

© 2019 - BIS / Business Imaging Systems, Inc. - Edmond, Oklahoma, U.S.A. All rights reserved.



Grooper WHAT'S NEW IN 2.9

Data Integration Improvements

- New MS Office Integration
- New Box integration
- Enhanced database export
- New document viewer, supporting multiple document views
- Data lookups can now be performed against CMIS sources
- Improved OCR synthesis and paragraph detection
- Classify, Extract, and Data review now support content type filtering

Improvements for Architects

- New Extractor weighting mechanisms
- TF/IDF training builds training batch
- Compile stats command for content containers
- New TF/IDF variants for lexical classification
- New weighted classification properties allow for new classification methodologies
- New Inheritance Architecture

Separation Improvements

- Improved unstructured document separation training
- Improved ESP auto-separation logic
- New settings on document types to aid separation
- New separation review UI
- New Data Annotation Option in data review
- Code expressions support LINQ and String Interpolation





MS OFFICE INTEGRATION

NATIVELY INGEST MULTIPLE FILE FORMATS GENERATED BY THE MS OFFICE SUITE:

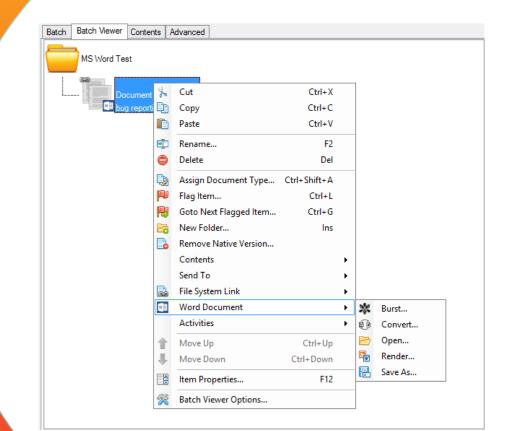
- Word document
- PowerPoint presentation
- Excel spreadsheet
- via CMIS, Outlook messages and mailboxes

NATIVELY PULL TEXT FROM ALL SUPPORTED FILETYPES

REMOVES THE NEED FOR SINGLE-THREADED RENDERING

MS OFFICE MUST BE INSTALLED ON MACHINE RUNNING DESIGN STUDIO

REPRESENTS BETTER INTEGRATION WITH CONTENT PRODUCTION SYSTEMS





NEW BOX INTEGRATION

CONNECT DIRECTLY TO A BOX ACCOUNT TO IMPORT OR EXPORT FILES

WORKS WITH BOX BUSINESS OR ENTERPRISE

TAKES ADVANTAGE OF GROOPER'S EXISTING CMIS+ ARCHITECTURE, GIVING YOU ACCESS TO METADATA MAPPING AND EXPRESSIONS, CMIS LOOKUPS, ETC.

OUR FIRST-EVER INTEGRATION WITH A CLOUD ONLY PROVIDER; POSITIONS DEPLOYMENTS TO TAKE ADVANTAGE OF ONE OF THE LEADING STORAGE PROVIDERS

CAN USE FOLDERS WITH GOVERNANCE LEVEL ASSIGNED TO TAKE ADVANTAGE OF GOVERNANCE FEATURES

MI	S Connection Contents Advanced	
	Save 🔞 Cancel 🛛 🗟 List Reposito	ories 💂 Import Repository
		Connection Properties
⊿	General	
	Connection Type	Box.com
⊿	Connection Settings	(Expand or click to edit)
	Box User ID	
	Repository Root Folder	
	Use Metadata	True
	Description	

		Prope	rties of Database Export Activity	()	
		General Database Connection Content Type Parent Content Types Export Mappings	Transcript (1 Content Type) (0 Export Mapping objects)		
	4	Processing Options Error Disposition	Flag, Log		
<u></u>		Maximum Consecutive Errors Concurrency Mode	0 Multiple		
		Q. France	t Manning Collection Editor		2 X
	Add 🤤	Export Delete 1 Move Up 4 Move Dow			
abase Table	🕽 Add 🤤 Data	Export Delete 1 Move Up 4 Move Dow a Element Scope	n 🖉 🖉 d General		? x
	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	/n ⊿ General Database Table	data	
abase Table	Add 🖨 Data (data	Export Delete 1 Move Up 4 Move Dow a Element Scope	/n ⊿ General Database Table Data Element Scope	data (data model)	
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	/n ▲ General Database Table Data Element Scope Identity Field	(data model)	
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	/n	(data model) (1 Column Mapping	
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	n ▲ General Database Table Data Element Scope Identity Field ▲ Column Mappings ▲ tconst	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	n ▲ General Database Table Data Bement Scope Identity Field ▲ Column Mappings ▲ t const Expression	(data model) (1 Column Mapping	
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	m	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	m	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	/n	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	m	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	n ▲ General Database Table Data Element Scope Identity Field ▲ Column Mappings ▲ tconst Expression ▶ title Type ▶ primary Title ▶ original Title ▶ isAdult	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	In	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	In	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	In	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	In	(data model) (1 Column Mapping Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	m	(data model) (1 Column Mapping Transcript.Full_Name Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	m	(data model) (1 Column Mapping Transcript.Full_Name Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	m	(data model) (1 Column Mapping Transcript.Full_Name Transcript.Full_Name)
abase Table a	Add 🖨 Data (data	Export Delete 1 Move Up J Move Dow a Element Scope a model)	m	(data model) (1 Column Mapping Transcript.Full_Name Transcript.Full_Name)



ENHANCED DATABASE EXPORT

Multiple exports can now be defined on a single database export step

- Must be within a single database; can span multiple tables
- Can choose separate data element scope on each table mapping object
- Greatly simplifies complicated or multipart database exports
- Supports SQL-server generated identity columns

© 2019 - BIS / Business Imaging Systems, Inc. - Edmond, Oklahoma, U.S.A. All rights reserved



NEW DOCUMENT VIEWER

ALLOWS FOR DISPLAY OF MULTIPLE DOCUMENT "RENDITIONS"



NEW VIEWER: NATIVE RENDITION

General	gn Document Type 🏴 Flag Item 🖹 Rem		Native (Mail Message)
Content Type			Native (Mail Message)
Flagged	False	From: Chris Dearner <cdearner@bisok.com></cdearner@bisok.com>	
Description		Sent: Monday, October 7, 2019 2:09 PM	
Attachment			
File Name	Piers Plowman.eml	To: Chris Dearner <cdearner@bisok.com></cdearner@bisok.com>	
File Name Extension	.eml	Subject: Piers Plowman	
MIME Type	message/rfc822		
File Type Description	Mail Message	IN a somer seson,	· · · · · · · · · · · · · · · · · · ·
MIME Type Info	Mail Message	Whan softe was the sonne	
Has Local Copy	True False	I shoop me into shroudes	
Has PDF Version Content Link	CMIS Document Link	As I a sheep weere	
Content Link	CMIS Document Link	In habite as an heremite	
Batch Folder		Unholy of werkes,	
Batch Folder is a container ob	ject within the hierarchical structure of a <u>Batch</u> , wh		
ed to represent logical folders	and multipage documents.	Wondres to here;	
emarks		Ac on a May morwenynge	
atch Folders may contain Batcl	<u>Page</u> objects or other folders as children.	On Malverne hilles	
atch Folder objects may be cre	ated in a variety of ways. If a <u>Separation Profile</u> is a	Me bifel a ferly,	
the Scan activity. then folders	will be created in real time during scanning. The Se	of fairye me thoghte.	
		I was wery for-wandred,	
Piers Plowman.msg (2)		
Piers Plowman.eml		And wente me to reste	
		Under a brood ban	
		By a bournes syde;	
		And as I lay and lenede,	
		And loked on the watres,	
		I slombred into a slepyng,	
		It sweyed so murye.	
		Thanne gan I meten	
		A merveillous swevene,	
		That I was in a wildernesse,	
		Wiste I nevere where,	
		And as I biheeld into the eest	
		An heigh to the sonne,	
		I seigh a tour on a toft	
		Trieliche y-maked,	
		A deep dale bynethe,	
		A dongeon therinne,	
		With depe diches and derke	
		And dredfulle of sighte.	
		A fair feeld ful of folk	
		Fond I ther bitwene,	
		Of alle manere of men,	
		The meene and the riche,	
		Werchynge and wandrynge,	
		As the world asketh.	

NEW VIEWER: ALTERNATIVE RENDITION

Save i Cancer i 🖓 Assig	n Document Type 🏴 Flag Item 📑 Remove N	ative version
General		RFC822 Source
Content Type		
Flagged	False	Content-Type: multipart/alternative; boundary="6a869d85c4844656aa2914b34f7a5f61"
Description		From: Chris Dearner <cdearner@bisok.com></cdearner@bisok.com>
Attachment		To: Chris Dearner <cdearner@bisok.com></cdearner@bisok.com>
File Name	Piers Plowman.eml	Subject: Piers Plowman
		Date: Mon, 7 Oct 2019 19:09:10 +0000 Received: from SN6PR1901MB2061.namprd19.prod.outlook.com
File Name Extension	.eml	(2603:10b6:803:2d::12) by SNGPR190IH202.nampro19.prod.outlook.com with
MIME Type	message/rfc822	HTTPS via SNAPRRØ71LCA032.NAMPRDØ7.PROD.OUTLOOK.COM; Mon, 7 Oct 2019 19:09:11
File Type Description	Mail Message	
MIME Type Info	Mail Message	MIME-Version: 1.0
Has Local Copy	True	
Has PDF Version	False	6a869d85c4844656aa2914b34f7a5f61
Content Link	CMIS Document Link	Content-Type: text/html; charset="utf-8"
Batch Folder		<pre><html xmins="http://www.w3.org/1K/REC-htm140" xmins:m="http://schemas.microsott.com/ot+lce/2004/12/omm1" xmlns:o="urn:schemas-microsoft-com:office:office" xmlns:v="urn:schemas-microsoft-com:vml" xmlns:w="urn:schemas-microsoft-
</pre></td></tr><tr><td>Batch Folder is a container obje</td><td>ect within the hierarchical structure of a Batch, which is</td><td>com:ot+lce:word"><head><meta< td=""></meta<></head></html></pre>
ed to represent logical folders ar		name=Generator content="Microsoft Word 15 (filtered medium)"> <style><!</td></tr><tr><td>marks</td><td></td><td>/* Font Definitions */</td></tr><tr><td></td><td>Page objects or other folders as children</td><td>@font-face</td></tr><tr><td></td><td>Page objects or other folders as children.</td><td>{font-family:"Cambria Math";</td></tr><tr><td></td><td>ed in a variety of ways. If a <u>Separation Profile</u> is assign</td><td></td></tr><tr><td>the Scan activity, then folders w</td><td>vill be created in real time during scanning. The Separat</td><td></td></tr><tr><td></td><td></td><td>{font-family:Calibri;</td></tr><tr><td>Piers Plowman.msg (2)</td><td></td><td>panose-1:2 15 5 2 2 2 4 3 2 4;}</td></tr><tr><td>(2)</td><td></td><td>@font-face</td></tr><tr><td>Piers Plowman.em</td><td></td><td>{font-family:"Roboto Light";</td></tr><tr><td></td><td></td><td>panose-1:0 0 0 0 0 0 0 0 0;}</td></tr><tr><td></td><td></td><td>/* Style Definitions */</td></tr><tr><td></td><td></td><td>p.MsoNormal, li.MsoNormal, div.MsoNormal</td></tr><tr><td></td><td></td><td>{margin:0in;</td></tr><tr><td></td><td></td><td>margin-bottom:.0001pt;</td></tr><tr><td></td><td></td><td>font-size:11.0pt; font-family:"Calibri",sans-serif;}</td></tr><tr><td></td><td></td><td></td></tr><tr><td></td><td></td><td>a:link, span.MsoHyperlink {mso-style-priority:99;</td></tr><tr><td></td><td></td><td>(mso-style-priority:99; color:#0563C1;</td></tr><tr><td></td><td></td><td>color:##05051; text-decoration:underline;}</td></tr><tr><td></td><td></td><td>avisited, span.MsoHyperlinkFollowed</td></tr><tr><td></td><td></td><td>(mso-style-priority:99;</td></tr><tr><td></td><td></td><td>color:#95472;</td></tr><tr><td></td><td></td><td>text-decoration:underline;}</td></tr><tr><td></td><td></td><td>span.EmailStyle17</td></tr><tr><td></td><td></td><td><pre>(mso-style-type:personal-compose;</pre></td></tr><tr><td></td><td></td><td>font-family:"Calibri",sans-serif;</td></tr><tr><td></td><td></td><td>color:windowtext;}</td></tr><tr><td></td><td></td><td>.MsoChpDefault</td></tr><tr><td></td><td></td><td><pre>{mso-style-type:export-only;</pre></td></tr><tr><td></td><td></td><td><pre>font-family:"Calibri",sans-serif;}</pre></td></tr><tr><td></td><td></td><td>@page WordSection1</td></tr><tr><td></td><td></td><td>{size:8.5in 11.0in;</td></tr><tr><td></td><td></td><td>margin:1.0in 1.0in 1.0in;}</td></tr><tr><td></td><td></td><td>div.WordSection1</td></tr><tr><td></td><td></td><td>{page:WordSection1;}</td></tr><tr><td></td><td></td><td>></style> [if gte mso 9] <xml></xml>
		<pre></pre>
		[endif] [if gte mso 9] <xml></xml>
		<o:shapelayout v:ext="edit"></o:shapelayout>
		<pre><c:idmap data="1" v:ext="edit"></c:idmap></pre>
		(original viewe care adda i //



CMIS DATA LOOKUPS

Functions similarly to a Database lookup and allows for population or validation of Grooper fields based on metadata located on CMIS objects

 Allows for better integration with CMIS data sources

0

• Allows for treating CMIS data sources as containing queryable information

Da	a Model Extraction Testing Scripting Contents Advance	ced	///
	Save 🔯 Cancel 🕟 Test Extraction		
⊿	General		
	Show Content Type	False	
	Description		
⊿	Multi-Level Display		
	Show Level Header	True	
	Level Header Alignment	MiddleLeft	
	Level Header Background Color	Control	
	Level Header Foreground Color	ControlText	
	Level Header Font Style	Regular	
⊿	Child Element Options		//
	Child Label Width	100	
	Show Fields In Grid	False	
⊿	Behavior		
	Lookups	(0 Lookup Specification objects)	
	Preserve Character Data	False	
	Rubberband OCR Profile		1
⊿	Appearance		
	Background Color	Control	
	Foreground Color	ControlText	
	Label Position	Left	
	Label Alignment	MiddleLeft	

OK S Cancel S Add - Delete Move Up Move Down Itest Lookup Lookup Type Lookup Fields Target Fields CMIS Lookup Itest Lookup Fields Itest Lookup Fields CMIS Lookup Itest Lookup Fields Itest Lookup Fields CMIS Lookup Itest Lookup Options Itest Lookup Options Trigger Mode Auto Moreal Mose Disposition Normal Conflict Disposition Field Population Overwrite Itest Lookup Fields Target Fields Target Fields Target Fields CMIS Lookup Performs a lookup against a CMIS repository. Moreal Remarks Executes a CMISOL Query against a CMIS Repository, and uses the result set to validate and/or populate fields. Query Editing The query editor provides advanced features such as syntax highlighting and intellisense which make ourser creation for a bezimer Soutax Witch make ourser creation for a bezimer Soutax			Lookup Specifi	catior	Collection Editor	?	x
CMIS Lookup Repository CMIS Lookup Options Auto Trigger Mode Auto Miss Disposition Normal Conflict Disposition Normal Conflict Disposition Normal Field Population Overwrite Image: Auto Auto Auto Auto Auto Auto Auto Auto	🗿 OK 🔞 Cancel	🕒 Add 🗕 😑 Delete 🛛	Move Up 📕 Move Down			🖉 Test Lo	okup
The query editor provides advanced features such as syntax highlighting and intellisense	Lookup Type		-	A A A A A A A A A A A A A A A A A A A	Repository CMIS Query Lookup Options Trigger Mode Miss Disposition Conflict Disposition Field Population Lookup Info Lookup Fields Target Fields MIS Lookup rforms a lookup against a CMIS marks tecutes a <u>CMISOL Query</u> against d'or populate fields.	Auto Nomal Nomal Overwrite	^
				Tł	e query editor provides advance	ed features such as syntax highlighting and intellisens asy and intuitive - even for a beginner Syntax	e Y



IMPROVED OCR SYNTHESIS AND PARAGRAPH DETECTION

Grooper can now detect paragraphs containing large sections of blank underlines correctly

There are additional various improvements to OCR Synthesis

		Pattern Editor	? X
📀 OK 🙁 Cancel		128 results in 37 ms Ref.	rence Material
Pattern Editor Properties		Image Text	
4 General			/5 📫 🕨
Value Type	String		
Mode	RegEx		~
Case Sensitive	False		
⊿ Preprocessing Options	Paragraphs		
Tab Marking	Disabled		
⊿ Paragraph Marking	Enabled		
Maximum Line Spacing	0.4		=
Line Spacing Ratio	125%		
Minimum Line Width	3		
Line Wrap Threshold	60%	OIL AND GAS AGREEMENT	
Maximum Horizontal Gap	0.5	OTE AND AND AGITELINEITT	
Consider Underline	True		
Indent Size	0.2		
Detect Bullets	True	THIS LEASE, Agreement, made and entered into this 13th day of April .	
Consider Underline		2009, by and between Charles E. Putt	
Type: Boolean, Default: Faise		Of 1598 Upper Birch Run Road, Allegany, New York 14706, hereinafter called the "Lessor" or	
	Iorizontal Gap will be made in cases where		
gap is underlined.	torizontal Gap will be made in cases where	Lessons, and onyder Diomens, inc. with onices at F. O. Dox 1022, 90 Glade Drive, Kittanning,	
Remarks		Pennsylvania 16201, hereinafter called the "Lessee".	
	documents which use underlining to indica		
"fill in the blanks" regions inside parag	aocuments which use underthining to indica		
This contract is between John D		WITNESSETH:	
This contract is between John L	Joe and Jane Smith , an		
atch: Paragraph Detection Test		1. Leasing Clause, Lessors in consideration of One (\$1.00) Dollar in hand paid by the	
Paragraph Detection Test			
		Lessee, receipt of which is hereby acknowledged, grant and convey unto the Lessee, its heirs,	
Paragraph Detection Test		executors, administrators, and assigns, all the oil, gas, surface and Drilling Rights in, on and under	
		executions, administrations, and assigns, an the on, gas, surface and Drining Fights in, on and under	
69		all that certain piece, parcel or tract of land situated in Allegany Township,	
Document (1)		Cattaraugus County, New York, bounded and described as follows:	
The second			
		On the North by lands of <u>David Ennis;</u>	
		Scale: 63 % 1685px x 2800px 8.43" x 14.00" 200 DPI Black & White	
		Results (128) Confidence Page No Index Length	
		James K. Griffith Instrument Number 100 % 1 0 35	
		CATTARAUGUS COUNTY CLERK *124386-004* 100 % 1 36 38	
		Cattaraugus County Center 303 Court Street 100 % 1 75 43	
		Little Valley, NY 14755 100 % 1 119 24	
		07(5) 338 9111 100 % 1 144 17	
		Fax: (76)3382773 100 % 1 162 20	
		100 % 1 183 1	



IMPROVEMENTS FOR ARCHITECTS



COMPILE STATS

COMPILE STATS PROVIDES COMPREHENSIVE STATISTICS ON CLASSIFICATION AND EXTRACTION ACROSS AN ENTIRE CONTENT MODEL

ASSISTS ADMINISTRATORS IN DEVELOPING AND TROUBLESHOOTING CONTENT MODELS, ALLOWING FOR MORE TARGETED WORK AND SHORTER DEVELOPMENT TIMES.

			Text Vie	wer	?	
📀 ок						
4,344 OBJECTS:						
OBJECT TYPE	COUNT					
Batch Page	985					
Data Type	862					
Data Format	755					
Training Page	752					
Batch Folder	308					
Data Field	151					
Page Type	138					
Document Type	109					
Folder	91					
Data Model	88					
Form Type	74					
Lexicon	13					
Content Category	8					
Batch Process Step	4					
Field Class	4					
Batch	1					
Batch Process	1					
DOCUMENT TYPE			XTRACT COUNT	FIELD NA	MES	
ATTACH•ACK	10	1		DocType		
ATTACH-ACK ATTACH-BLM Attachment		1		DocType		
ATTACH•BEN ACCACHMENC		1		DocType		
				DocType		
ATTACH•SIG						
ATTACH•SIG BLM•CORR BLM	2	1	0	DocType		
ATTACH•SIG BLM•CORR BLM BLM•LEASE CORR BLM	2	1	0 0	DocType DocType		
ATTACH•SIG BLM•CORR BLM BLM•LEASE CORR BLM BLM•TITLE ASMT CORR BLM	2 2 3	1 1 1	0 0 0	DocType DocType DocType		
ATTACH•SIG BLM•CORR BLM BLM•LEASE CORR BLM	2 2 3 2	1 1 1 1 1 1	0 0 0	DocType DocType DocType DocType		
ATTACH•SIG BLM•CORR BLM BLM•LEASE CORR BLM BLM•TITLE ASMT CORR BLM BLM•TITLE CA CORR BLM	2 2 3	1 1 1 1	0 0 0 0	DocType DocType DocType DocType DocType	. Sender. Recipient. Date Sent	
ATTACH•SIG BLM•CORR BLM BLM•LEASE CORR BLM BLM•TITLE ASMT CORR BLM BLM•TITLE CA CORR BLM BLM•TITLE MERGER CORR BLM	2 2 3 2 2 2	1 1 1 1 4	0 0 0 0 3	DocType DocType DocType DocType DocType DocType,	, Sender, Recipient, Date Sent	
ATTACH-SIG BLM-CORR BLM BLM-LEASE CORR BLM BLM-TITLE ASMT CORR BLM BLM-TITLE CA CORR BLM BLM-TITLE MERGER CORR BLM CORR-CORR EML	2 2 3 2 2 53	1 1 1 1 4	0 0 0 3 0	DocType DocType DocType DocType DocType DocType, DocType	, Sender, Recipient, Date Sent , Sender, Recipient, Document Date	
ATTACH-SIG BLM-CORR BLM BLM-LEASE CORR BLM BLM-TITLE ASMT CORR BLM BLM-TITLE CA CORR BLM BLM-TITLE HERGER CORR BLM CORR-CORR EML CORR-CORR FML	2 2 3 2 2 53 12 79	1 1 1 1 4 1 4	0 0 0 3 0 3	DocType DocType DocType DocType DocType DocType, DocType DocType,		
ATTACH-SIG BLM-COR BLM BLM-CHR BLM BLM-TITLE ASMT CORR BLM BLM-TITLE CA CORR BLM BLM-TITLE CA CORR BLM CORR-CORR FAX CORR-CORR FAX CORR-CORR FAX CORR-CORR LTR	2 2 3 2 2 53 12 79	1 1 1 1 1 4 1 4 2	0 0 0 3 0 3	DocType DocType DocType DocType DocType DocType, DocType, DocType,	, Sender, Recipient, Document Date	
ATTACH-SIG BLM-CORR BLM BLM-LEASE CORR BLM BLM-TITLE ASMT CORR BLM BLM-TITLE CA CORR BLM BLM-TITLE KRGER CORR BLM CORR-CORR FML CORR-CORR FML CORR-CORR FAX COYER SHEET CORR-CORR LTR CORR-CORR MAIL RECEIPT	2 2 2 2 2 2 53 12 79 19	1 1 1 1 4 1 4 2 4	0 0 0 3 0 3 1 3	DocType DocType DocType DocType DocType, DocType, DocType, DocType, DocType,	, Sender, Recipient, Document Date , Document Date	

					~ .		
de	av N	Javi	rat	tor.	Set	ttin	nc

? X

×

V

v

0	OK 🔀 Cancel	
⊿	General	
	Flag Invalid Items	False
	Processing Level	Level1
	Display Parent Folder(s)	False
	Auto-Load Next Invalid Document	False
	Content Type Filter	(0 Content Type objects)
⊳	Flag Messages	(empty)

Content Type Filter

Type: List of <u>Content Type</u> An optional list of <u>Content Types</u> to which the review process should be restricted.

Remarks

When this list is empty, all documents in scope will be included in the review process. If a list is provided, only document types matching one of the list entries will be included.

	Pro	operties of Extract Activity
⊿	General	
	Mode	Normal
	Default Content Type	
	Flag Invalid Items	False
	Content Type Filter	(0 Content Type objects)
⊿	Processing Options	
	Error Disposition	Flag, Log
	Maximum Consecutive Errors	0
	Concurrency Mode	Multiple
-		
	Pro	operties of Classify Activity
⊿	General	
	Content Model Scope	Blump University
	Classification Level	DocType
	Output Level	DocType
	Apply To	All
	Reclassify Mode	Overwrite
	Supress Candidate List	False
	Model Refresh Interval	60
	Deserving Ontions	



CONTENT TYPE FILTERING

SUPPORTED IN EXTRACTION, CLASSIFICATION, AND DATA REVIEW ACTIVITIES

- Appears in "index navigator settings" on Data Review Activity
- Allows for more focused or partial classification, extraction, and review
- Allows classification, extraction, and review to proceed in stages for larger or more complicated projects
- Supports targeted expansion or improvement
- Supports tiered classification decisioning
- Allows for classification to target only unclassified documents ("All," "Classified", or "Unclassified")

© 2019 - BIS / Business Imaging Systems, Inc. - Edmond, Oklahoma, U.S.A. All rights reserved



⊿ General		
Auto Snap	Disabled	
Output Case Remove Control Characters	None	
	False False	
Trim Whitespace Output Confidence	0%	
Confidence Multiplier	1.2	
Output Value	1.4	
Confidence Multiplier		
Confidence Multiplier Type: <u>Double</u> , Default: <u>1</u>		^
Type: <u>Double</u> , Default : <u>1</u>	boost or reduce the confidence of each result.	^
Type: <u>Double</u> , Default : <u>1</u>	boost or reduce the confidence of each result.	^

NEW EXTRACTOR WEIGHTING MECHANISMS

CONFIGURABLE OUTPUT CONFIDENCE AND OUTPUT MUTIPILIERS ON DATA TYPES

- Manually set extraction confidence and multiply confidence by a set amount
- Allows for waterfall extraction techniques: most reliable methods can be backed up by less-reliable methods by using "order by confidence"

© 2019 - BIS / Business Imaging Systems, Inc. - Edmond, Oklahoma, U.S.A. All rights reserved.

Doc	ument Type	Data Element Overrides	Weightings	Contents	Advanced				
*::) X	🔚 Save 💿 Cancel 🛯 🏙 Compile Stats 🛜 Purge Training 🛜 Rebuild Training								
⊿	General								
	Name Field								
	Allow Training			True					
	Page Scope - Data Extraction			(unlimited)					
	Base Content Type			Document					
	Description								
⊿	Classificati	on							
⊳	Positive Extra	actor		Reference	e: Positive	Rule			
	Negative Extractor			(none)					
	Minimum Pages			0					
	Maximum Pag	ges		0					
	Scoring Limit			0%					
4	Separation								
	Pagination			Unstructured					
	Combine Contiguous Documents			False					
	Prioritize EPI			True					
	Training Sco	1	Nomal						
	Secondary P	age Extractor	(none)					

		Result Options ?
0 🖸	K 🔞 Cancel	
⊿ (General	
1	Auto Snap	Disabled
	Dutput Case	None
	Remove Control Characters	False
	Trim Whitespace	False
	Dutput Confidence	0%
	Confidence Multiplier	1.2
(Dutput Value	
	idence Multiplier	
Type	: <u>Double</u> , Default : <u>1</u>	
Type	: <u>Double</u> , Default : <u>1</u>	oost or reduce the confidence of each result.
<i>Type</i> Spec	: <u>Double</u> , Default : <u>1</u>	oost or reduce the confidence of each result.
<i>Type</i> Spec Rem The o	: <u>Double</u> , Default: <u>1</u> ifies a multiplier which can be used to be arks	ied by this value to produce the final output confidence. If a value is



WEIGHTED CLASSIFICATION PROPERTIES AND WATERFALL CLASSIFICATION

ADVANCED CLASSIFICATION STRATEGIES FOR COMPLICATED DOCUMENT SETS

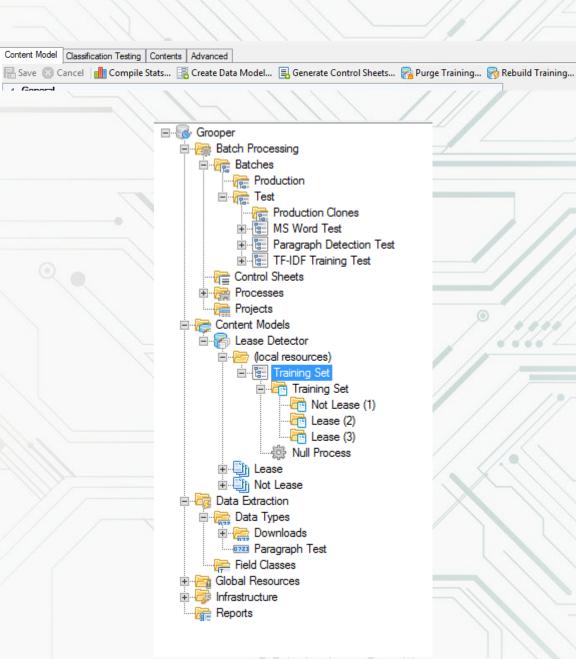
- Positive extractors now classify documents with the confidence of their referenced extractor
- Using the confidence multiplier and output confidence configuration items on data types allows for the creation of waterfall classification strategies
- This allows for documents that are highly lexically dissimilar within a group or highly lexically similar across groups to be targeted with better accuracy (e.g. legal documents, oil and gas leases vs lease acknowledgements, etc.)



AUTOBUILT TF/IDF TRAINING BATCH

TF/IDF training now automatically builds training batch, allowing for easier training after changes to feature collection or TF/IDF settings.

"Rebuild training" allows for tuning and A/B testing using identical training set and document training decisions



_	_	E E LEEN X							
Γ	Cor	tent Model Classification Testing	Contents	Advanced					
📳 Save 🔞 Cancel									
l	⊿	General							
⊿		Classification Method		Lexical					
l		Text Feature Extractor		"\ w +"					
l		Image Feature Extractor							
l		EPI Extractor		(empty)					
l		Bullet Extractor		(none)					
l		Use Class Frequency		True					
l		Use Confidence		False					
l		Term Frequency Mode		Augmented					
l		Frequency Scaling		75%					
l		Document Frequency Mode		Nomal					
l		Default Content Type							
l		Page Scope - Classification		(unlimited)					
l		Page Scope - Data Extraction		(unlimited)					
l		Base Content Type		Document					
l		Description							
l	⊿	Classification Tuning							
l		Minimum Similarity		60%					
l		Minimum Difference		2%					

0%

Use Confidence

Type: <u>Boolean</u>, **Default**: <u>False</u> Specifies how the confidence of each feature instance should be calculated into the term frequency.

Term Frequency Mode

Type: TfModes, Default: Normal

Minimum Training Similarity

Specifies how the term frequency (TF) of features should be calculated. Can be one of the following values:

- Normal Term frequency is normalized to the size of the document. This mode provides matching which is completely independent of document size. In some classification scenarios, however, this mode can produce undesirable false positives, as it allows a very short document to match a very long document with a high degree of similarity.
- · Logarithmic Term frequency is scaled logarithmically. This mode is provided for backwards compatibility.
- Augmented Term frequency is normalized to the frequency of the most-frequent feature. Provides an alternative to Normal
 mode where the size of the document plays more of a role in classification decisions. This mode also provides a Dampering
 Factor

Document Frequency Mode

Type: IdfModes, Default: Normal

Specifies how the inverse document frequency (IDF) of features should be calculated. Can be one of the following values:

- Normal Standard IDF.
- Smooth Smooth IDF.



NEW TF/IDF VARIANTS

- 'Use confidence' allows extraction confidence to factor into TF/IDF feature scoring
- 'Augmented' TF mode (in addition to logarithmic/'sublinear') allows for TF/IDF to partially account for document length (standard TF/IDF tends to favor longer documents)
- Smooth TF/IDF selectable under "Document Frequency Mode"



Document Type Data Element Overrides Weightings Contents Advanced Rave Cancel I Test Extraction E-B LEASE Data Field Preview No fields are visible Filename Lease Number No Issues Document View Text View Instance View /5 📥 🔰 PDF James K. Griffith Instrument Number Test Batch Property Overrides *124386-004* CATTARAUGUS COUNTY CLERK ⊿ General Value Type String Cattaraugus County Center 303 Court Street Little Valley, NY 14755 Value Extractor (none) Minimum Confidence (716) 938-9111 Sub-Element Name Fax: (716) 938-2773 Description Appearance Alignment Left 300 Display Width No. of Pages: Delivered By . SNYDER BROTHERS INC Error Color Window WindowText Background Color Foreground Color 124386 Return To: Receipt No Visible False SNYDER BROTHERS INC Label Position Left Label Alignment MiddleLeft P O BOX 1022 DATE: 07/16/2009 Behavior KITTANNING PA 16201 Auto Complete Disabled ENVELOPE (4) Disabled Input Mask Time: 03:07 PM Data Element Profil Defines one or more property overrides which apply to a Data Element for a particular Content Type. Used By Document Type: LEASE ontent Category, Content Model, Document Type, Form Type, Page Type Parties To Transaction: O&G-PUTT TO SNYDER BROTHERS INC Scale: 50 % 1685px x 2200px 8.43" x 11.00" 200 DPI Black & White

NEW INHERITANCE ARCHITECTURE

DATA ELEMENT PROFILES ARE GONE; LONG LIVE DATA ELEMENT OVERRIDES

- Data element profiles have been replaced by data element overrides
- Every property is now overridable
- Data Element Overrides tab includes built-in tester
- A fuller realization of Grooper's inheritancebased architecture; allows for more robust, more scalable data models.
- Saves time and reduces complexity when architecting large systems



SEPARATION IMPROVEMENTS

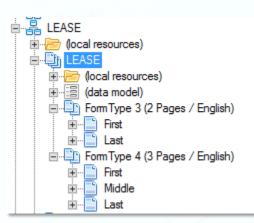
© 2019 - BIS / Business Imaging Systems, Inc. - Edmond, Oklahoma, U.S.A. All rights reserved.



ESP AUTO-SEPARATION IMPROVEMENTS

SEPARATION IS MORE ROBUST AND ACCURATE THAN EVER

- Improved ESP auto-separation logic
- New Secondary Page Extractor
- Unstructured Documents now consider only first, middle, last pages; each n-page version is no longer considered a different formtype



Document Type Data Element Overrides Weightings Contents Advanced

🔚 Save 💿 Cancel 🔚 Compile Stats... 🛜 Purge Training... 🛜 Rebuild Training...

⊿	General				
	Name Field				
	Allow Training	True			
	Page Scope - Data Extraction	(unlimited)			
	Base Content Type	Document			
	Description				
⊿	Classification				
⊳	Positive Extractor	Reference: Positive Rule			
	Negative Extractor	(none)			
	Minimum Pages	0			
	Maximum Pages	0			
	Scoring Limit	0%			
⊿	Separation				
	Pagination	Unstructured			
	Combine Contiguous Documents	False			
	Prioritize EPI	True			
	Training Scope	Normal			
	Secondary Page Extractor	(none)			





New Separation review UI

SIMILAR TO ESP AUTO-SEPARATION TESTER

	Processing Task Batch Help						
	Start Stop						477 c:
		lose Task 📸 Complete Task 🛛					🎇 Settings
	72 53 🖍		138 items 🛛 🗞 🍓 🗔 💯	🖻 🔍 🔍 🍕 🔍 🔍 🖨 🕅 🎇			Display Image
	Name Document Type	Page No EPI	Similarity A				
	Page 1 MOR	1 1 of 2	80%				
	Page 2 MOR	2 2 of 2	0% =				
	Page 3 MOR	1	77%				
	Page 4 TITLE DEED	1 1 of 2	100%				
	Page 5 TITLE DEED (ACK)	2 2 of 2	0%				
	Page 6 -confidence too low- Page 7 MOR		81%		LIMITED OWNE	ERSHIP REPORT	
	Page 7 MOR Page 8 MOR	2 2	0%		LIMITED OWN		
	Page 9 -unknown-	2 2	0%				
	Page 10 LEASE	1	100%			Dated: May 5, 2017	
	Page 11 LEASE	2	0%			Updated: April 23, 2019	· · · · · · · · · · · · · · · · · · ·
	Page 12 LEASE	3	0%			Research Dated: May 4, 2017	
	Page 13 LEASE (ACK)	4	0%			Research Update: April 17, 2019	
	Page 14 -unknown-	1 4	0%				
	Page 15 LEASE	1	65%				
	Page 16 LEASE	2	0%				
	Page 17 LEASE	3	0%	LANDS COVERED BY THIS REPOR	RT:	WELD COUNTY, COLORADO	
	Page 18 LEASE (ACK)	4	0%				
	Page 19 LEASE	1 1 of 4	80%	Township 1 North - Range 67 West, of th	he 6th P.M.		x
· · · · ·	Page 20 LEASE	2 2 of 4 3 3 of 4	0% 0%	Section 33: Lot B of Recorded Exemption	on RECX-16-0076, b	being a part of the SE ¹ / ₄	-
		Search:	0%				
	Document Types			Containing 71.45 acres, more or less			
	Document Type	Similarity	^				
	MOR	80%					
	HIST+HIST LDS LEASE+LEASE RATIF	43%		SURFACE OWNERSHIP:			
	PMTS+PMTS INV	42%		WW, LLC			
-	LEASE LEASE RELEASE	36%		9457 S. University Boulevard, Apt. 401			
				Highlands Ranch, CO 80126			
7	TITLE TITLE STIP	36%					
7.	TITLE+TITLE STIP TITLE+TITLE AOGL	36% 34%		Highlands Ranch, CO 80126			
7	TITLE • TITLE STIP TITLE • TITLE AOGL TITLE • TITLE AORI	36% 34% 34%		Highlands Ranch, CO 80126			
1	TITLE • TITLE AOGL	34% 34% 32%		0			
1 in the second se	TITLE • TITLE AOGL TITLE • TITLE AORI	34% 34%		MINERAL OWNERSHIP:			
1	TITLE+TITLE AOGL TITLE+TITLE AORI PMTS+PMTS RECEIPT	34% 34% 32%	v	MINERAL OWNERSHIP:			
, i i i i i i i i i i i i i i i i i i i	TITLE • TITLE AOGL TITLE • TITLE AORI PMTS-PMTS RECEIPT TITLE • TITLE ASMT OP RTS General	34% 34% 32%	v V	0	INTEREST	LEASEHOLD STATUS	
Í	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name	34% 34% 32%	Core Process	MINERAL OWNERSHIP:	INTEREST	LEASEHOLD STATUS	
i	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name Batch Name	34% 34% 32%	Separation Test	MINERAL OWNERSHIP:			
· ·	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name Batch Open Time	34% 34% 32%		MINERAL OWNERSHIP: OIL AND GAS OWNERSHIP Timothy J. Pivonka	INTEREST 34,00%	LEASEHOLD STATUS HELD BY PRODUCTION	
Í.	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name Batch Open Time Review	34% 34% 32%	Separation Test	MINERAL OWNERSHIP: OIL AND GAS OWNERSHIP Timothy J. Pivonka 2116 14 th Street SE			
i i i	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name Batch Open Time	34% 34% 32%	Separation Test	MINERAL OWNERSHIP: OIL AND GAS OWNERSHIP Timothy J. Pivonka			
i i i	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name Batch Open Time Review	34% 34% 32%	Separation Test 00:01:17 1	MINERAL OWNERSHIP: OIL AND GAS OWNERSHIP Timothy J. Pivonka 2116 14 th Street SE Loveland, Colorado 80537			
i i i i i i i i i i i i i i i i i i i	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name Batch Open Time Review	34% 34% 32%	Separation Test 00:01:17 1 Scale: 67 % 11	MINERAL OWNERSHIP: OIL AND GAS OWNERSHIP Timothy J. Pivonka 2116 14 th Street SE Loveland, Colorado 80537 00px x 2200px [8.50" x 11.00" [200 DPI [24-Bit RGB]			
	TITLE-TITLE AOGL TITLE-TITLE AORI PMTS-PMTS RECEIPT TITLE-TITLE ASMT OP RTS General Batch Process Name Batch Open Time Review Unseparate	344, 344, 327, 327,	Separation Test 00:01:17 1 Scale: 67 % 11 2 2 2 2 2	MINERAL OWNERSHIP: OIL AND GAS OWNERSHIP Timothy J. Pivonka 2116 14 th Street SE Loveland, Colorado 80537	34.00%		



© 2019 - BIS / Business Imaging Systems, Inc. - Edmond, Oklahoma, U.S.A. All rights reserved.

	CATTARAL Cattaraugus C Litt	mes K. Griffith JGUS COUNTY CLE county Center 303 Court St le Valley, NY 14755 (716) 938-9111 Fax: (716) 938-2773		
No. of Pages:	4	Delivered By:	SNYDER BROTHERS INC	
Receipt No.	124386	Return To:		
			ROTHERS INC	
DATE:	07/16/2009	P O BOX 1	022 NG PA 16201	
		ENVELOPE		6
Time:	03:07 PM			
				1
Document Type:	LEASE			1
	LEASE			01
Parties To Trans	ac Test Field G-PUTT TO	SNYDER BROTHERS INC	C	

Deed init	onnation	Mortgage Informa		
Consideration:	\$0.00	Mortgage Amount		
Transfer Tax:	\$0.00	Basic Mtge. Tax:		
RETT No:	03172	Special Mtge. Tax:		
		Additional Mtge. Tax:		
State of New Yo	rk			

Cattaraugus County Clerk

Mortgage Serial No.:

This sheet constitutes the Clerk endorsement required by Section 316-A(5) & Section 319 of the Real Property Law of the State of New York. DO NOT DETACH Activate Windows



NEW DATA ANNOTATION OPTION IN DATA REVIEW

DISPLAYS EXTRACTED DATA AT THE EXTRACTION LOCATION ON THE DOCUMENT

- Greatly speeds up review
- Configurable properties include color, location
- Can disable for simple extraction cases or documents.



CODE EXPRESSIONS: LINQ AND STRING INTERPOLATION SUPPORT IN EXPRESSIONS

DO MORE WITH EXPRESSIONS

- .NET string interpolation now supported in expressions
- LINQ now supported in expressions

Expressions

Default Value Expression Generated Default Value Calculate Expression Validate Expression **\$"Today is {DateTime.Today}"** Today is 1/27/2020 12:00:00 AM

....