

Award Category | Digital Transformation

Initiative Title | *Parsing Bot*

Name of Awardees | Shithi Maitra, Asst. Manager, CSE; Samia Tasnim, Manager, Procurement Ops.

What is *Parsing Bot*?

Parsing Bot is an in-house developed tool for digging through large, complex documents and extracting out useful information from them, in a tabular form analyzable using Excel. In more technical terms, it is a process that transforms unstructured text into more useful structured data using regular expressions, within a fraction of a second. At Unilever Bangladesh Ltd, this capability has been developed by Supply Chain CoE and is currently being leveraged by the Procurement and Logistics team.

What are *Parsing Bot*'s capabilities?

In simpler terms, *Parsing Bot* takes as input 5 main banks' LCs and accurately outputs 18 useful data points, which would otherwise be a tedious and time-consuming task for a human. The project automatically takes input from Outlook emails, performs the computing on a local machine and communicates the output to respective stakeholders through an automated, periodic email. *Parsing Bot* impacts the following business aspects.

	A	B	C	D	E	F	G	H	I	J
1	pdf_name	dc_no	dc_curr	dc_amt	beneficiary	issue_date	issue_date_refined	insurance_no	insurance_date	insurance_date_ref
2	BDHSBC4A04863380-T02.PDF	DPCDAK400064	USD	11450,00	SARAF FOODS LTD	240104	04-Jan-24	UIC/CZO/MC-2093/12/2023-(CO-20DEC2023	20-Dec-23	
3	BDHSBC3K09285756-T02.PDF	DPCDAK312292	EUR	2309,38	UNILEVER EUROPE BV WEENA	231109	09-Nov-23	UIC/CZO/MC-1846/10/2023-(CO-23OCT2023	23-Oct-23	
4	BDHSBC3K08269895-T02.PDF	DPCDAK312535	USD	83596,80	LUBRIZOL ADVANCED MATERIA	231108	08-Nov-23	UIC/CZO/MC-1887/11/2023-(CO-01NOV2023	01-Nov-23	
5	BDHSBC3K06246318-T02.PDF	DPCDAK312445	USD	52160,00	SARAF FOODS LTD	231106	06-Nov-23	UIC/CZO/MC-1866/10/2023-(CO-30OCT2023	30-Oct-23	
6	BDHSBC3K05229581-T02.PDF	DPCDAK312446	USD	33626,00	UNILEVER INDIA EXPORTS LIM	231105	05-Nov-23	UIC/CZO/MC-1879/10/2023-(CO-31OCT2023	31-Oct-23	
7	BDHSBC3J30170498-T02.PDF	DPCDAK312289	USD	6128,00	INTERNATIONAL FLAVOURS AN	231030	30-Oct-23	UIC/CZO/MC-1845/10/2023-(CO-23OCT2023	23-Oct-23	
8	BDHSBC3J29156155-T02.PDF	DPCDAK312272	USD	18525,00	SARAF FOODS LTD	231029	29-Oct-23	UIC/CZO/MC-1844/10/2023-(CO-23OCT2023	23-Oct-23	
9	BDHSBC3J16032646-T02.PDF	DPCDAK311941	USD	7355,00	SARAF FOODS LTD	231016	16-Oct-23	UIC/CZO/MC-1782/10/2023-(CO-11OCT2023	11-Oct-23	
10	BDHSBC3J10965608-T02.PDF	DPCDAK311709	USD	47376,00	CHEMSPAC CHEMICAL PVT. LT	231010	10-Oct-23	UIC/CZO/MC-1747/10/2023-(CO-10OCT2023	02-Oct-23	
11	BDHSBC3J10965251-T02.PDF	DPCDAK311703	USD	227360,00	DOW CHEMICAL PACIFIC (SING	231010	10-Oct-23	UIC/CZO/MC-1744/10/2023-(CO-10OCT2023	02-Oct-23	
12	BDHSBC3J10965046-T02.PDF	DPCDAK311669	GBP	20290,00	UNIVAR SOLUTIONS B.V. SCHO	231010	10-Oct-23	UIC/CZO/MC-1717/09/2023-(CO-25SEP2023	25-Sep-23	
13	BDHSBC3J09957324-T02.PDF	DPCDAK311638	USD	21822,50	ASIAN DISTRIBUTION TRADE C	231009	09-Oct-23	UIC/CZO/MC-1712/09/2023-(CO-24SEP2023	24-Sep-23	
14	BDHSBC3J09953028-T02.PDF	DPCDAK311667	USD	64550,00	KUSUMA PHARMA ADDRESS: S	231009	09-Oct-23	UIC/CZO/MC-1716/09/2023-(CO-25SEP2023	25-Sep-23	
15	BDHSBC3J05924774-T02.PDF	DPCDAK311620	USD	59371,20	DOW EUROPE GMBH BACHTO	231005	05-Oct-23	UIC/CZO/MC-1711/09/2023-(CO-24SEP2023	24-Sep-23	
16	BDHSBC3J05924697-T02.PDF	DPCDAK311625	USD	55025,00	UNILEVER INDIA EXPORTS LIM	231005	05-Oct-23	UIC/CZO/MC-1719/09/2023-(CO-26SEP2023	26-Sep-23	
17	BDHSBC3J05924456-T02.PDF	DPCDAK311373	USD	16720,00	SOHO ANECO CHEMICALS CO.	231005	05-Oct-23	UIC/CZO/MC-1693/09/2023-(CO-21SEP2023	21-Sep-23	
18	BDHSBC3J05924365-T02.PDF	DPCDAK311377	USD	8866,00	DSM NUTRITIONAL PRODUCTS	231005	05-Oct-23	UIC/CZO/MC-1700/09/2023-(CO-21SEP2023	21-Sep-23	

~150 Manual Hours Saved, Yearly: The month of June 2023 witnessed 150 LCs being issued, the tracking of each of which required ~5 mins for a 3P resource. Along this line, the project saves 150*5*12 mins = 150 man-hours annually.

Readable Data Transformation: Different banks provide numbers, dates etc. in different formats. *Parsing Bot* transforms the data types into one uniform format and compiles all LCs' data in one place for easier tracking.

5 Main Banks Automated: Banks covering 80% of LCs: HSBC, SCB, MTB, Brac Bank Ltd., Prime Bank Ltd. have been automated.

18 Datapoints Parsed: pdf_name, dc_no, dc_curr, dc_amt, beneficiary, issue_date, issue_date_refined, insurance_no, insurance_date, insurance_date_refined, material_name, payment_term, bb_ref, hs_code_importer, hs_code_exporter, hs_code, upas_tenor_1, upas_tenor_2, inv_no, bank_name

Identification of Incomplete Documents: *Parsing Bot* also identifies and reports premature LCs and extracts useful info. only from complete documents.

What business results has *Parsing Bot* delivered?

The business results from *Parsing Bot* are mainly 100% accuracy, easier control, fewer manual hours and error detection. The parsing statistics since 26-Sep-23 are shown in the following table.

Bank Name	LCs Received	LCs Parsed	LCs Incomplete	Incomplete LC Docs	Sec to Parse
HSBC	285	25	260	BDHSBC4A04865892-T02.PDF, BDHSBC4A04864182-T02.PDF, BDHSBC4A04863668-T02.PDF, ...	82
SCB	215	150	65	422010628118-AMD001-MT707-ADV.pdf, 422010627413-AMD001-MT707-ADV.pdf, 422010626735-AMD001-MT707-ADV.pdf, ...	150
BBL	244	4	240	F08102314517D001_SWIFT_071557613.pdf, 308523025845_SWIFT_125158823.pdf, 308523025748_SWIFT_201604644.pdf, ...	46

As shown in the above table, it took just 82 sec for *Parsing Bot* to process 285 HSBC LCs, i.e., ~3 sec per document, which is **100 times faster than a human** taking 300 sec.