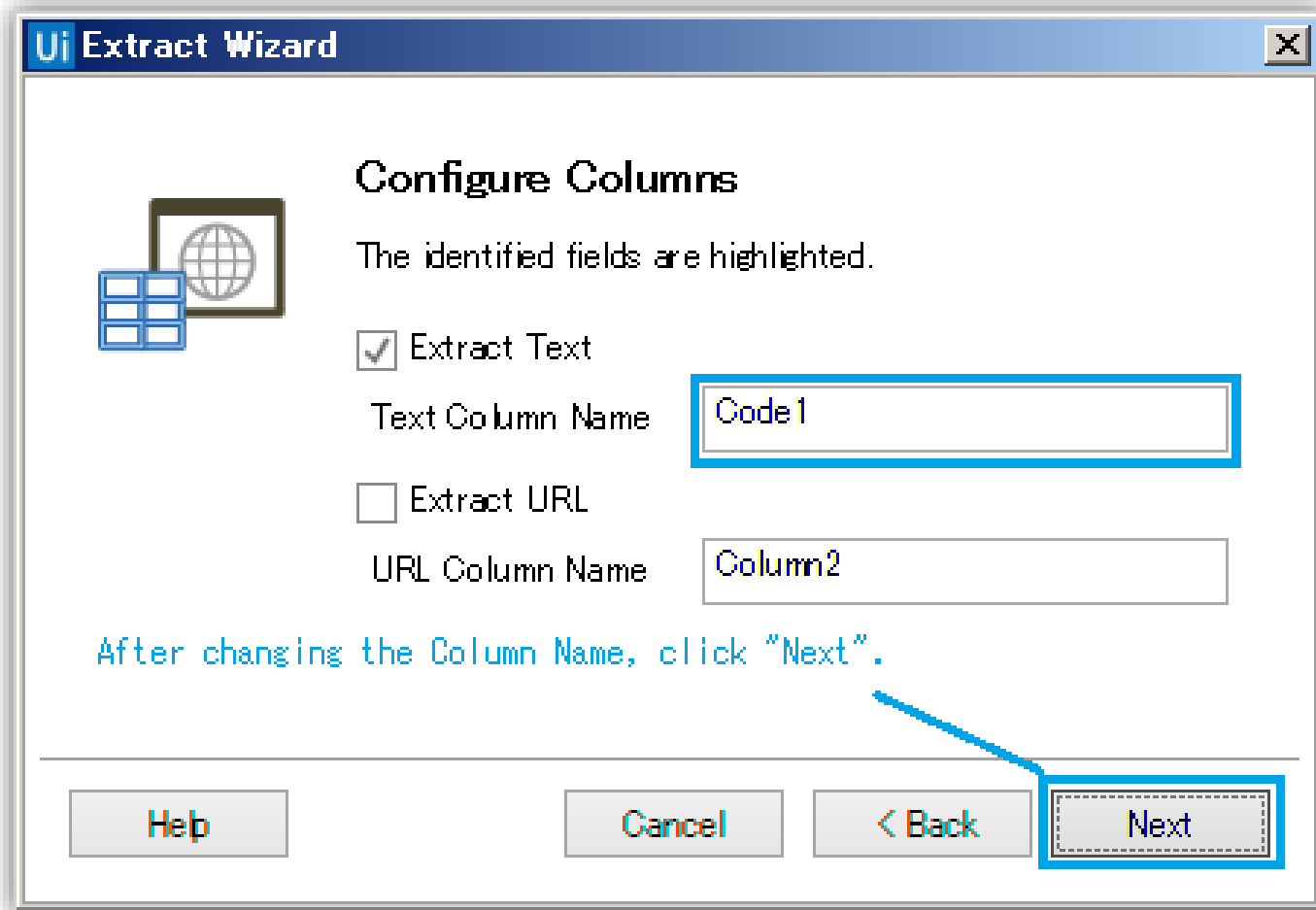
Step9 : Click the "Next" on the Wizard

Step10 : Check the "Preview Data" and then click "Extract Correlated Data" for the next data column

Step11 : Click the cell indicated the next data column

Step12 : Click the "Next" on "Select Second Element"

Extract Wizard

Select Second Element

To create a pattern you need to indicate a similar field, preferable the last in collection.

Click "Next".
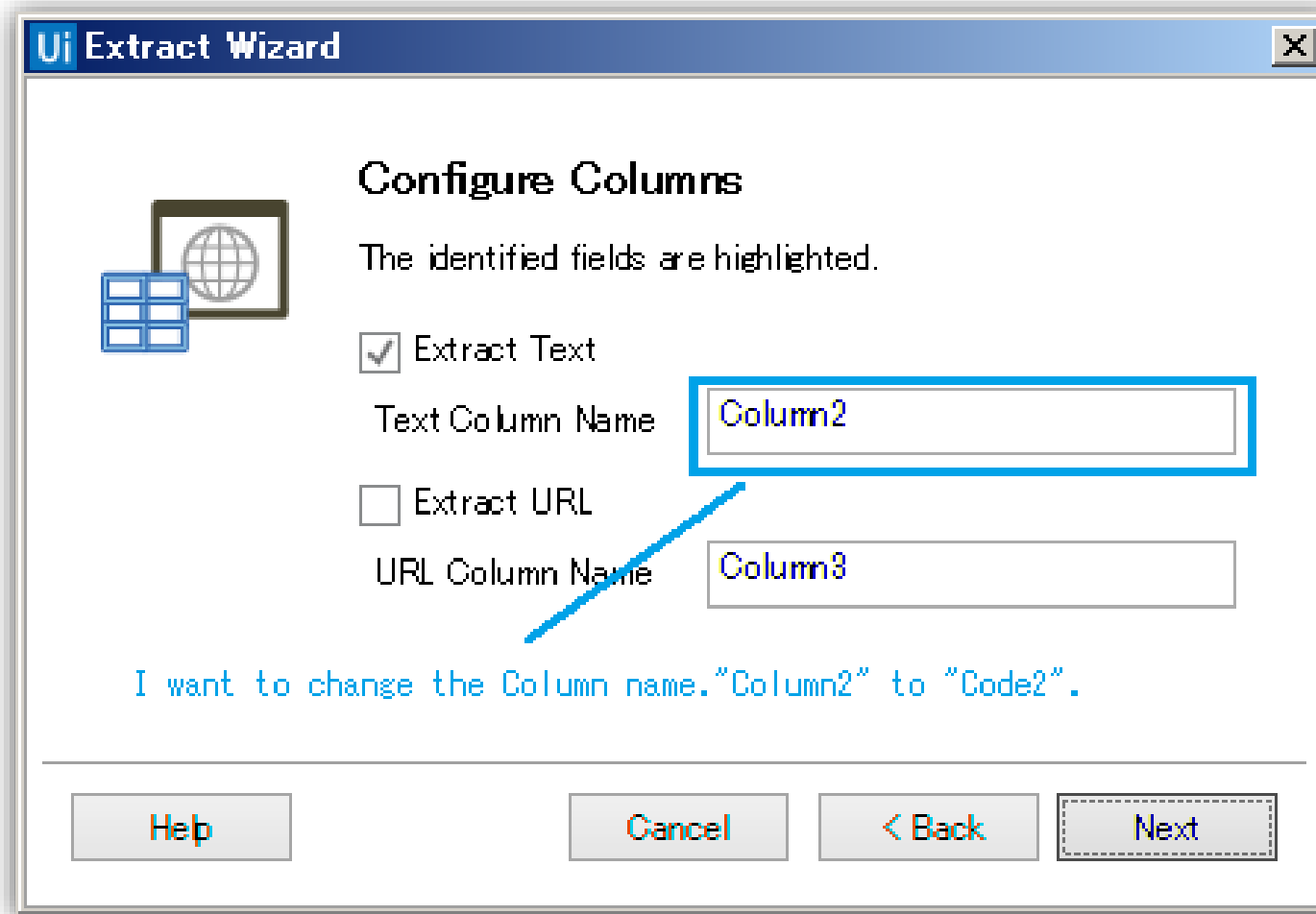
| Help | | Cancel | < Back | Next |
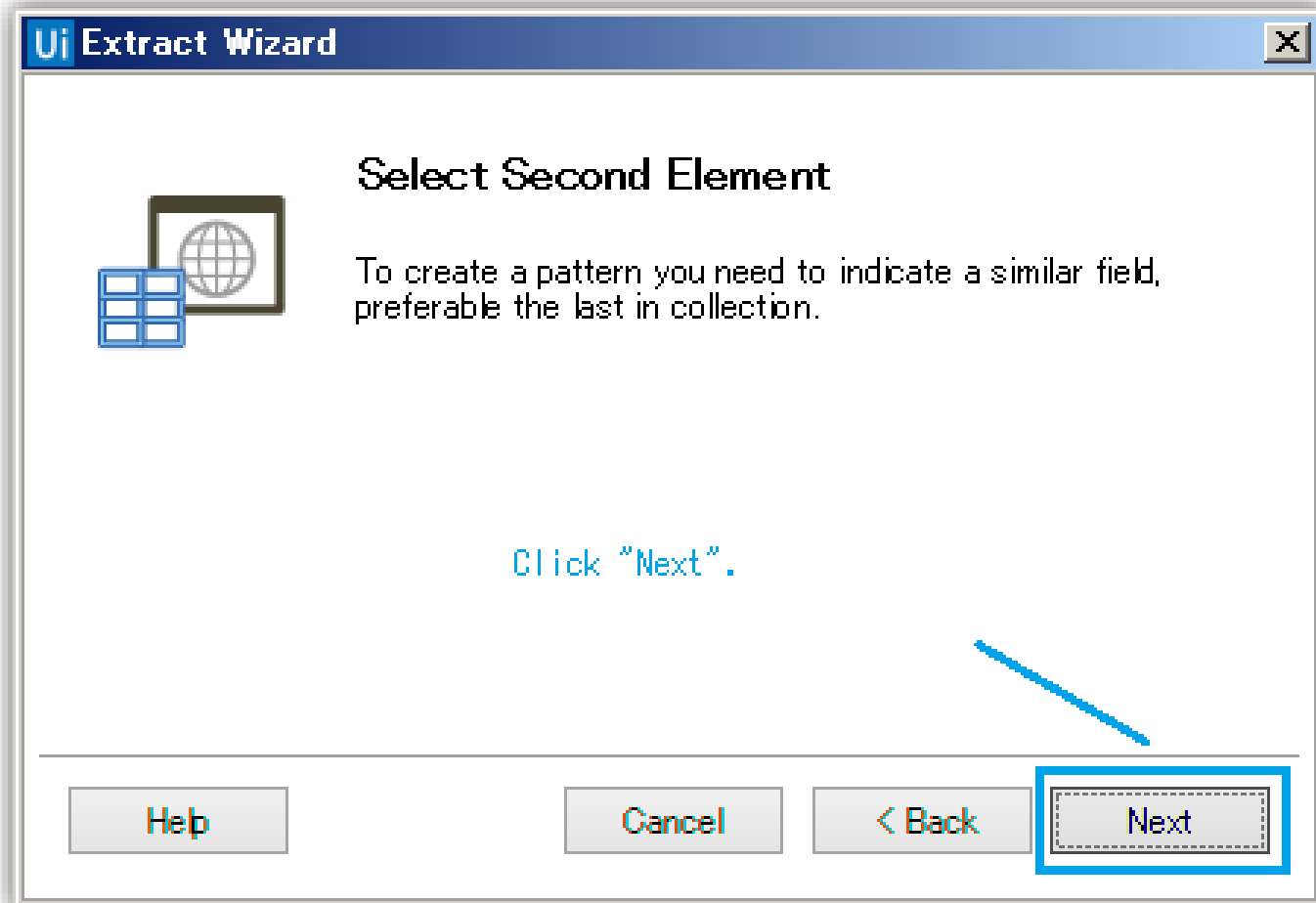
Step13 : Click below cell that colored by Wizard

Legal status codes description for European (EP) documents

Click a below cell, and wait for few seconds, and then display "Wizard".

Code			Short description	Latest update	Detailed descriptio
EP AA		P P	PRIORITY	May 97	📄 PDF
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97	📄 PDF
EP AA	A3	P P	DIVIDED OUT OF	May 97	📄 PDF
EP AA	D	P P	PRIORITY (APPLICATION OF DESIGN)	May 97	📄 PDF
EP AA	E	P P	PRIORITY (EXHIBITION PRIORITY)	May 97	📄 PDF
EP AA	U	P P	PRIORITY (APPLICATION OF UTILITY MODEL)	May 97	📄 PDF
EP AA	W	P P	PCT-APPLICATION	May 97	📄 PDF
EP AC		P R	DIVISIONAL APPLICATION (ART. 76) OF:	May 97	

Step14 : Change the Column Name."Column2" to "Code2"

Step15 : Click the "Next" on the Wizard
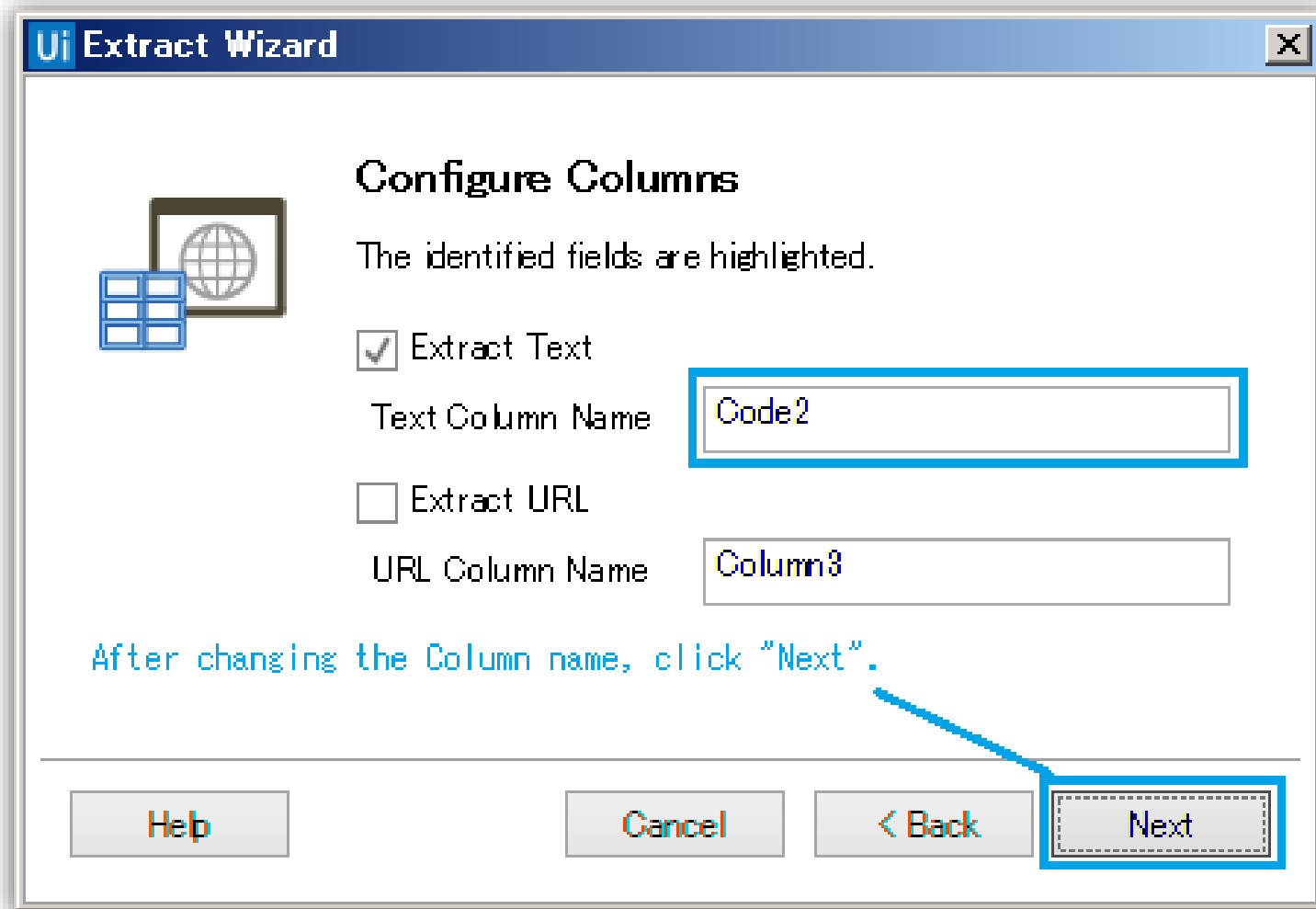
Step16 : Check the "Preview Data" and then click "Extract Correlated Data" for the next data column

Step17 : Click the cell indicated the next data column

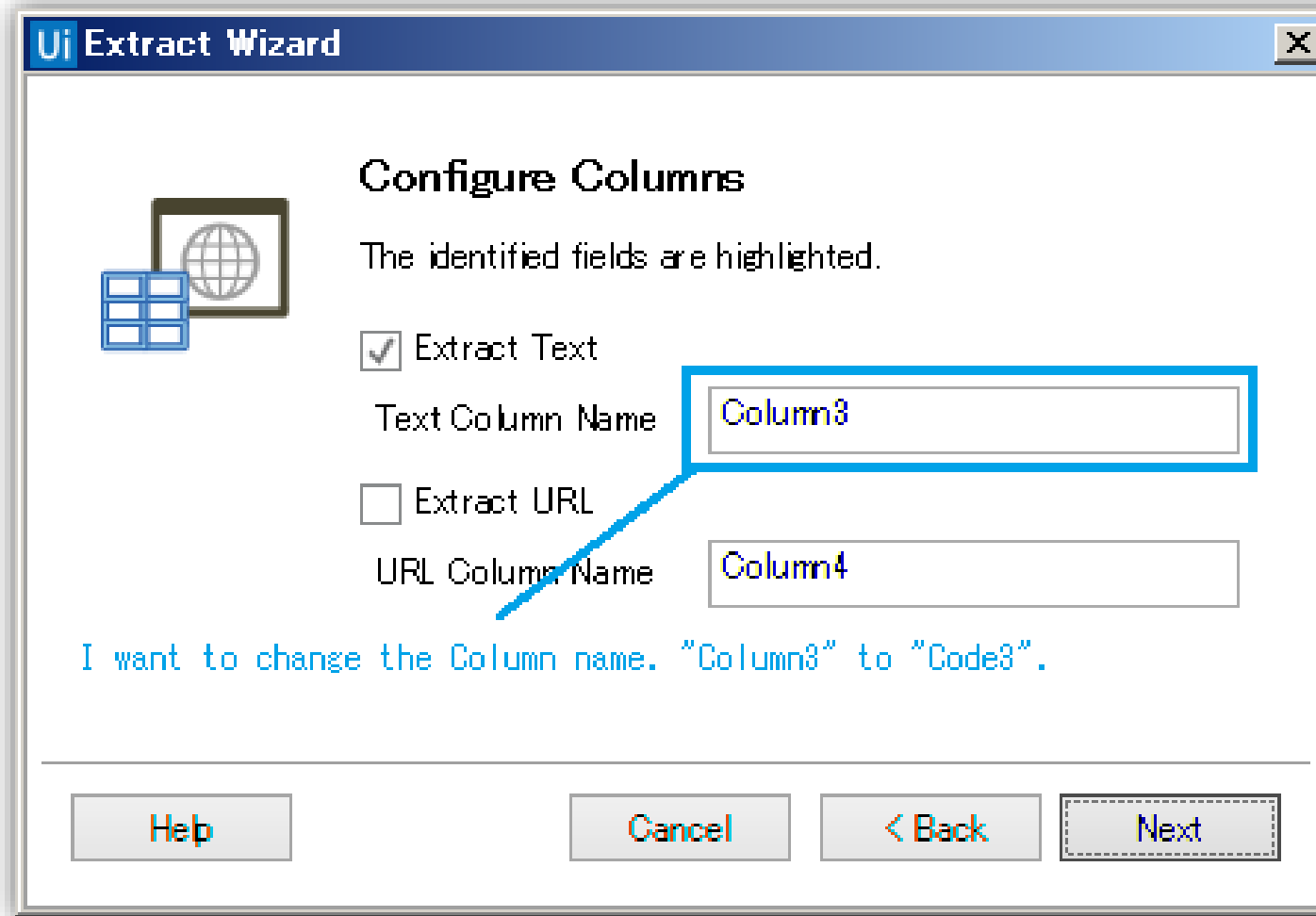| Code			Short description	Latest update
EP AA		P P	PRIORITY	May 97
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97
EP AA	A3	P P	DIVIDED OUT OF	May 97
EP AA	D	P	PRIORITY	May 97

Click the cell.

Step18 : Click the "Next" on "Select Second Element"
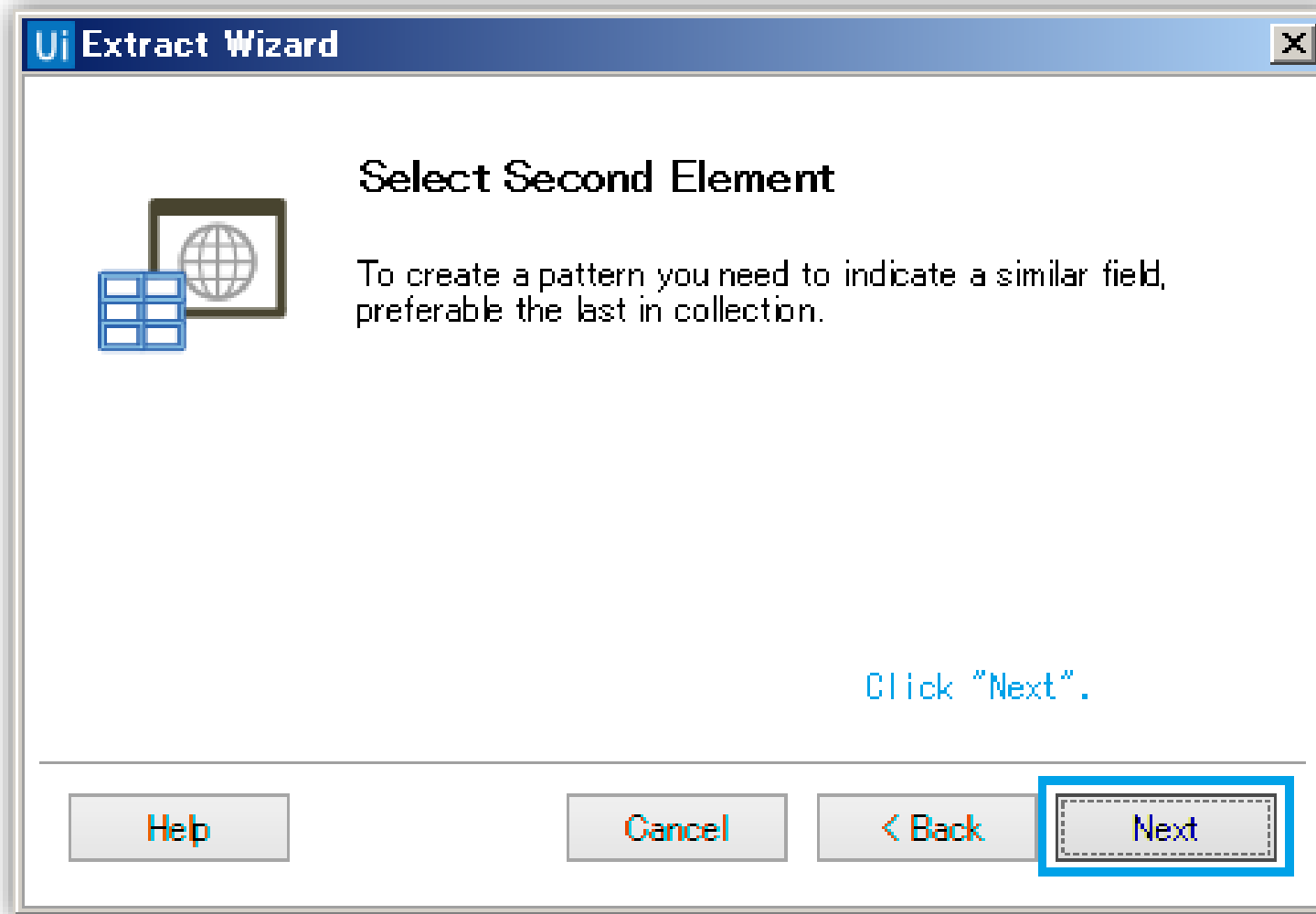
Step19 : Click below cell that colored by Wizard

Code			Short description	Latest update	Deta descr
			Click a below cell, and wait for few seconds, and then display "Wizard".		
EP AA		P P	PRIORITY	May 97	☐ PI
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97	☐ PI
EP AA	A3	P P	DIVIDED OUT OF	May 97	☐ PI
EP AA	D	P P	PRIORITY (APPLICATION OF DESIGN)	May 97	☐ PI
EP AA	E	P P	PRIORITY (EXHIBITION PRIORITY)	May 97	☐ PI
EP AA	U	P P	PRIORITY (APPLICATION OF UTILITY MODEL)	May 97	☐ PI
EP AA	W	P P	PCT-APPLICATION	May 97	☐ PI
EP AC		P R	DIVISIONAL APPLICATION (ART. 76) OF:	May 97	

Step20 : Change the Column Name."Column3" to "Code3"

Step21 : Click the "Next" on the Wizard

Step22 : Check the "Preview Data" and then click "Extract Correlated Data" for the next data column

Step23 : Click the cell indicated the next data column

Code Click the cell.				Short description	Latest update	Detailed description
EP AA		P P		PRIORITY	May 97	🗋 PDF
EP AA	A	P P		PRIORITY (PATENT APPLICATION)	May 97	🗋 PDF
EP AA	A3	P P		DIVIDED OUT OF	May 97	🗋 PDF
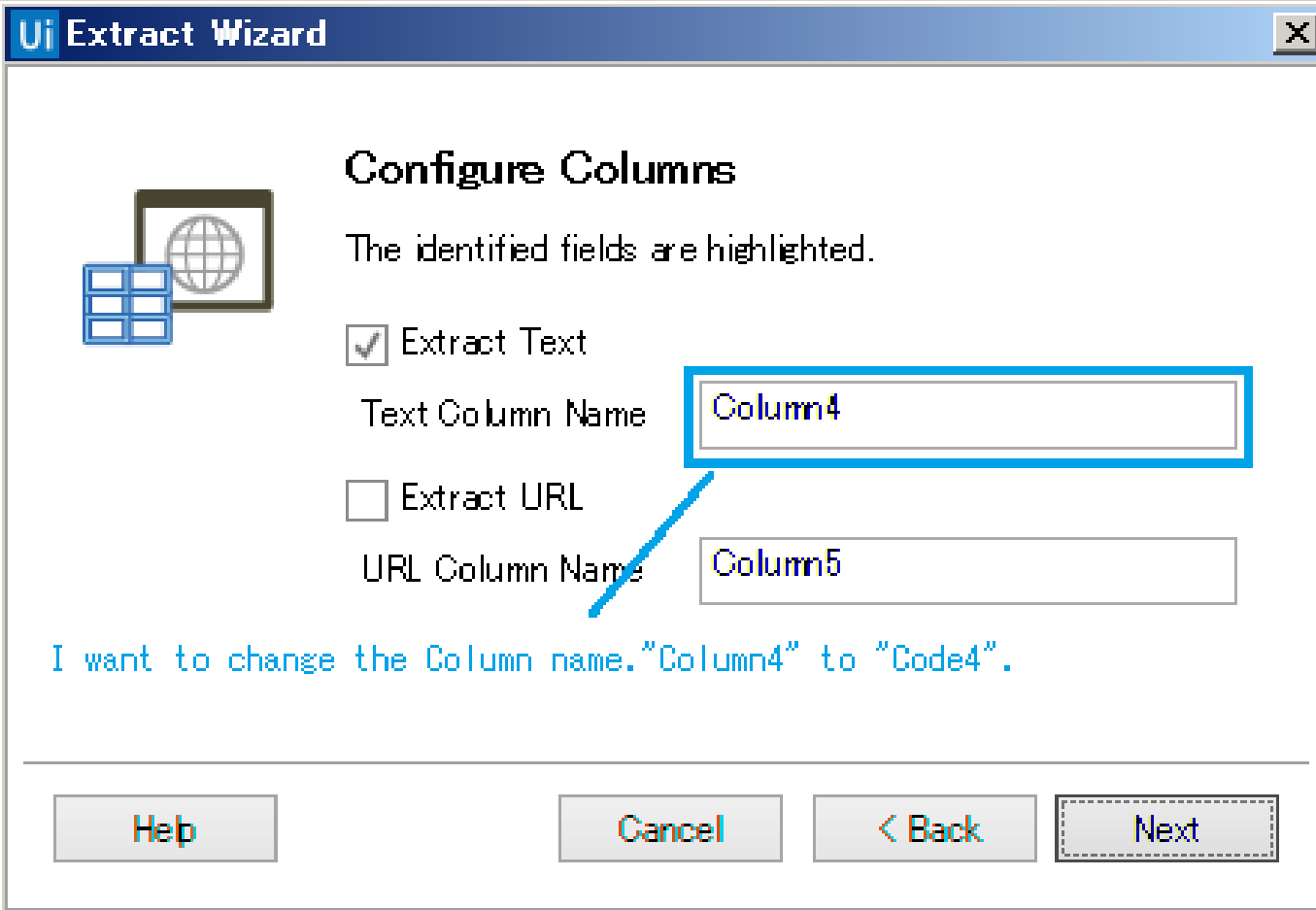
Step24 : Click the "Next" on "Select Second Element"

Extract Wizard

Select Second Element

To create a pattern you need to indicate a similar field, preferable the last in collection.

Click "Next"

| Help | | Cancel | < Back | Next |
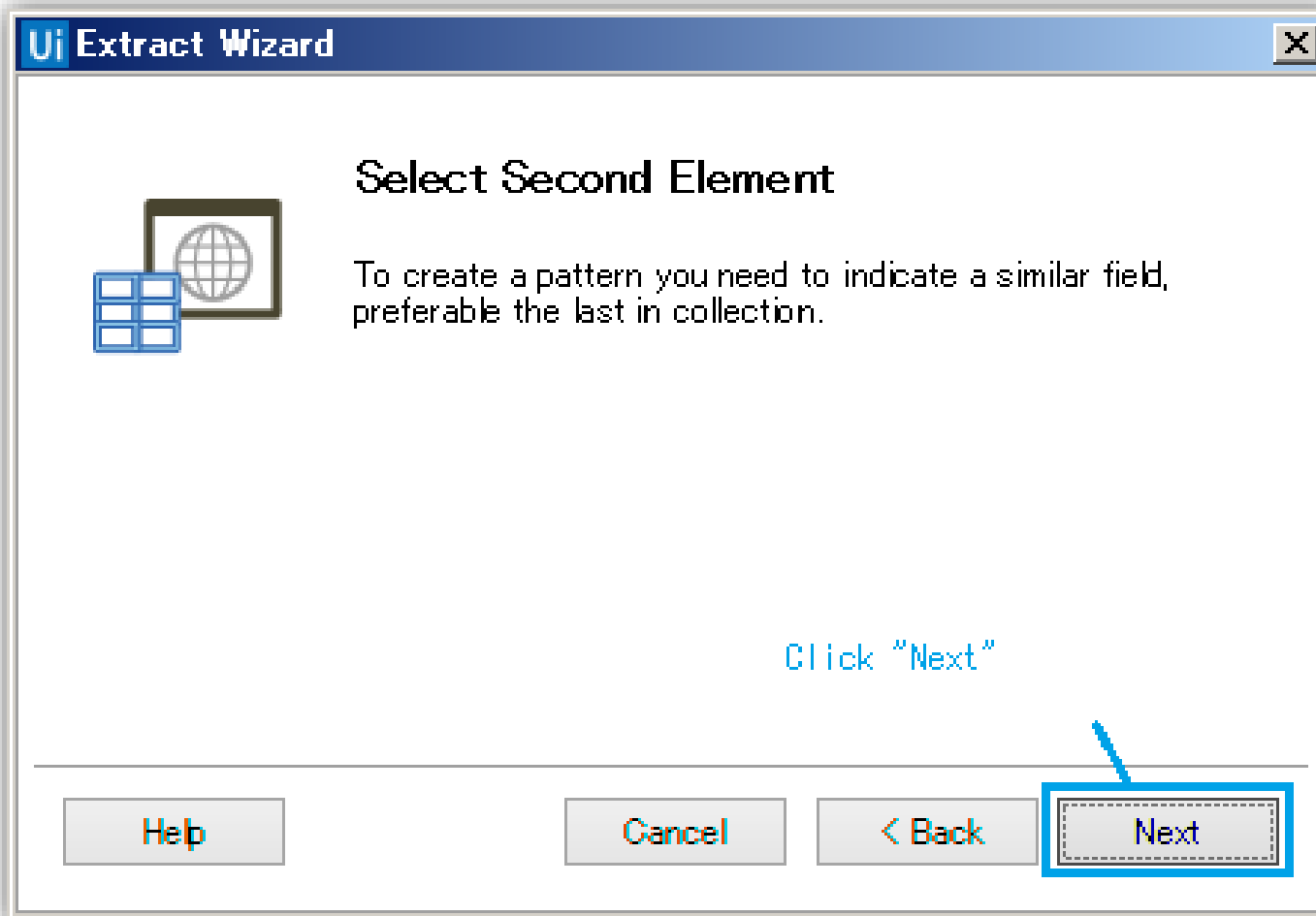
Step25 : Click below cell that colored by Wizard

Code			Short description	Latest update	Detailed description
Click a below cell, and wait few seconds, and then display "wizard".					
EP AA		P P	PRIORITY	May 97	📄 PDF
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97	📄 PDF
EP AA	A3	P P	DIVIDED OUT OF	May 97	📄 PDF
EP AA	D	P P	PRIORITY (APPLICATION OF DESIGN)	May 97	📄 PDF
EP AA	E	P P	PRIORITY (EXHIBITION PRIORITY)	May 97	📄 PDF
EP AA	U	P P	PRIORITY (APPLICATION OF UTILITY MODEL)	May 97	📄 PDF
EP AA	W	P P	PCT-APPLICATION	May 97	📄 PDF
EP AC		P R	DIVISIONAL APPLICATION (ART. 76) OF:	May 97	

Step26 : Change the Column Name."Column4" to "Code4"

Step27 : Click the "Next" on the Wizard

Step28 : Check the "Preview Data" and then click "Extract Correlated Data" for the next data column
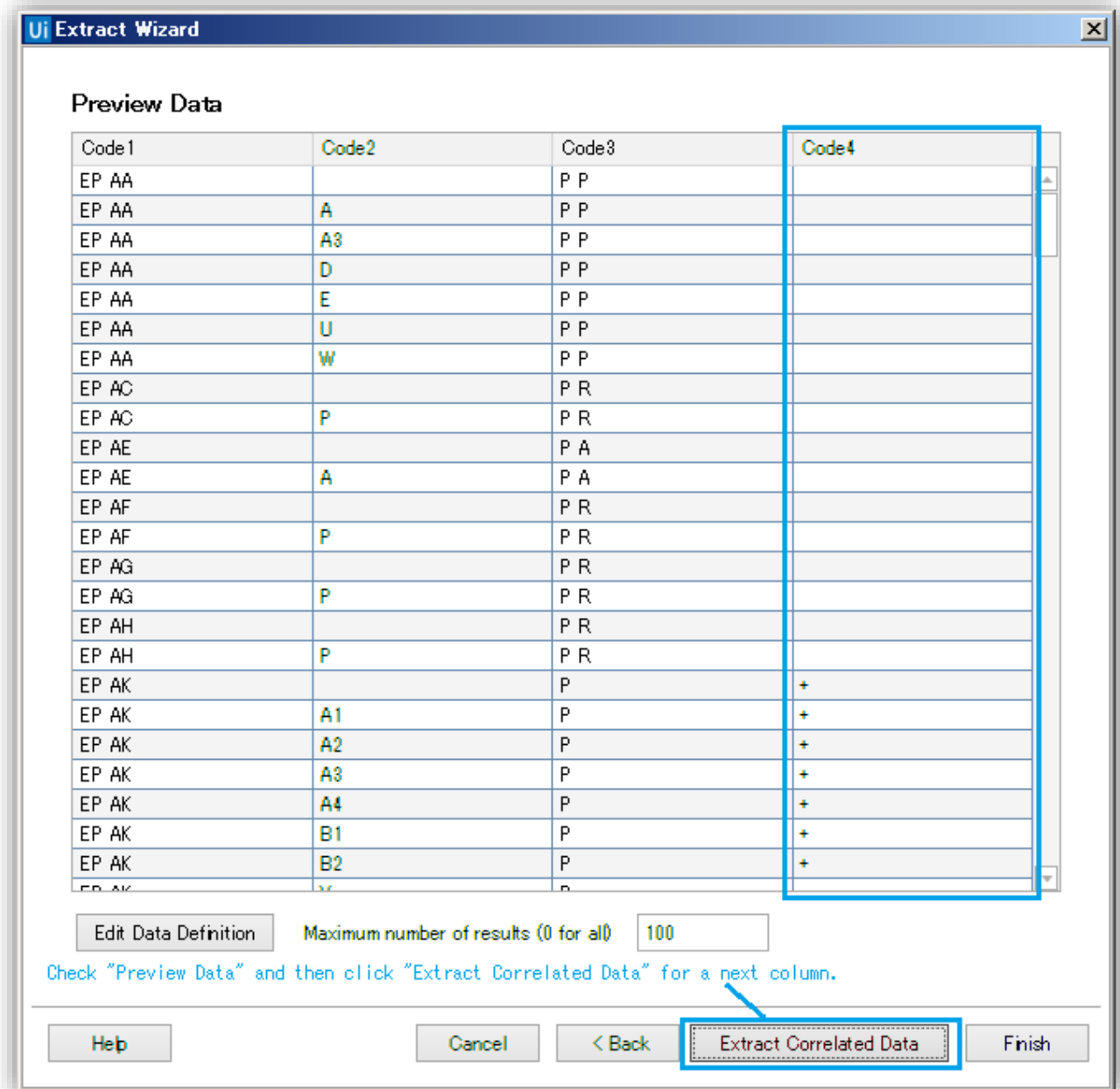
Step29 : Click the cell indicated the "Short description" data column

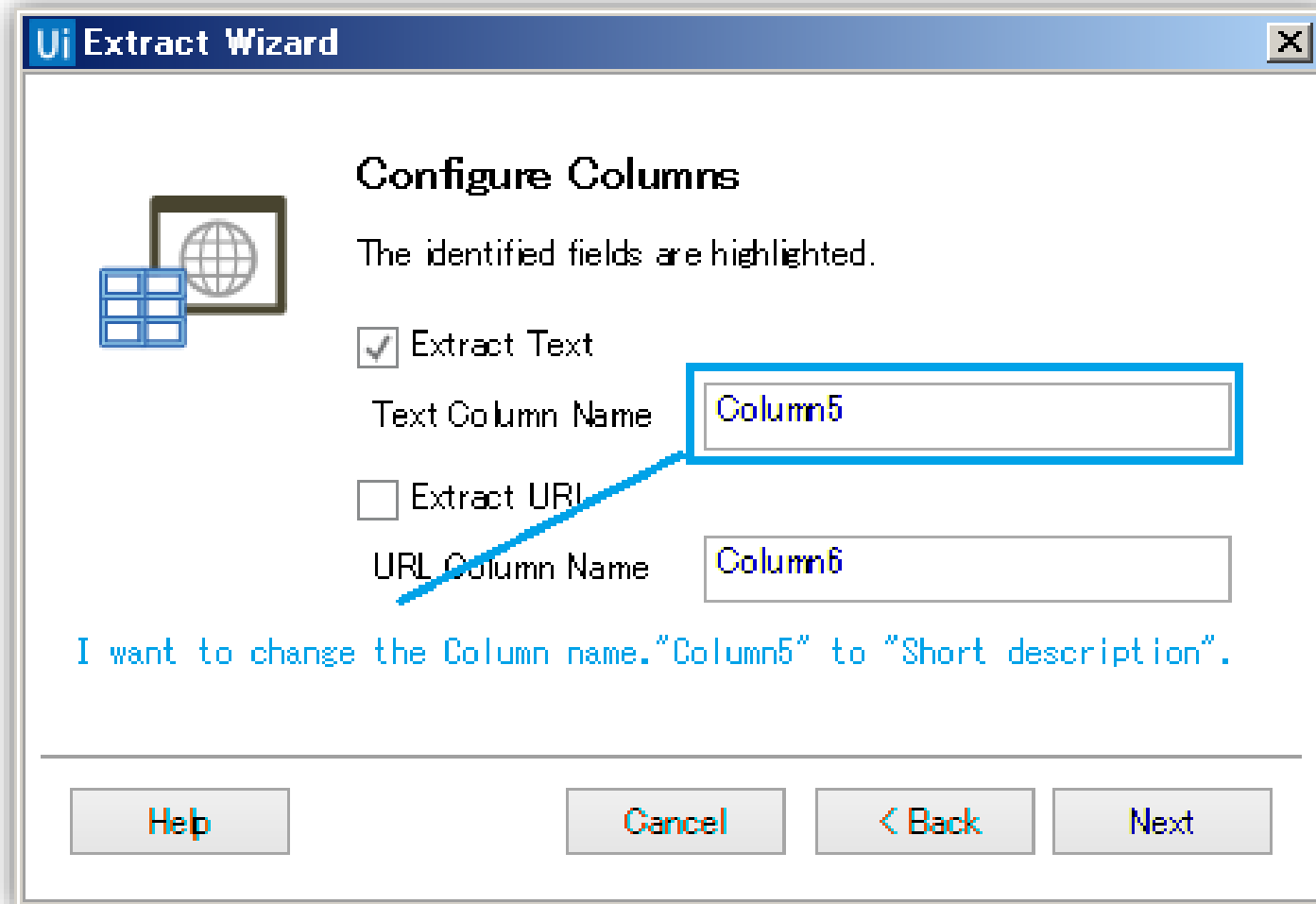| Code			Short description	Latest update	Detailed description
			Click the cell.		
EP AA		P P	PRIORITY	May 97	📄 PDF
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97	📄 PDF
EP AA	A3	P P	DIVIDED OUT OF	May 97	📄 PDF

Step30 : Click the "Next" on the Wizard
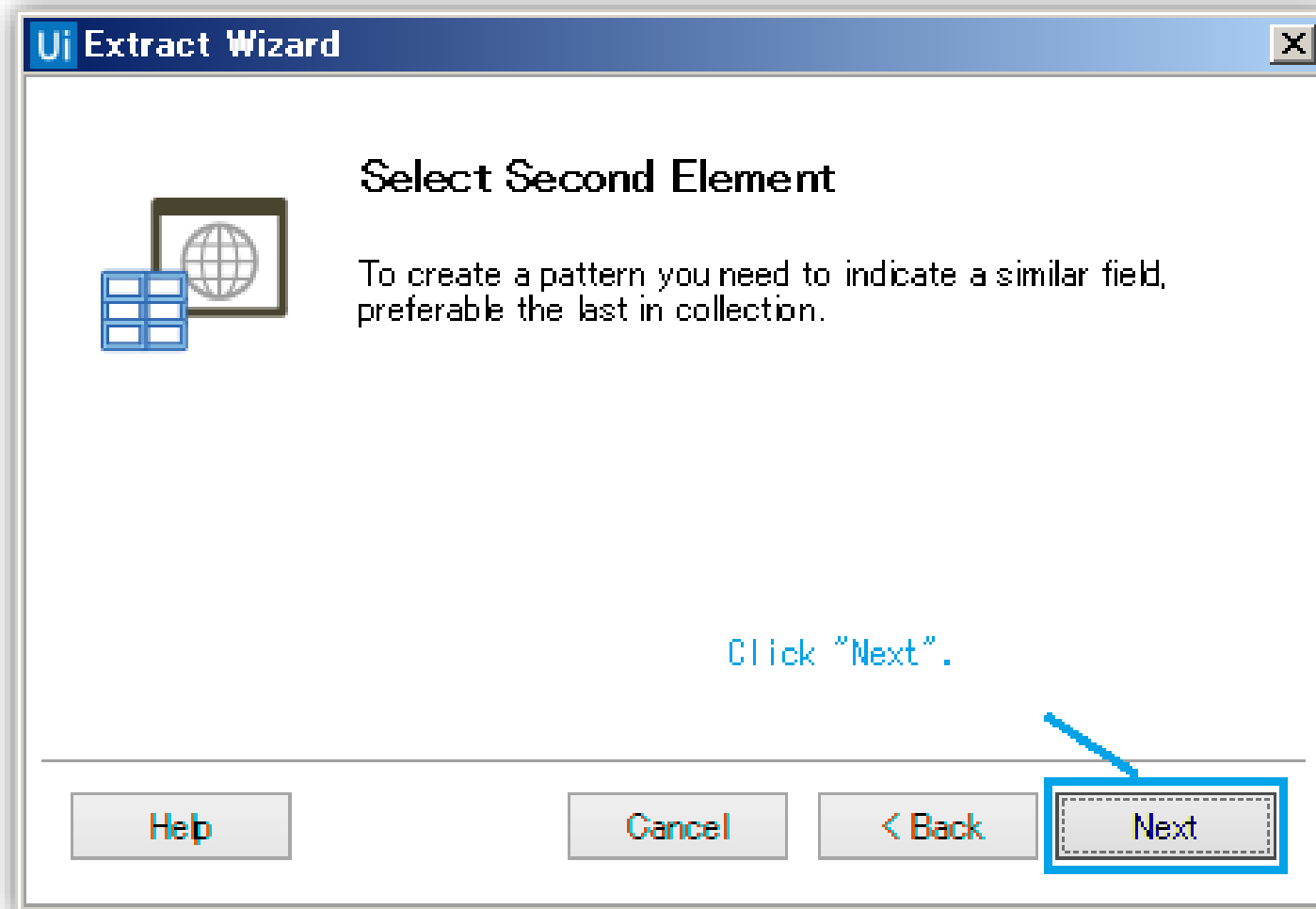
Step31 : Click below cell that colored by Wizard

Code			Short description	Latest update	Detailed description
Click a below cell and wait for few seconds and then dispaly "Wizard".					
EP AA		P P	PRIORITY	May 97	📄 PDF
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97	📄 PDF
EP AA	A3	P P	DIVIDED OUT OF	May 97	📄 PDF
EP AA	D	P P	PRIORITY (APPLICATION OF DESIGN)	May 97	📄 PDF
EP AA	E	P P	PRIORITY (EXHIBITION PRIORITY)	May 97	📄 PDF
EP AA	U	P P	PRIORITY (APPLICATION OF UTILITY MODEL)	May 97	📄 PDF
EP AA	W	P P	PCT-APPLICATION	May 97	📄 PDF
EP AC		P R	DIVISIONAL APPLICATION (ART. 76) OF:	May 97	

Step32 : Change the Column Name."Column5" to "Short description"

Step33 : Click the "Next" on the Wizard

Step34 : Check the "Preview Data" and then click "Extract Correlated Data" for the next data column

Step35 : Click the cell indicated the "Latest update" data column

Step36 : Click the "Next" on the Wizard
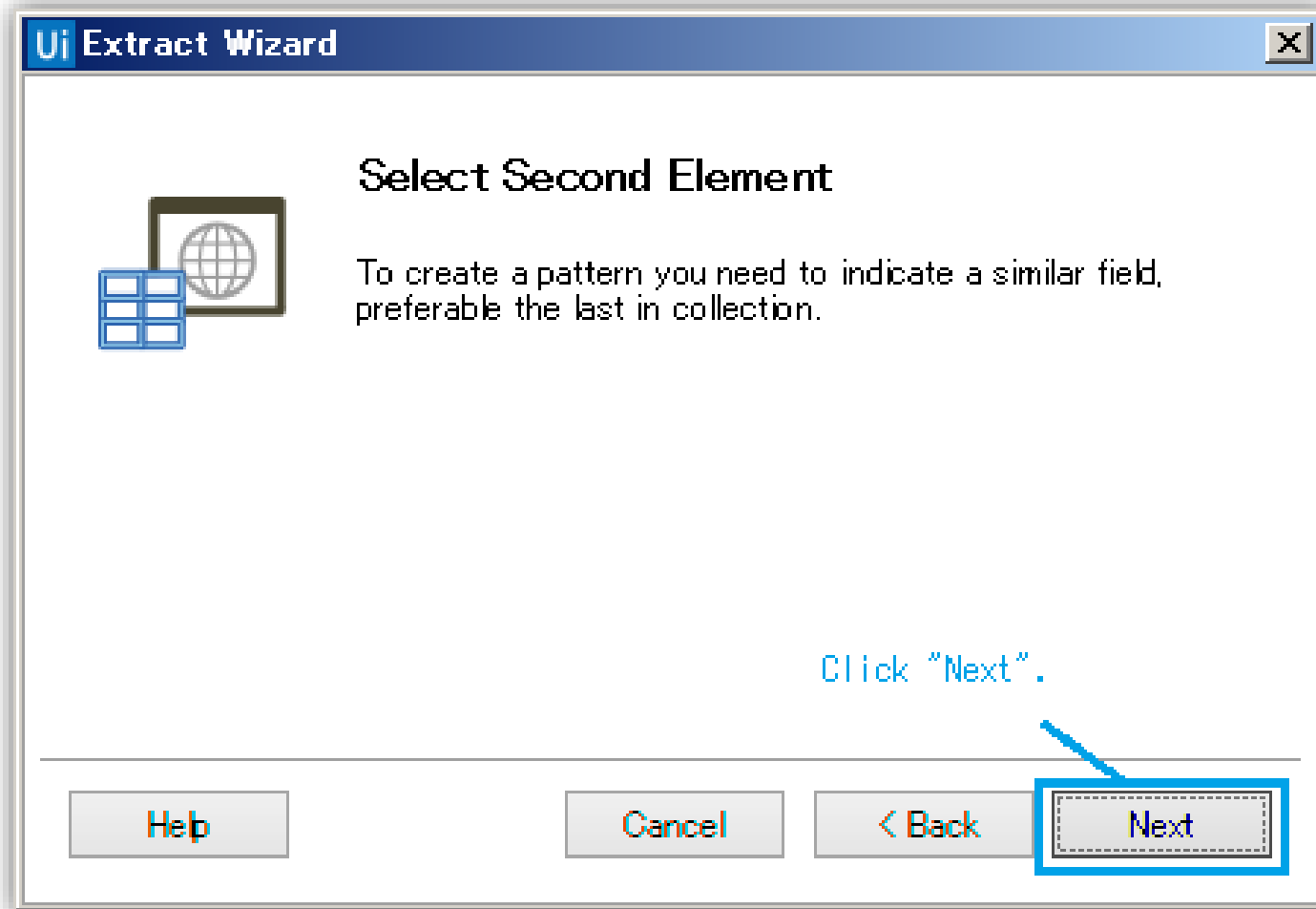
Step37 : Click below cell that colored by Wizard

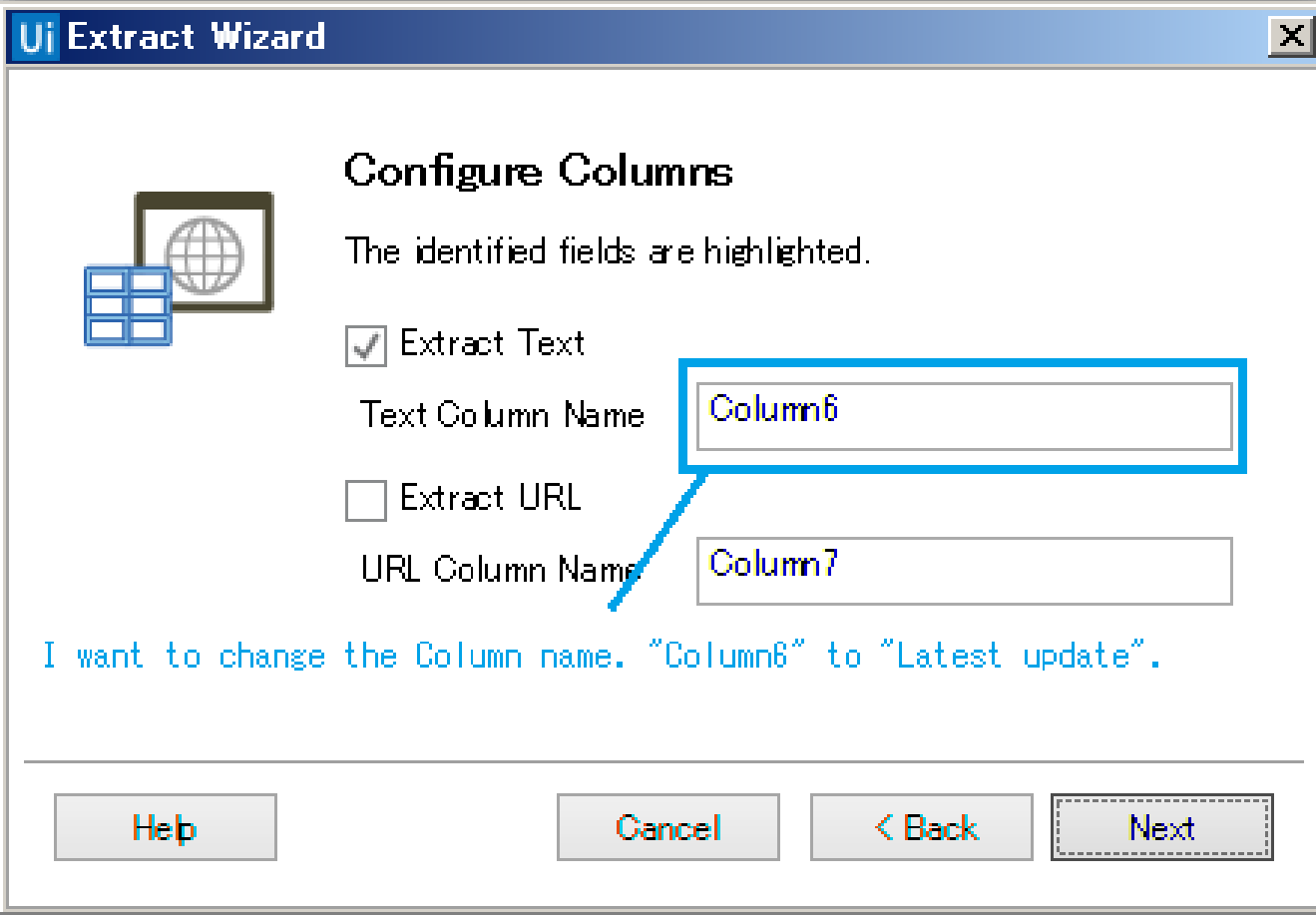| Code			Short description	Latest update	Detailed description
Click a below cell, and wait for few minutes, and then display "Wizard".					
EP AA		P P	PRIORITY	May 97	📄 PDF
EP AA	A	P P	PRIORITY (PATENT APPLICATION)	May 97	📄 PDF
EP AA	A3	P P	DIVIDED OUT OF	May 97	📄 PDF
EP AA	D	P P	PRIORITY (APPLICATION OF DESIGN)	May 97	📄 PDF
EP AA	E	P P	PRIORITY (EXHIBITION PRIORITY)	May 97	📄 PDF
EP AA	U	P P	PRIORITY (APPLICATION OF UTILITY MODEL)	May 97	📄 PDF
EP AA	W	P P	PCT-APPLICATION	May 97	📄 PDF
EP AC		P R	DIVISIONAL APPLICATION (ART. 76) OF:	May 97	

Step38 : Change the Column Name."Column6" to "Latest update"

Step39 : Click the "Next" on the Wizard

Step40 : Check the "Preview Data" and then click "Finish"

Step41 : Click the "No" on the "Indicate Next Link" dialog
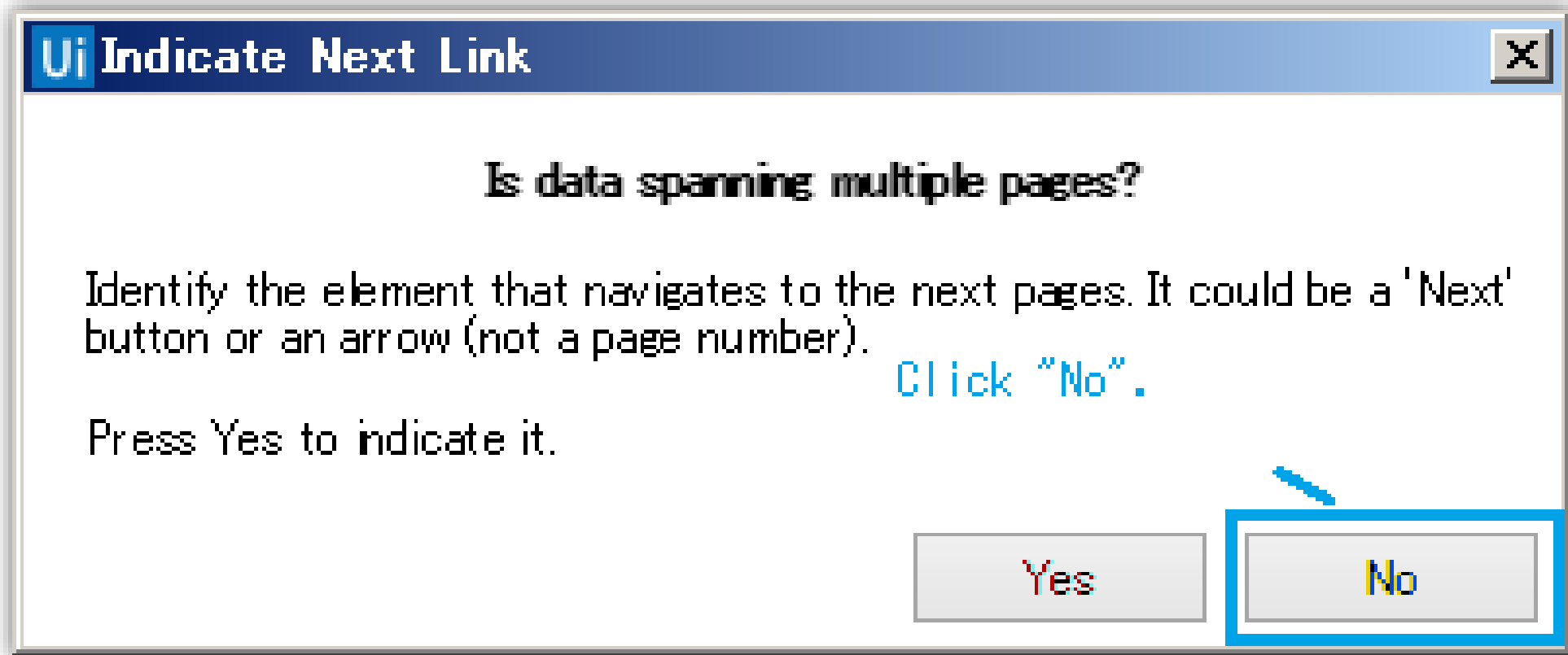
Step42 : On the studio, there is the Sequence activity that is created by the "Data Scraping" Wizard
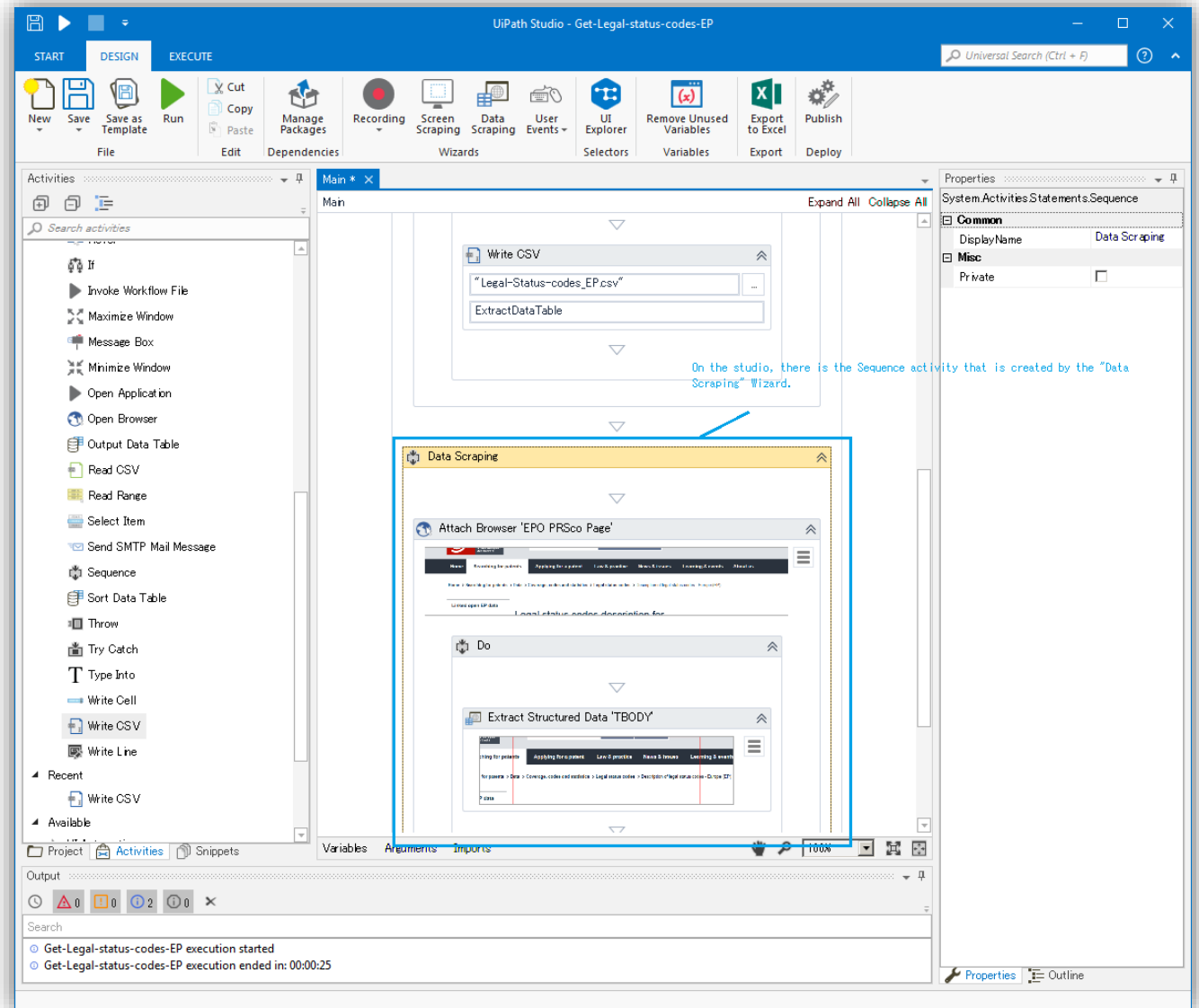
Step43 : Copy the "Extract Structured Data" activity

Step44 : Pest the activity
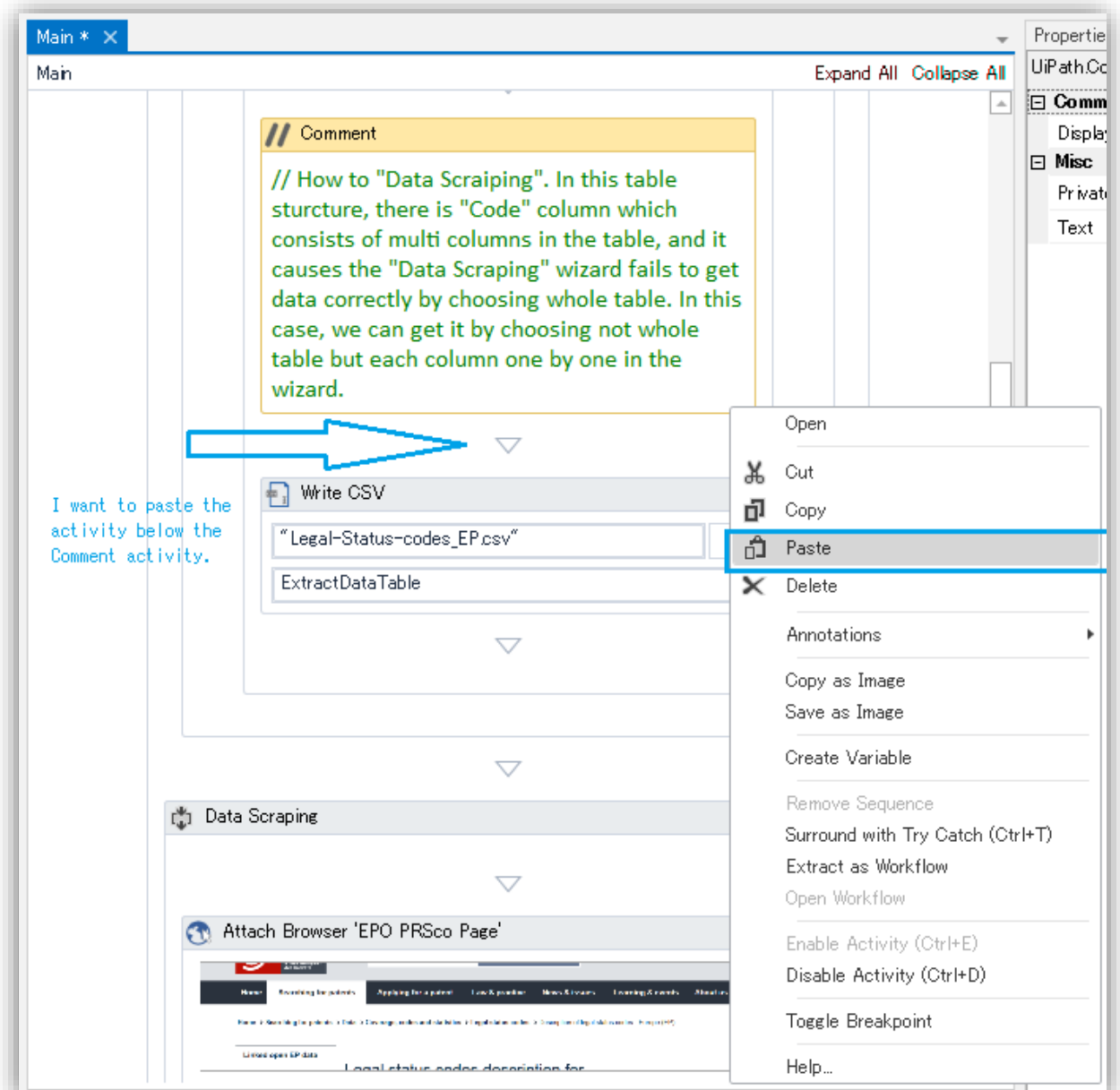
Step45 : Change the "MaxNumberOfR esults" property. 100 to 0.

Step46 : Delete the Sequence activity
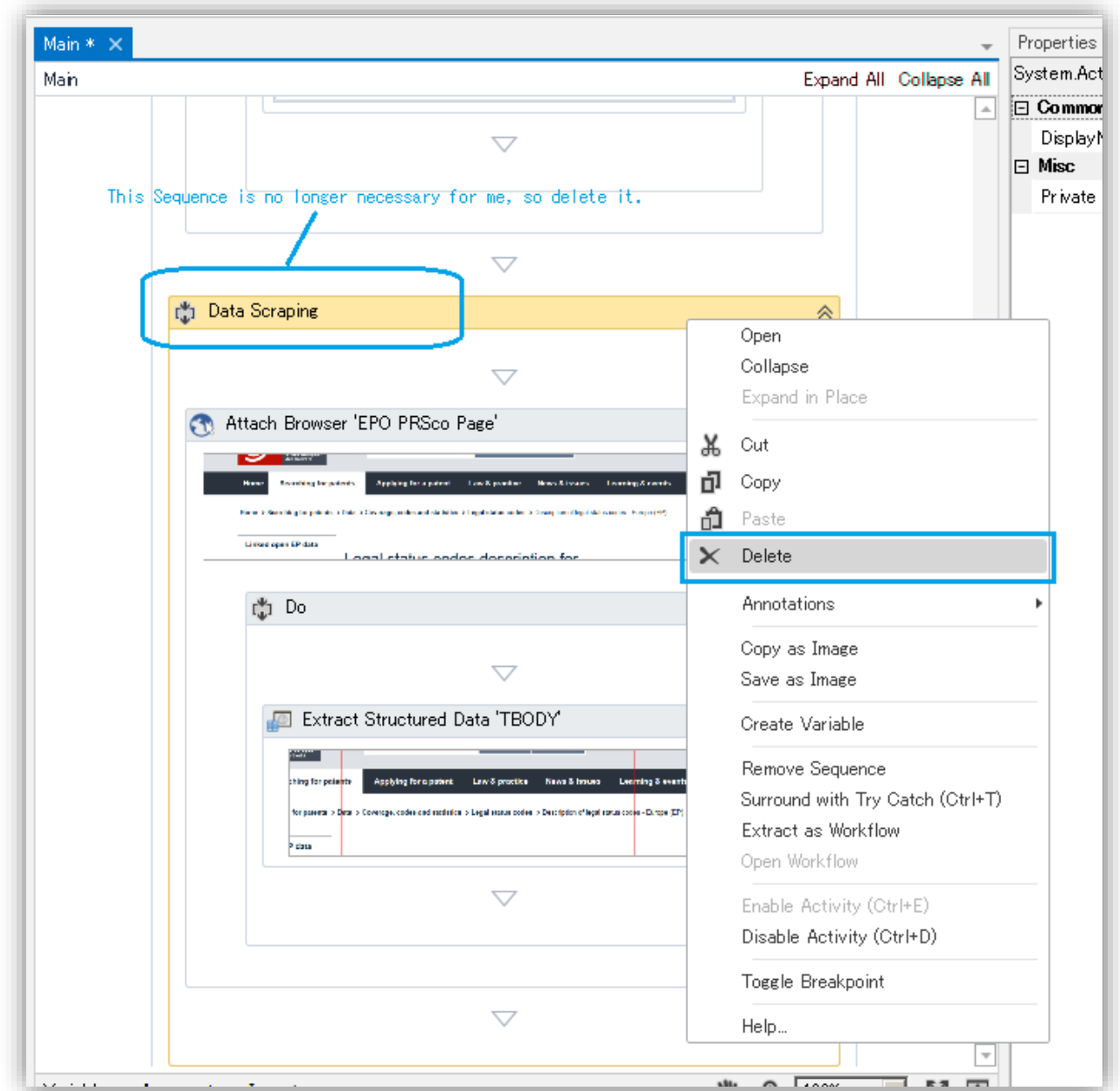
Step47 : Save

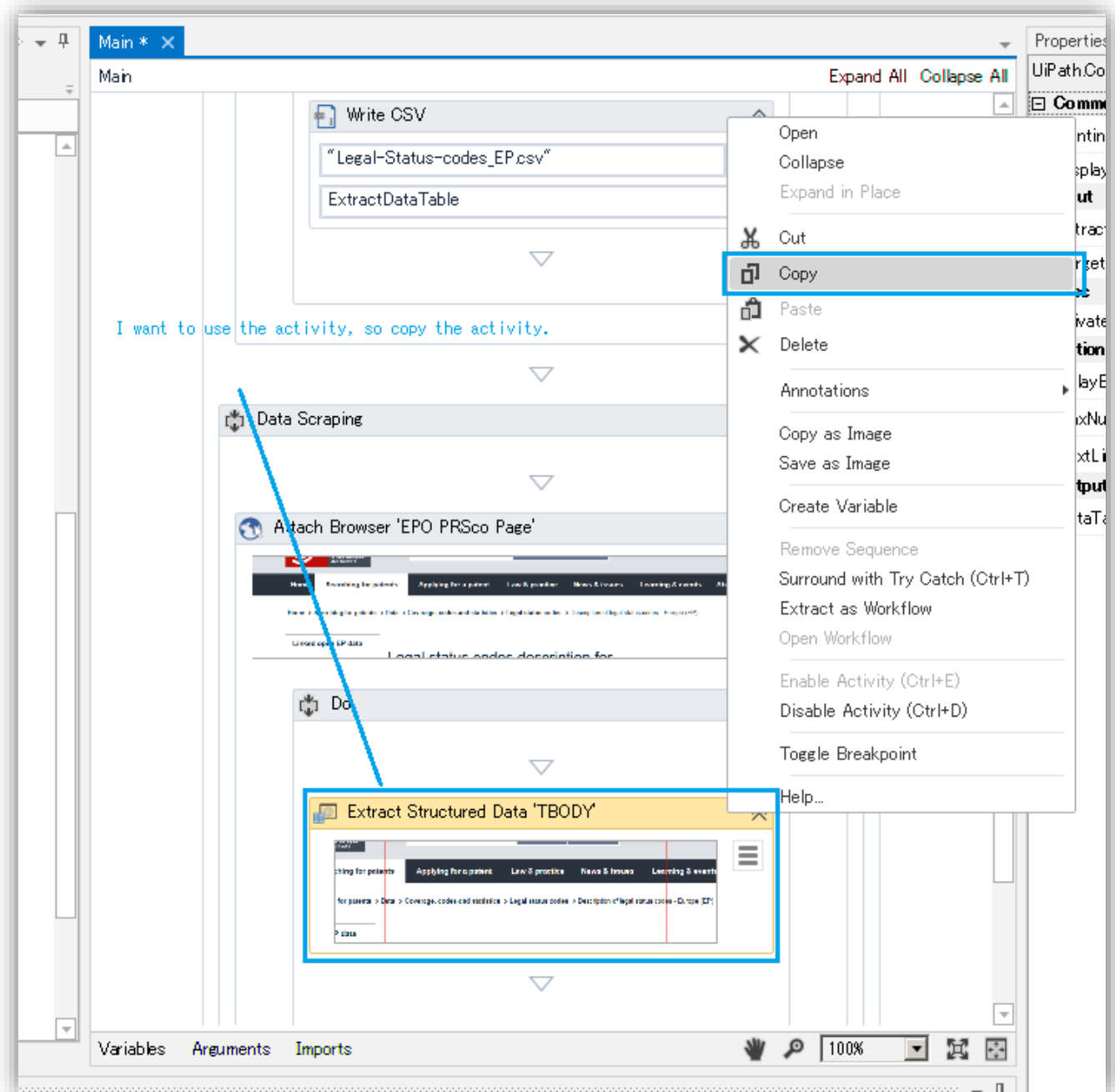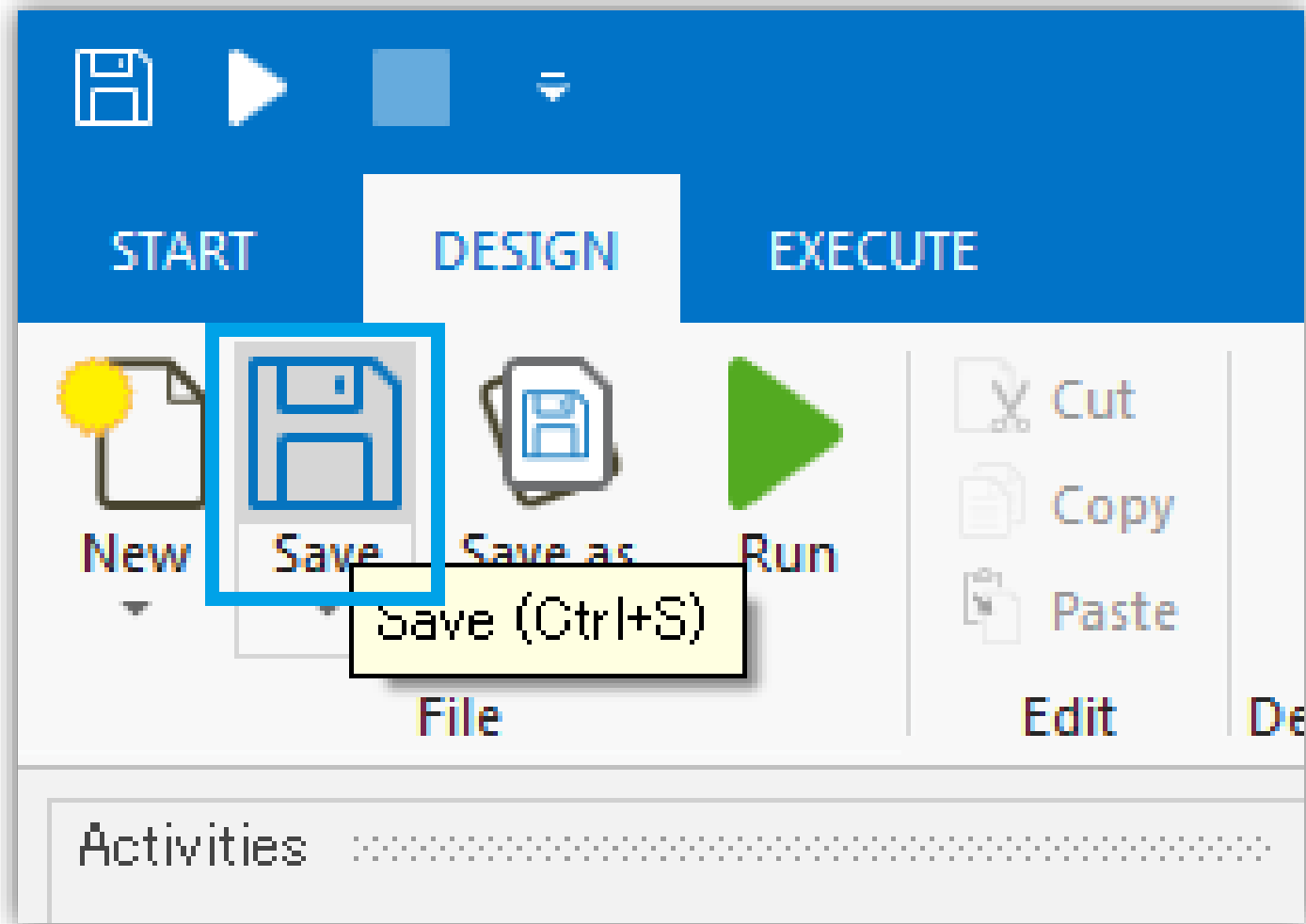# Step48 : Run for test
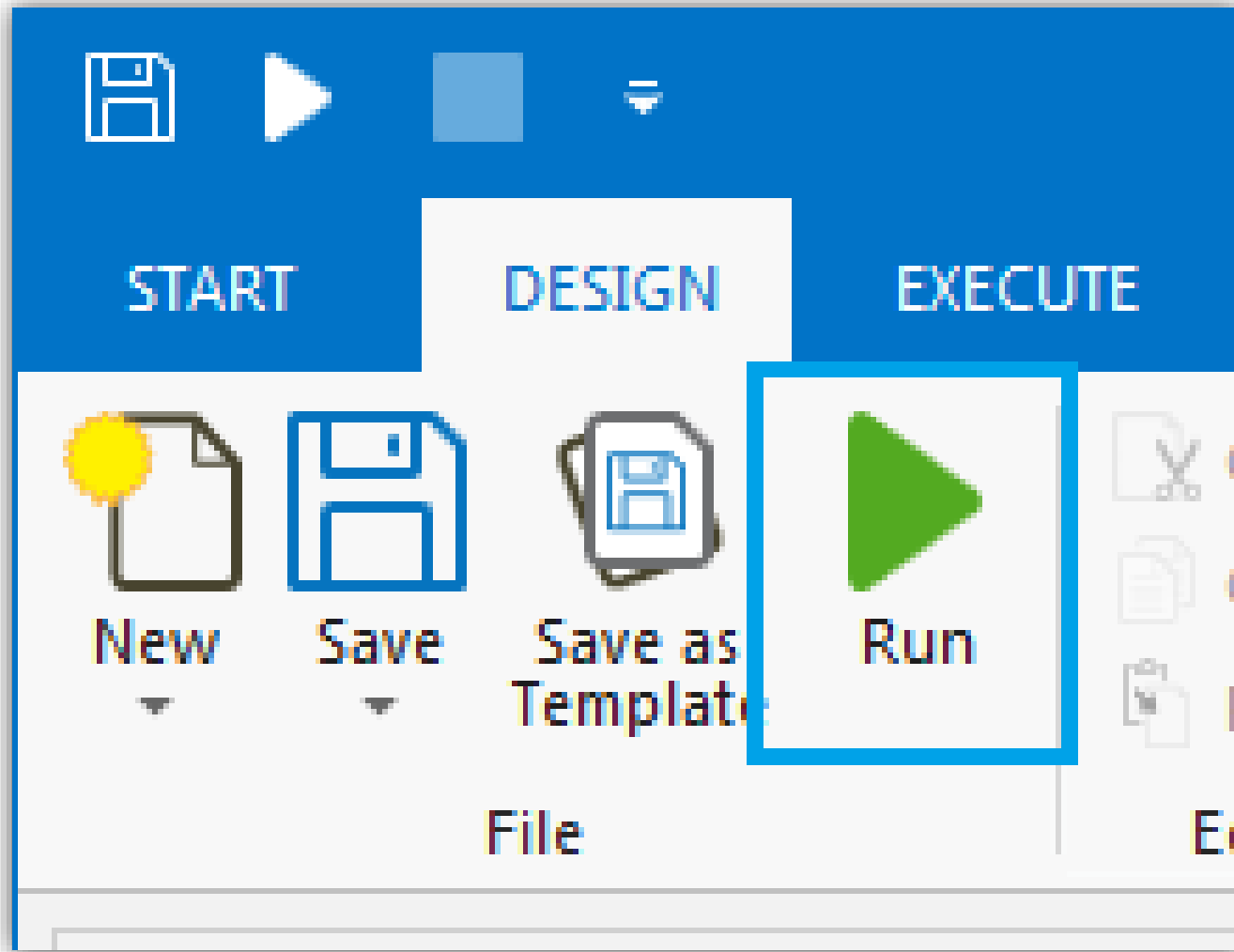
Step49 : Output the csv file

📁	How to scrape data	2019/02/05 12:30
📁	screenshots	2019/02/05 12:06
📁	.local	2019/02/05 9:36
🔵	Main.xaml	2019/02/05 12:28
📊	Legal-Status-codes_EP.csv	2019/02/05 10:40
📄	project.json	2019/02/05 9:35
🖼	DatascrapingOnEP-Page.png	2019/02/04 18:43
🖼	EP-Page.png	2019/02/04 18:33

I got it.

Step50 : Open the csv file on Excel

Done!

	A	B	C	D	E	F
1	Code1	Code2	Code3	Short desc	Latest update	
2	EP AA			PRIORITY	May-97	
3	EP AA	A		PRIORITY	May-97	
4	EP AA	A3		DIVIDED C	May-97	
5	EP AA	D		PRIORITY	May-97	
6	EP AA	E		PRIORITY	May-97	
7	EP AA	U		PRIORITY	May-97	
8	EP AA	W		PCT-APPL	May-97	
9	EP AC			DIVISIONA	May-97	
10	EP AC	P		DIVISIONA	May-97	
11	EP AE			EP-APPLIC	May-97	
12	EP AE	A		EP-APPLIC	May-97	
13	EP AF			SUCCESSI	May-97	
14	EP AF	P		SUCCESSI	May-97	
15	EP AG			SUCCESSI	May-97	
16	EP AG	P		SUCCESSI	May-97	
17	EP AH			DIVISIONA	May-97	
18	EP AH	P		DIVISIONA	May-97	
19	EP AK		+	DESIGNAT	May-97	
20	EP AK	A1	+	DESIGNAT	May-97	
21	EP AK	A2	+	DESIGNAT	May-97	
22	EP AK	A3	+	DESIGNAT	May-97	
23	EP AK	A4	+	DESIGNAT	May-97	
24	EP AK	B1	+	DESIGNAT	May-97	
25	EP AK	B2	+	DESIGNAT	May-97	
26	EP AK	X	−	DESIGNAT	May-97	
27	EP AKNL			NL: CORRE	May-97	
28	EP A1		+	PUBLICAT	May-97	
29	EP A2		+	PUBLICAT	May-97	
30	EP A3		+	SEPARATE	May-97	
31	EP A4		+	SUPPLEME	May-97	
32	EP A5		+	SEPARATE	May-97	
33	EP BECA			BE: CHANG	May-97	