



ABBYY RECEIPT CAPTURE SDK 1.1 R2 FOR WINDOWS

RELEASE NOTES

6/11/2018

1 CONTENTS

1	Contents.....	1
2	About This Release	3
2.1	Purpose.....	3
2.2	Key Product Features	3
2.3	Build Information	3
3	Upgrading from Previous Versions and Releases	4
3.1	Binary Incompatibility	4
4	New and Updated Features.....	4
4.1	Release 1	4
4.1.1	Confidence of extracted value for Vendor, Date, Total, and Time fields	4
4.1.2	ZIP, Administrative region, and City components for Address field.....	5
4.1.3	SKU property of a line item.....	5
4.1.4	Region property of a field	5
4.1.5	Support for PDF file format as input.....	5
4.1.6	XML parameters.....	6
4.1.7	XML schema update.....	6
4.1.8	More pre-trained templates for vendors	6
4.1.9	Preview of Line Item measurement units.....	7
4.1.10	CEM support for France	7
4.1.11	Expanded beta-testing program	8
4.2	Release 2	9
4.2.1	CEM support in USA.....	9
4.2.2	CEM support in France update	10
4.2.3	Loyalty Management support in USA update.....	10
4.2.4	Vendor Name classifier update for USA	10
5	Customer Support.....	11
5.1	Contacts.....	11



5.2 Information required..... 11

Appendix A. The list of supported US vendors. 13

2 ABOUT THIS RELEASE

2.1 PURPOSE

This document contains a description of the Release 2 of the ABBYY Receipt Capture SDK 1.1.

The new Release of the product that includes:

- Official support for Corporate Travel Expense Management (CEM) scenario in USA.
- Fine purchase type classification for CEM in France by new field: Guest Count.
- Technical preview of automatic country detection function.

2.2 KEY PRODUCT FEATURES

The product officially supports retail tape receipts from the US and Corporate Travel Expense Management (CEM) scenario for USA and France. All other countries support is for demo purposes only.

Key features:

- Confidence and IsSuspicious properties for Vendor, Date, Total, and Time fields.
- SKU, City, Administrative region, and ZIP fields.
- Region property of a file for easy file coordinates access.
- Support for input in PDF format.
- Improved accuracy for Total, Date, and Line Items fields on Japanese receipts
- Additions to XML exporting parameters to support new field properties.
- XML schema update to reflect new API additions.
- Pre-trained template support for more vendors.
- Measurement unit field preview for LIs.
- Changes to the Demo sample so it displays all extracted fields and hides empty ones.
- Official support for Corporate Travel Expense Management (CEM) scenario for USA and France.
- Fine purchase type classification for CEM in France by new field: Guest Count.
- Technical preview of automatic country detection function.
- New countries in technology preview for beta-testing: Turkey, Poland.
- Improvements in technology preview by prospects' requests for UK and Australia.

2.3 BUILD INFORMATION

Part Number: 1358/7

Build Number:

1.1.173.10

3 UPGRADING FROM PREVIOUS VERSIONS AND RELEASES

3.1 BINARY INCOMPATIBILITY

It is necessary to recompile a host application regardless of the SDK version used previously. Product doesn't work with RC1.0 license. New licenses for RC1.1 are required.

4 NEW AND UPDATED FEATURES

4.1 RELEASE 1

4.1.1 CONFIDENCE OF EXTRACTED VALUE FOR VENDOR, DATE, TOTAL, AND TIME FIELDS

New properties for the Vendor, Date, Total, and Time fields helping with routing of an extracted value were added: if the product is confident with it there is no need to send the value to a human for verification.

For your convenience, there are:

- `IRceiptVendorField::IsSuspicious` returning FALSE if the processing engine is confident that the vendor has been determined correctly. It allows 3% of errors which is known level of human accuracy.
- `IRceiptVendorField::Confidence` returning the confidence of vendor attribution as a number from 0 to 100 which estimates the probability that the vendor is recognized correctly for this receipt. This is for those who want to adjust its own percentage of errors that the product does.

Please remember that allowed % of errors corresponds to % of receipts found. If you allow less errors, you will receive less receipts belonging to a certain vendor. The dependency is not linear so it might be that 1% of errors corresponds to dozens % of extracted receipts.

The property `IsSuspicious` has been trained on so it shows the following accuracy:

	Date	Time	Total	Vendor
Precision	97,47%	96,43%	97,00%	97,24%
Recall	93,36%	98,37%	86,59%	94,73%

Batch	Precision	Recall
Photo.1000	97.75%	93.23%
Photo.11000	99.32%	92.67%
GeneralRetail/USA/NewReceipts/Photo	97.41%	97.41%
GeneralRetail/USA/Photo	96.73%	98.57%
Test.Photo	98.61%	91.02%
Others.Photo	-	-

Precision shows how many correct vendor name you get by filtering the returning values with IsSuspicious=FALSE.

Recall indicates how good the technology finds and extract (not missing) values with IsSuspicious=FALSE.

4.1.2 ZIP, ADMINISTRATIVE REGION, AND CITY COMPONENTS FOR ADDRESS FIELD

In addition to the recognized text API returns components of an address:

- ZIP – postal code found in the recognized text.
- Administrative region – state/region found in the recognized text.
- City – city name found in the recognized text.

4.1.3 SKU PROPERTY OF A LINE ITEM

New property and component of the IReceiptLineItem interface return the stock-keeping unit identifier (SKU) of the product detected in the recognized text.

4.1.4 REGION PROPERTY OF A FIELD

Each recognizable field now has the Region property providing an access to coordinates of a field value on a pre-processed image of a receipt.

This property is helpful if one needs to indicate field values on a receipt image to an end-user in order to improve UX and ease the verification process.

4.1.5 SUPPORT FOR PDF FILE FORMAT AS INPUT

Starting from Beta 8 release RC SDK 1 supports PDF file format for input images. This receipt image format is popular for e-mail attachments.

To enable the feature, one needs a license with corresponding option switched on.

4.1.6 XML PARAMETERS

As RC SDK abilities in XML export grow IXmlExportParams transforms becoming more powerful.

Now instead of simple DetailsLevel parameter one can find many more as Boolean properties at root level:

Parameter	Description
WriteFieldConfidenceLevels	Specifies if the confidence levels and flags, when available, should be saved to the resulting XML file. The default value of this property is FALSE.
WriteRecognizedText	Specifies if the recognized text, for the fields and the receipt as a whole, should be saved to the resulting XML file. The default value of this property is FALSE.
WriteSuspiciousCharacterFlags	Specifies if uncertainly recognized characters should be marked in the recognized text. This property makes sense only if the recognized text is exported, that is, WriteRecognizedText is set to TRUE. The default value of this property is FALSE.

This 'family' of parameters is going to grow more in later releases giving you flexibility in forming desired XML output.

4.1.7 XML SCHEMA UPDATE

XML exporting schema become wider along with new capabilities of the product. From now on you can find the following new elements and attributes in the outputted XML file:

- Vendor field sub-elements:
 - VatNumber
- Address field sub-elements:
 - City
 - ZipCode
 - AdministrativeRegion
- Line Item sub-elements
 - SKU
 - Amount

4.1.8 MORE PRE-TRAINED TEMPLATES FOR VENDORS

Comparing to R1 in this release

- New 19 vendors are supported in the classifier;
- 17 vendors have got the template technology support;
- 11 vendors are removed from the classifier to make it more accurate.

The full list you can find in the [Appendix A](#) to this document.

4.1.9 PREVIEW OF LINE ITEM MEASUREMENT UNITS

The release previews new IReceiptLineItem interface property for extracting product measurement unit. AmountUnits can return one of the following values:

- Unknown
- Piece
- Liter
- Milliliter
- Fluid Ounce
- Gallon
- Kilogram
- Gram
- Quintal
- Pound
- Ounce
- Meter
- Centimeter
- Foot
- Inch
- Point

4.1.10 CEM SUPPORT FOR FRANCE

Corporate Travels Expense Management (CEM) needs an employee to submit a report on its expenses during a business trip. The report includes data on an expense type, date, and amount.

With the help of our customer we shaped out critical requirements to the scenario in France and deliver corresponding technology that extracts:

- Purchase type
- Date of a purchase
- Total amount of a purchase

for 'receipts' (receipts, tickets, bills, ...) from restaurants, gasoline stations, toll roads, car parking, hotels.

As it is very important for meal expense accounting the technology also extracts applied taxes from restaurant receipts.

4.1.11 EXPANDED BETA-TESTING PROGRAM

The technology preview list gets new countries:

- **Turkey.** LM scenario. Expanded receipt fields support for 'A101', 'Bim', and 'Şok' vendors: Cash Register ID and #, Receipt #, and Store #.
- **Poland.** LM scenario. Support for VAT #, Date, Time, and Total fields. Vendor field is trained to 13 names identified by a prospect:

1. VISTULA GROUP
2. TCHIBO WARSZAWA
3. SMYK
4. Sklep Original Marines
5. RESERVED
6. PBH S.A
7. H&M Hennes&Mauritz
8. GREENPOINT
9. DEICHMANN-OBUIWIE
10. COSTA COFFEE
11. CARREFOUR
12. C&A
13. Bijou Brigitte

- **UK.** Previously announced **LM** scenario support has been updated with improved receipt fields extraction accuracy for ASDA and Tesco vendors.
- **CEM** scenario for **Australia**. In a course of PoC (Proof of Concept) RC SDK has support for VAT number field and improved support for Vendor Name field.

The technology was trained to recognition of 25 vendor names in Australia (see the list below).

The functionality seems to be not enough for MVP in CEM scenario, thus it is still in the 'technical preview' status. Further development is scheduled for this year.

The list of vendors the technology is trained for recognition:

1. 7-Eleven
2. ALDI
3. Australia Post
4. Big W
5. Bunnings
6. Caltex
7. Coles
8. Coles Express
9. Dan Murphy's

10. David Jones
11. Foam Coffee Bar
12. IGA
13. Jaycar Electronics
14. JB Hi-Fi
15. Kmart
16. Lucky 7
17. Masters Home Improvement
18. McDonald's
19. Metro Petroleum
20. Officeworks
21. Spar
22. Supercheap Auto
23. United
24. Woolworths
25. Woolworths Petrol

4.2 RELEASE 2

4.2.1 CEM SUPPORT IN USA

CEM (Corporate (Travel) Expense Management) feature helps to automate the process of submitting a travel expense report to your organization after or during a business trip. RC technology is capable of extracting required data from a document picture (scan or photo).

With the technology you can process documents proofing different purchasing types:

- Meal at restaurants
- Fuel purchases at gasoline stations
- Accommodating in hotels
- Taking a taxi

The purchasing types above form the most of a business person expenses in a travel.

Of course more types are there like air/train bookings, toll road passing, etc., but their share is insignificant in reports - one can process them manually w/o noticeable burden.

The following data is extracted automatically from the documents:

- Document/purchase type
- Date/Check-in date/Check-out date
- Total
- Line items: date, description, total
(applicable for hotel bills)

Note: at the moment each page of a multipage hotel bill is processed individually and a RC SDK user is responsible for combining results from all pages.

4.2.2 CEM SUPPORT IN FRANCE UPDATE

The technology is capable of assigning the purchase type 'Restaurant' to receipts from restaurants. This is not enough for the financial law in France: a company needs to attribute a receipt either to a meal (employee nutrition) or to a customer/partner invitation (business meetings).

There is the simple workaround: the technology marks a receipt as a restaurant receipt and extracts a number of guests printed on a receipt, then a customer makes a final decision by the number of guests - 1 guest means 'meal', 2 and more guests mean 'invitation'.

4.2.3 LOYALTY MANAGEMENT SUPPORT IN USA UPDATE

Loyalty Management scenario gets continuous improvements.

In this release:

- JCPenney vendor has got 'templated' technology support which means higher extraction accuracy for all fields and especially for line items and its components. In addition the technology extract 'Cashier' field from receipts of this vendor. The field is required for de-duplication process and fraud protection in LM scenario.
- XML export presents <streetAddress> element containing Receipt::AddressField::RecognizedText value for the sake of clarity. Previously XML element <address> represented both a full vendor address and a street vendor address depending on the technology capabilities for different countries and/or document types. Now a full address goes to <fullAddress>, a street address goes to <address> elements respectively.

4.2.4 VENDOR NAME CLASSIFIER UPDATE FOR USA

Vendor Name field employs several technologies in order to extract correct value from a receipt.

They are:

1. Vendor Logo classifier. It knows some picture specific to particular vendors and can recognize those pictures in a receipt. This is very accurate classifier: if it fails then most probably a receipt belongs to unknown vendor.
2. Receipt Text classifier. It works with full-text OCR result and knows specific words (keywords) for each known vendor. This classifier works if Logo Classifier fails. Moreover, a customer is allowed to append pre-trained classifier or train own one.
3. 'Templated' classifier. It knows a vendor receipt layout: static text, keywords, and field values relative positions. Receipt layouts are pretty unique, but creating them is time consuming, thus the product includes templates for the most important vendors. This classifier works last in a row and adjusts decisions of Logo and Text classifiers.

4. Generic technology finder. This is an algorithm guessing a vendor name out of full-text OCR result by common keywords, position on a receipt, font size, similarity to an URL at the bottom of a receipt, etc. This technology is much less accurate than all above, that is why it works when all above failed.

In this release Text classifier has got 3x times bigger coverage: 337 known vendors. The name list is below.

Notes:

- The updated classifier can't be extended with new names by now, this is scheduled for the next release.
If you need a custom text classifier or want to append pre-trained one please use legacy version of Text classifier which is available provided you set
`IRceiptSynthesisParams::UseAdvancedVendorClassifier = false`.
- The updated classifier works for USA only and it bypasses 'Templated' classifier. Balanced superposition of Logo, updated Text, 'Templated' classifiers is scheduled for the next release.

5 CUSTOMER SUPPORT

The ABBYY SDK Support team is ready to help you. Please refer to the contact information and hours below.

5.1 CONTACTS

ABBYY Technology Portal: <https://abbyy.technology>

Customer Support Management (CSM) Portal: www.abbyy.com/csm

Developer Support for Receipt Capture SDK Beta 9 for Windows:

Email: barbara.gross@abbyyusa.com

Office Hours: Monday – Friday from 9AM to 6PM PST

5.2 INFORMATION REQUIRED

When opening a support case or contacting support, please be prepared to provide the following information:

- Description of the issue

- Sample images for testing
- Run and send Ainfo Report from the Bin/Support folder of the installation directory
- Full error messages that have occurred
- Any additional information you feel may be helpful for the investigation

The information above will assist the ABBYY Support team in investigating your issue and in providing a prompt response.

APPENDIX A. THE LIST OF SUPPORTED US VENDORS.

Added support is marked with **yellow background**. Removed items are ~~stroke out~~.

#	Vendor	Classification support	Template Support
1	7-Eleven	Supported	Supported
2	99 Cents Only Stores	Supported	Supported
3	Ace Hardware	Supported	
4	ACME	Supported	Supported
5	Aeropostale	Supported	
6	Albertsons	Supported	Supported
7	ALDI	Supported	Supported
8	Amazon.com	Supported	
9	Aplus	Supported	Supported
10	Applebee's Neighborhood Grill & Bar	Supported	
11	Apple Store		
12	Au Bon Pain	Supported	
13	Baker's	The Kroger Co.	The Kroger Co.
14	Bashas'	Supported	Supported
15	Bath & Body Works	Supported	Supported
16	Bed Bath & Beyond	Supported	Supported
17	Best Buy	Supported	Supported
18	Big Lots	Supported	Supported
19	Big Y	Supported	
20	BI-LO	Supported	Supported
21	BJ's	Supported	Supported
22	Bloomingdales	Supported	
23	Bunn's Natural Foods		
24	BURGER KING	Supported	Supported
25	CafePress	Supported	
26	Century 21	Supported	
27	Chick-fil-A	Supported	Supported
28	Chipotle	Supported	
29	Circle K	Supported	Supported
30	City Market	The Kroger Co.	The Kroger Co.
31	Cosi	Supported	
32	Costco Wholesale	Supported	Supported
33	Cub Foods	Supported	Supported

34	CVS/pharmacy	Supported	Supported
35	Defense Commissary Agency	Supported	Supported
36	Dillons	The Kroger Co.	The Kroger Co.
37	Dollar General Store	Supported	Supported
38	Dollar Tree	Supported	Supported
39	DSW	Supported	Supported
40	Duane Reade	Supported	
41	Dunkin' Donuts	Supported	Supported
42	eBay	Supported	
43	Einstein Bros. Bagels	Supported	
44	Exchange	Supported	Supported
45	Express		
46	Family Dollar	Supported	Supported
47	Fareway	Supported	Supported
48	Fine Wine & Good Spirits	Supported	Supported
49	Food 4 Less	The Kroger Co.	The Kroger Co.
50	Food Lion	Supported	Supported
51	Foodland	Supported	Supported
52	Fred Meyer	The Kroger Co.	The Kroger Co.
53	Fry's	Supported	Supported
54	Fry's Food Stores	The Kroger Co.	The Kroger Co.
55	Gap		
56	GapKids		
57	Giant Eagle	Supported	Supported
58	GIANT Food Stores	Supported	Supported
59	Grocery Outlet	Supported	Supported
60	Hannaford	Supported	Supported
61	Hardee's	Supported	Supported
62	Harris Teeter	The Kroger Co.	The Kroger Co.
63	H-E-B	Supported	Supported
64	Hilton Hotels	Supported	
65	Homeland	Supported	
66	Hy-Vee	Supported	
67	Ingles	Supported	Supported
68	JCPenney	Supported	Supported
69	Jewel-Osco	Supported	Supported
70	Justice	Supported	
71	Kangaroo Express	Supported	Supported
72	King Soopers	The Kroger Co.	The Kroger Co.
73	Kmart	Supported	Supported
74	Kohl's	Supported	Supported

75	Kroger	The Kroger Co.	The Kroger Co.
76	LOFT	Supported	
77	Lord & Taylor		
78	Lowe's	Supported	Supported
79	Lowe's Foods		
80	Lucky Supermarkets	Supported	Supported
81	LUSH	Supported	
82	Macy's	Supported	Supported
83	Marc's	Supported	Supported
84	Mariano's	Supported	Supported
85	Market Basket	Supported	Supported
86	Mars Super Markets		
87	Marshalls & HomeGoods	Supported	Supported
88	MARTIN'S	Supported	Supported
89	Martin's Super Markets	Supported	Supported
90	McDonald's	Supported	Supported
91	Meijer	Supported	Supported
92	Michaels Stores	Supported	Supported
93	Nordstrom	Supported	Supported
94	Norman's Hallmark	Supported	Supported
95	Panera Bread	Supported	Supported
96	Papa John's	Supported	Supported
97	Party City		
98	Pathmark		
99	Pay Less Super Markets	The Kroger Co.	The Kroger Co.
100	Petco	Supported	
101	PetSmart	Supported	Supported
102	Pick 'n Save	Supported	Supported
103	Piggly Wiggly	Supported	Supported
104	Price Chopper	Supported	Supported
105	Publix	Supported	Supported
106	Ralphs	The Kroger Co.	The Kroger Co.
107	Randalls	Supported	Supported
108	Rite Aid	Supported	Supported
109	Ross Stores	Supported	Supported
110	Safeway	Supported	Supported
111	Sam's Club	Supported	Supported
112	Sbarro	Supported	
113	Schnucks	Supported	Supported
114	Sephora	Supported	
115	Shaw's	Supported	Supported

116	Shoppers Food and Pharmacy	Supported	Supported
117	ShopRite	Supported	Supported
118	Smart & Final	Supported	Supported
119	Smith's	The Kroger Co.	The Kroger Co.
120	South Moon Under		
121	Sprouts Farmers Market	Supported	Supported
122	Staples	Supported	Supported
123	Starbucks	Supported	Supported
124	Stater Bros	Supported	Supported
125	Stein Mart	Supported	
126	Stop & Shop	Supported	Supported
127	SUBWAY	Supported	
128	Taco Bell	Supported	Supported
129	TARGET	Supported	Supported
130	The Home Depot	Supported	Supported
131	The Kroger Co.	The Kroger Co.	The Kroger Co.
132	TJ Maxx	Supported	Supported
133	Tom Thumb	Supported	Supported
134	TOPS	Supported	Supported
135	Trader Joe's	Supported	Supported
136	Uber	Supported	
137	Vons	Supported	Supported
138	Waldbaums	Supported	Supported
139	Walgreens	Supported	Supported
140	Walmart	Supported	Supported
141	Wawa	Supported	
142	Wegmans	Supported	Supported
143	Weis Markets	Supported	Supported
144	Wendy's	Supported	Supported
145	Whole Foods Market	Supported	Supported
146	WinCo Foods	Supported	Supported
147	Winn-Dixie Stores	Supported	Supported