Aquaforest

Making SharePoint
Content Findable

Office 365

SharePoint

Microsoft Azure

OCR PDF

Taxonomy

Tags

Entities

Aquaforest Searchlight

Automated Metadata and OCR

C#263
Rules

Zonal

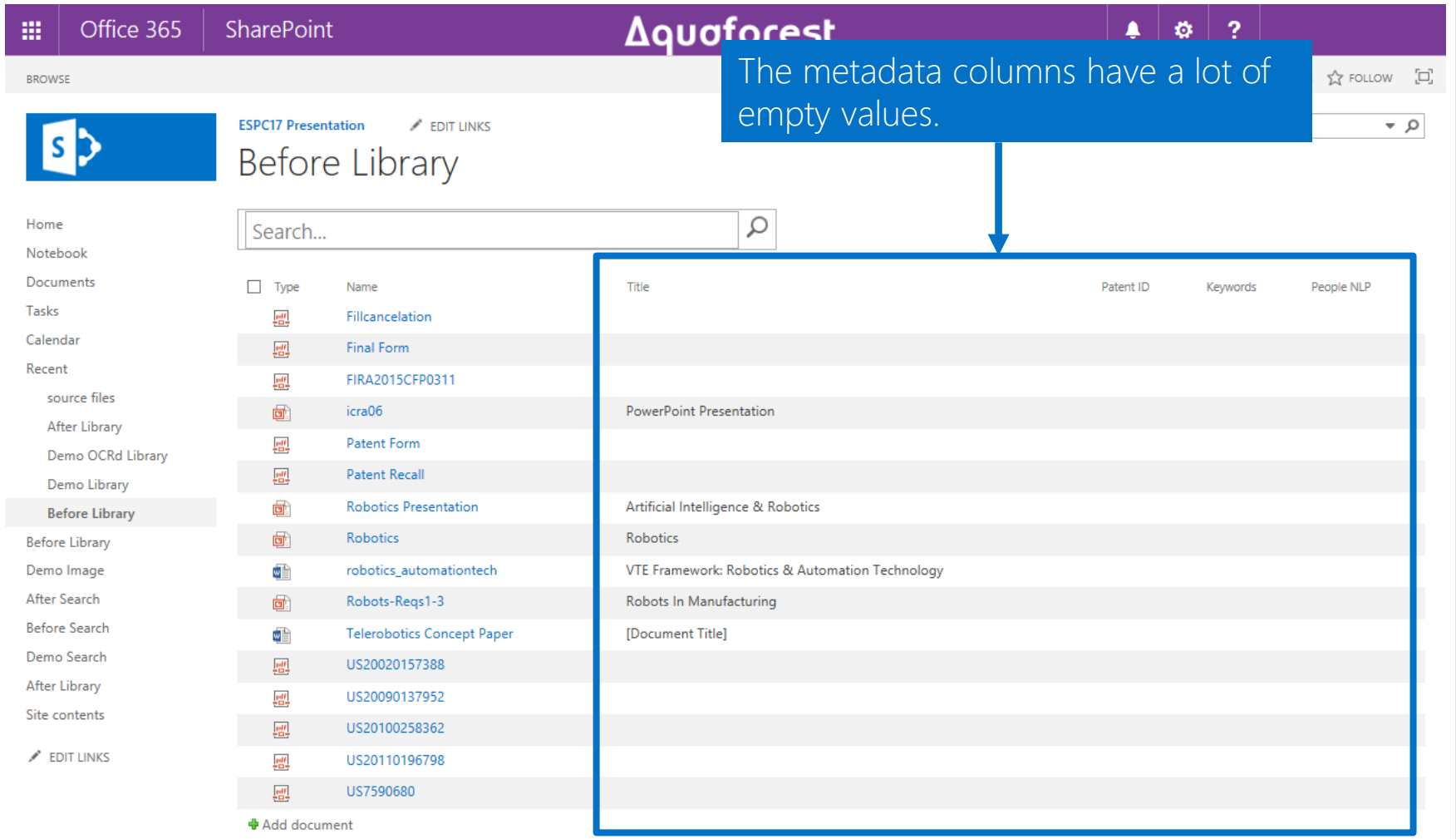
Forms

Barcode

Dramatically Improve your Search with Aquaforest Searchlight.

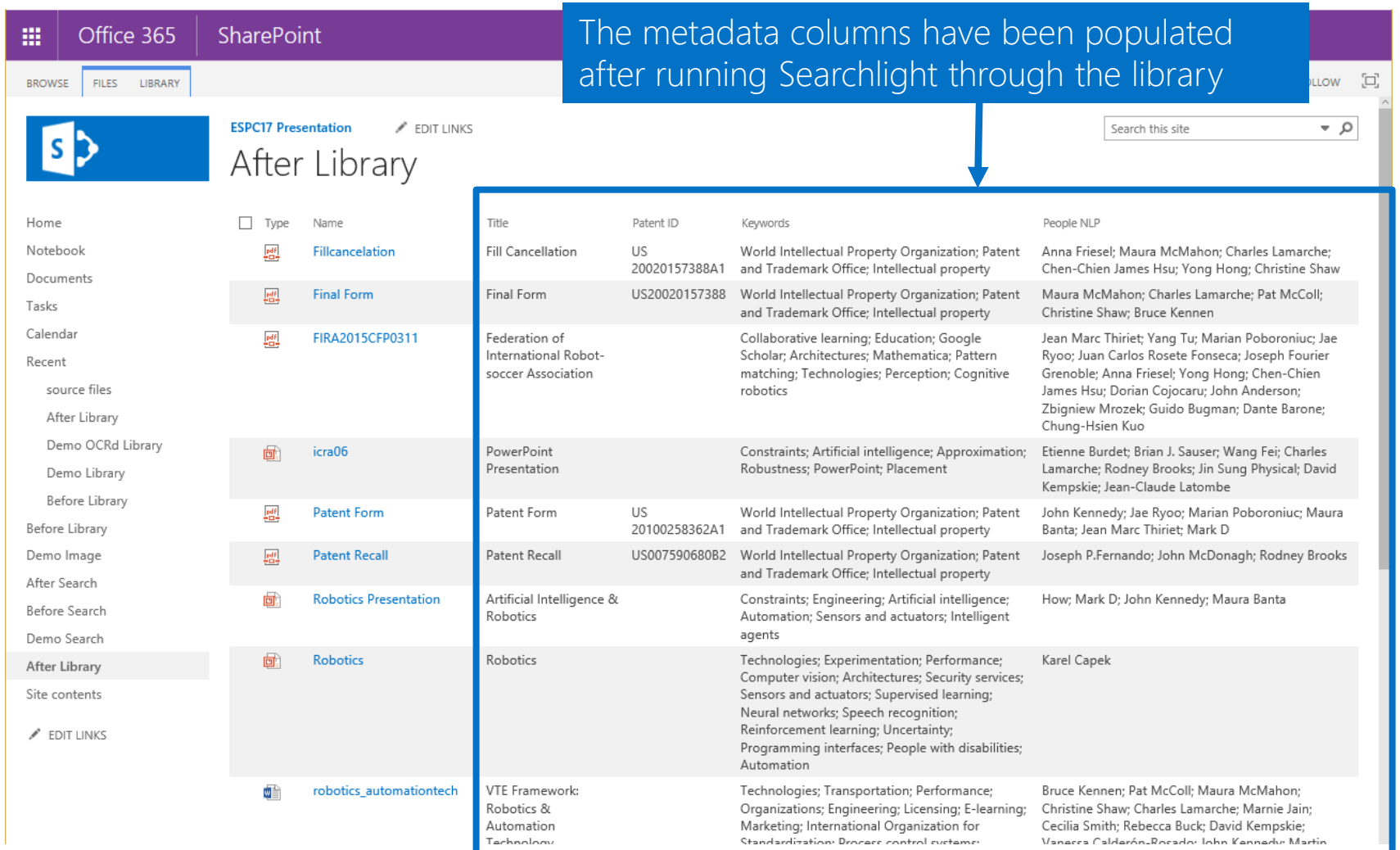
Aquaforest

Before Automated Metadata



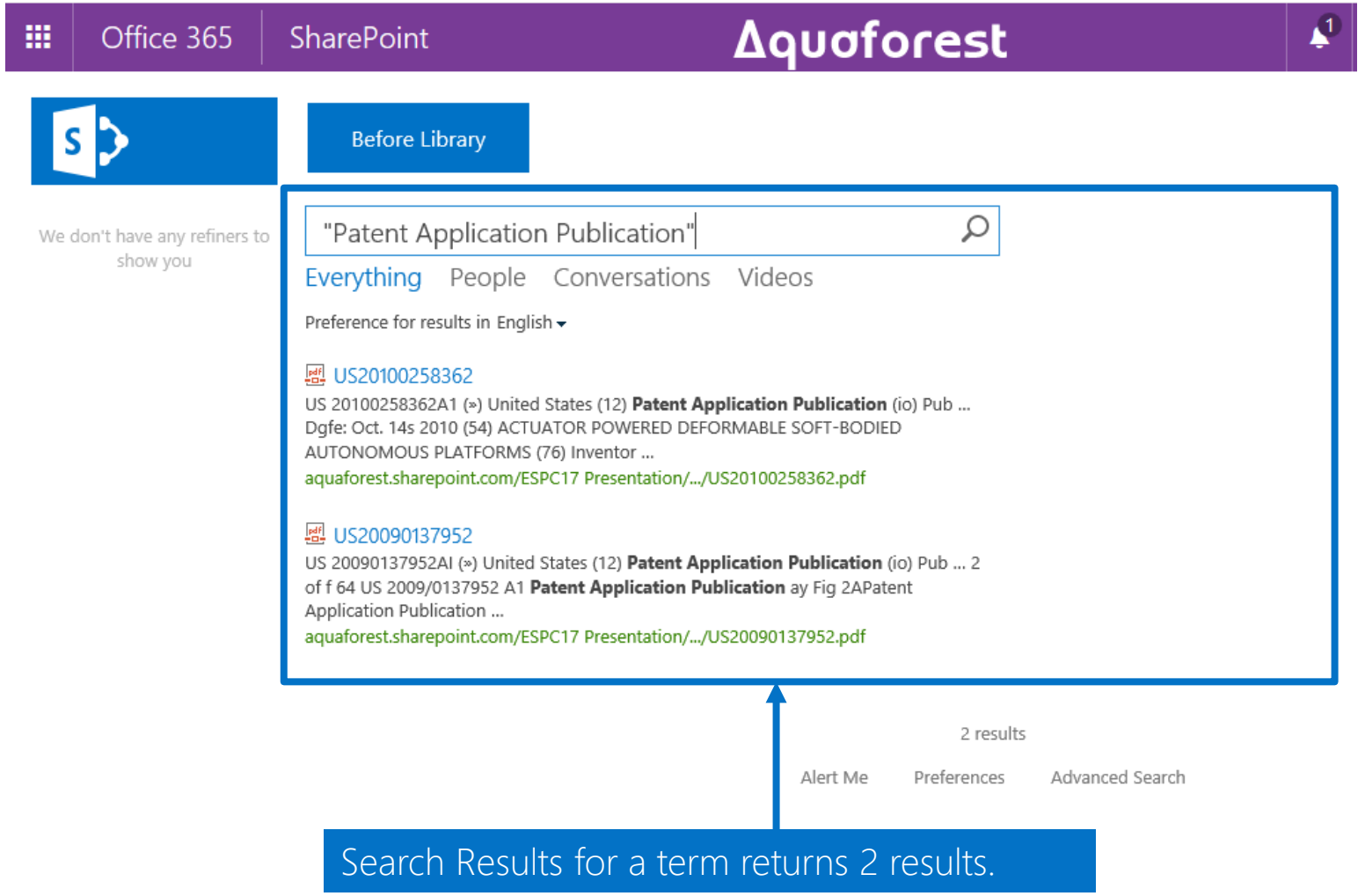
Aquaforest

After Automated Metadata



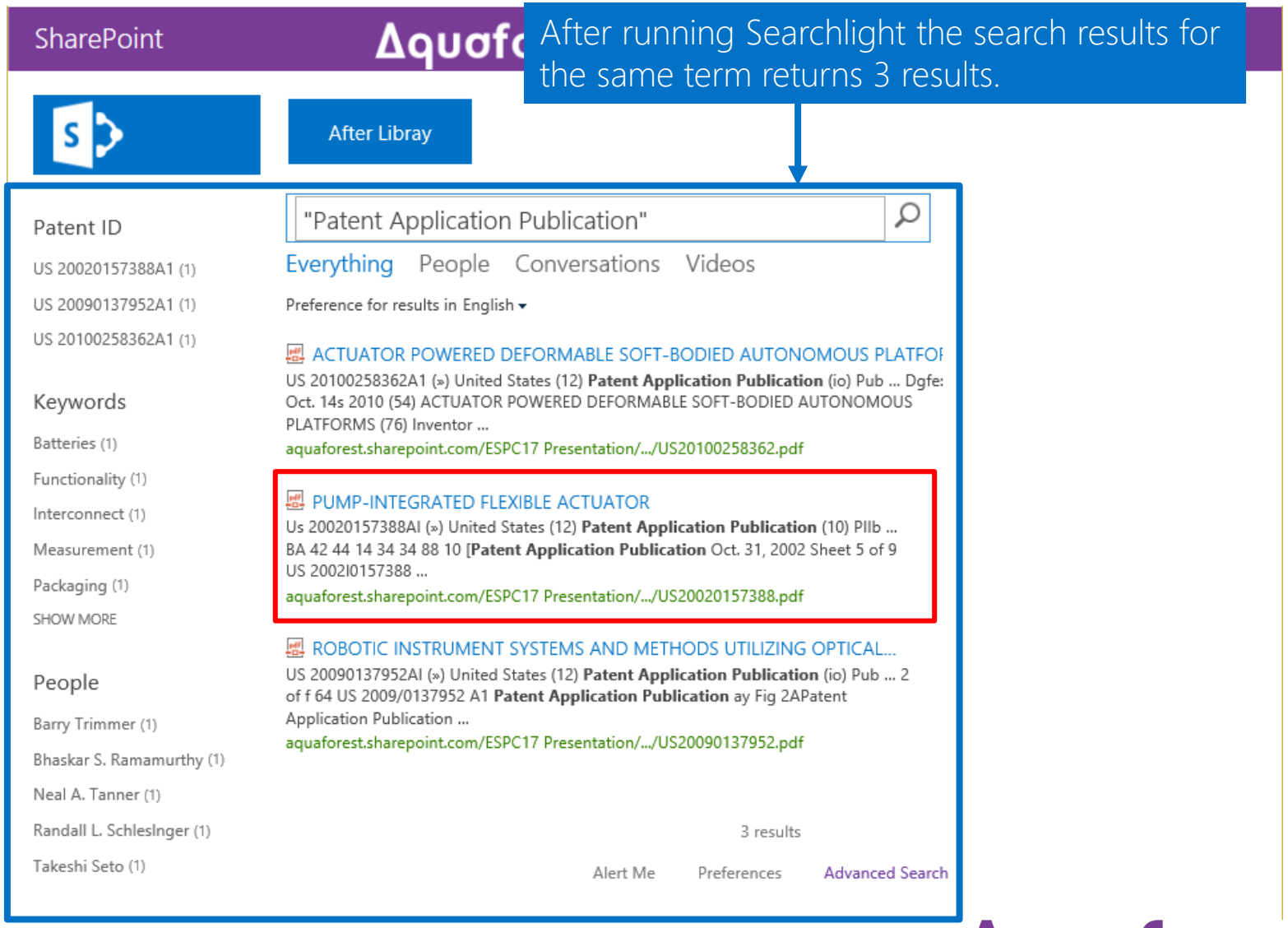
Aquaforest

Before Searchable PDF Processing



Aquaforest

After Searchable PDF Processing



Aquaforest

How does it work?

Aquaforest

Aquaforest

Making SharePoint
Content Findable

Office 365

SharePoint

Microsoft Azure

OCR PDF

Taxonomy

Tags

Entities

Aquaforest Searchlight

Automated Metadata and OCR

C#263
Rules

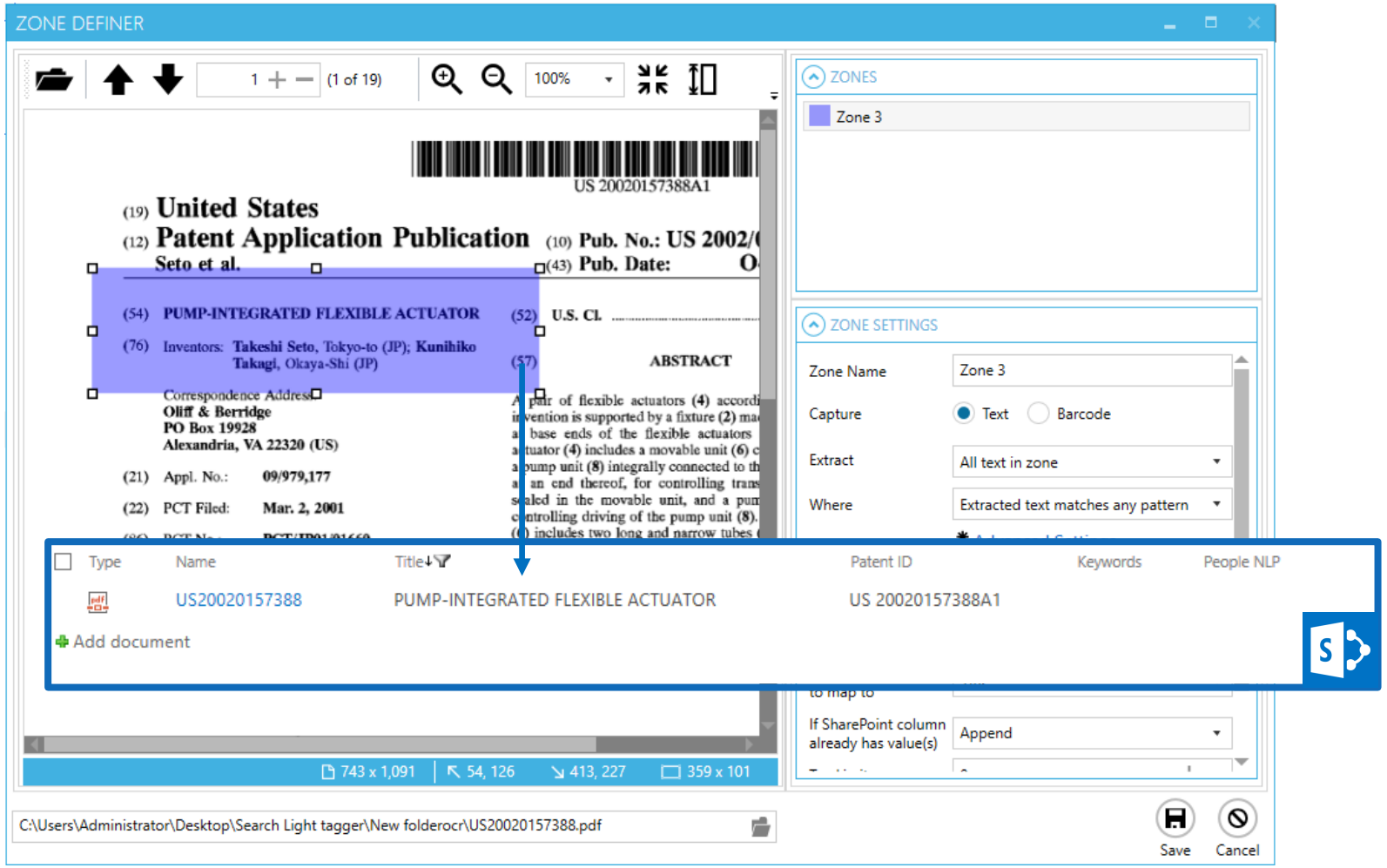
Zonal

Forms

Barcode

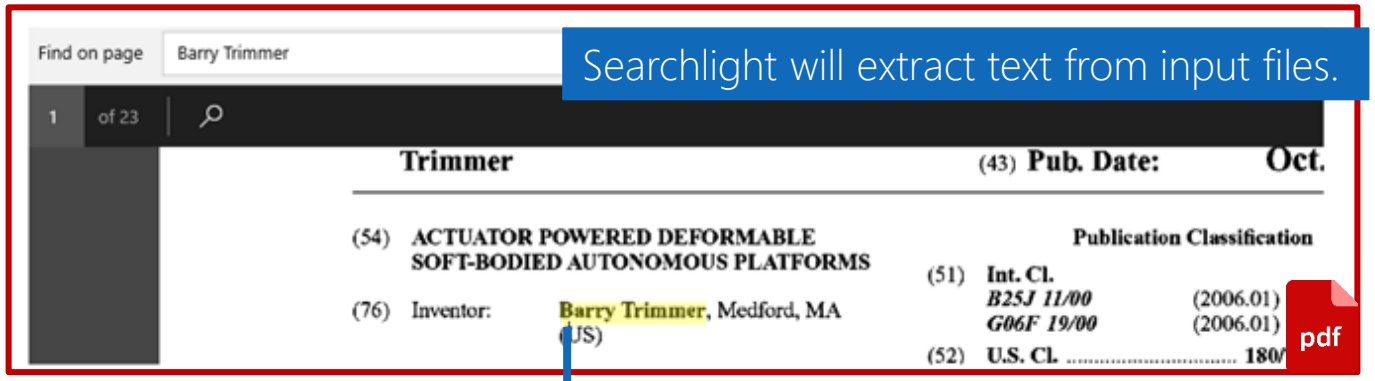
Zonal Extraction

Extract text from zones based on text rules and regular expressions

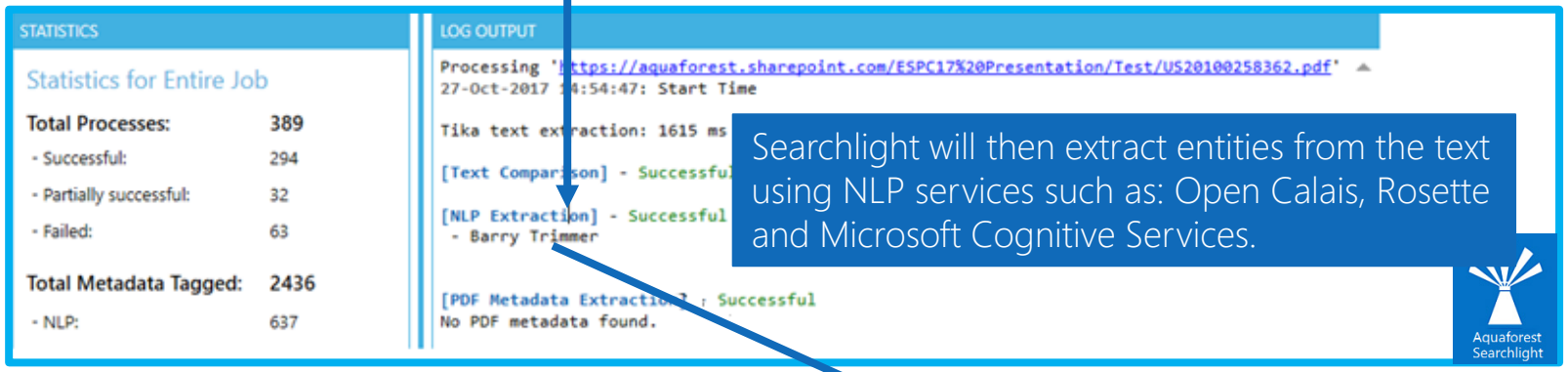


Aquaforest

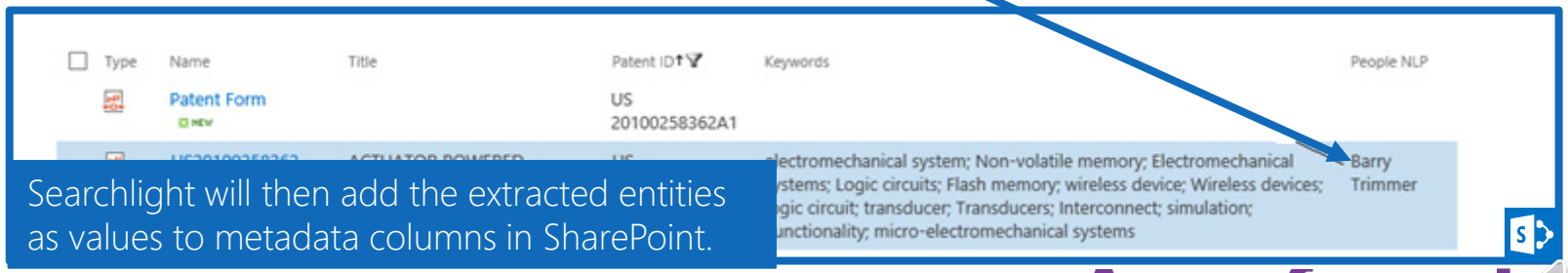
Entity Extraction



Searchlight will extract text from input files.



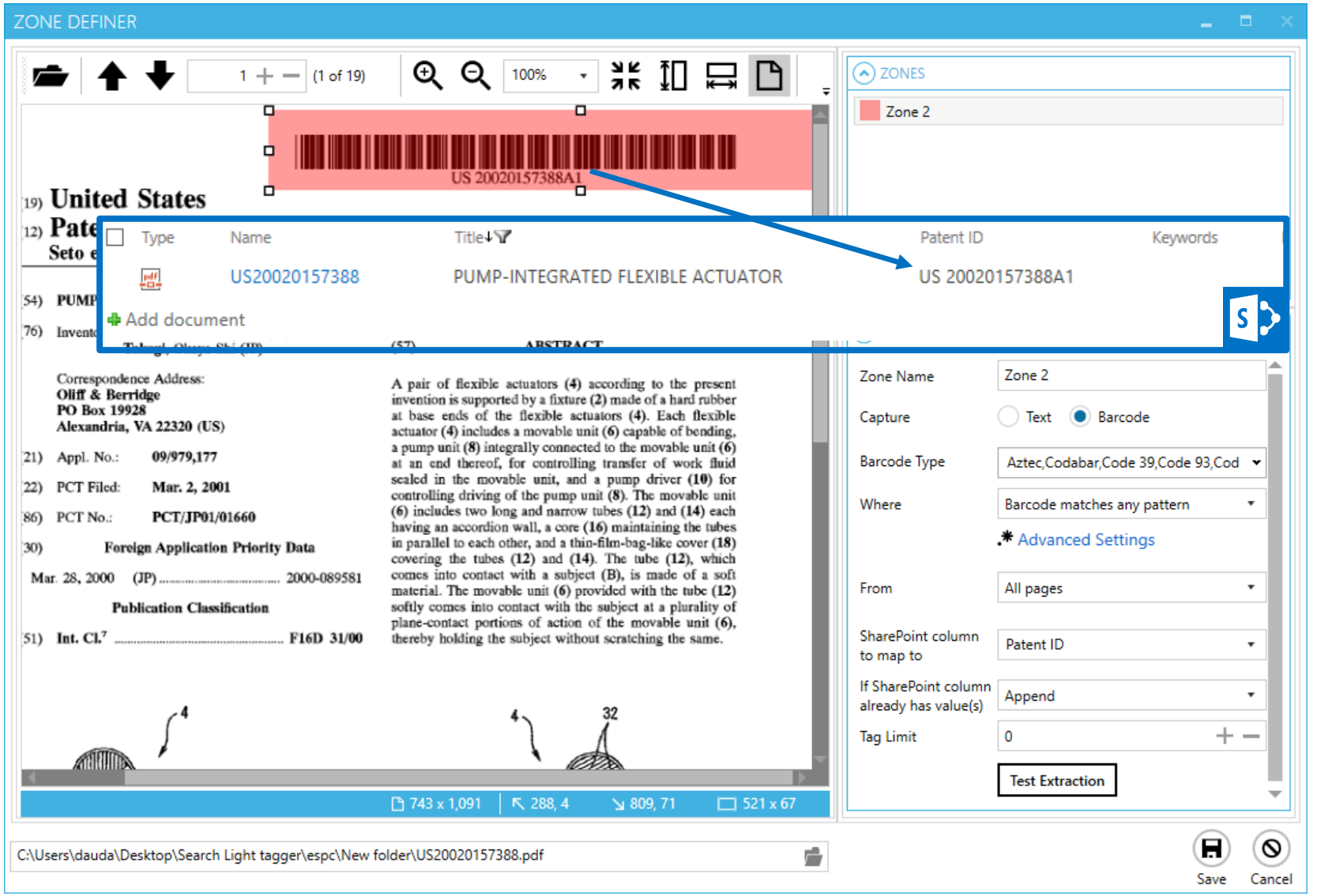
Searchlight will then extract entities from the text using NLP services such as: Open Calais, Rosette and Microsoft Cognitive Services.



Searchlight will then add the extracted entities as values to metadata columns in SharePoint.

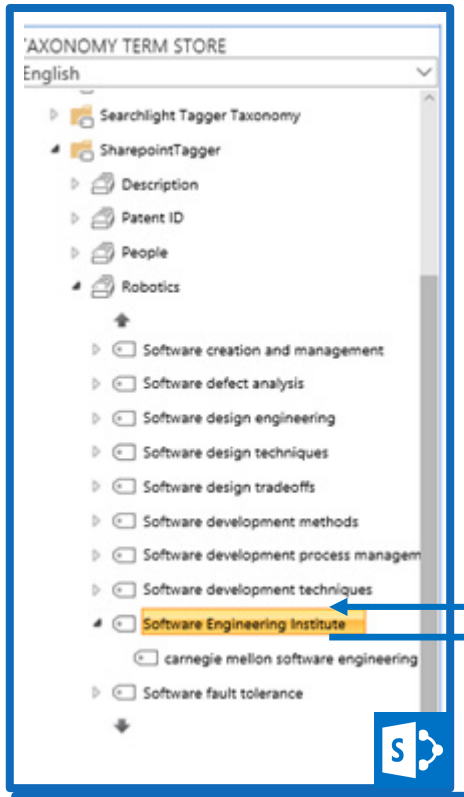
Aquaforest

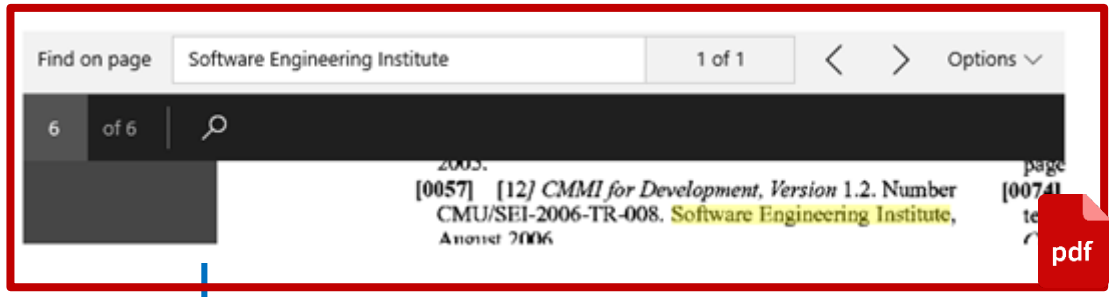
Barcode Values



Aquaforest

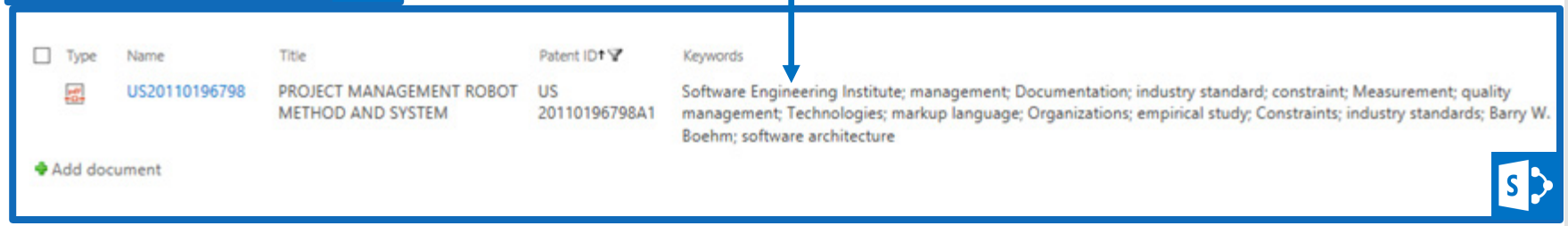
Taxonomy Matching





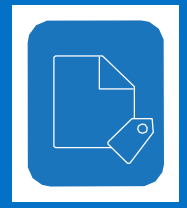
Searchlight matches the content of your documents to a term store in SharePoint to see if any terms appear in the document

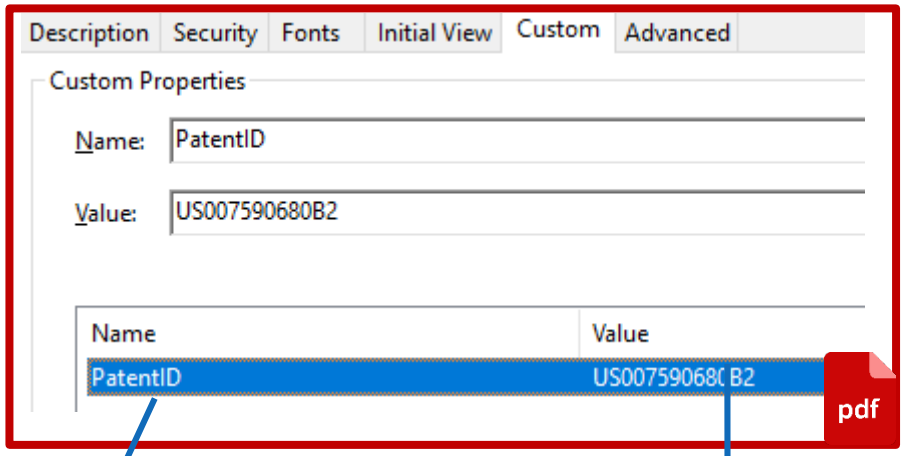
If any hits are found, Searchlight will then add them as values to the Managed Metadata columns



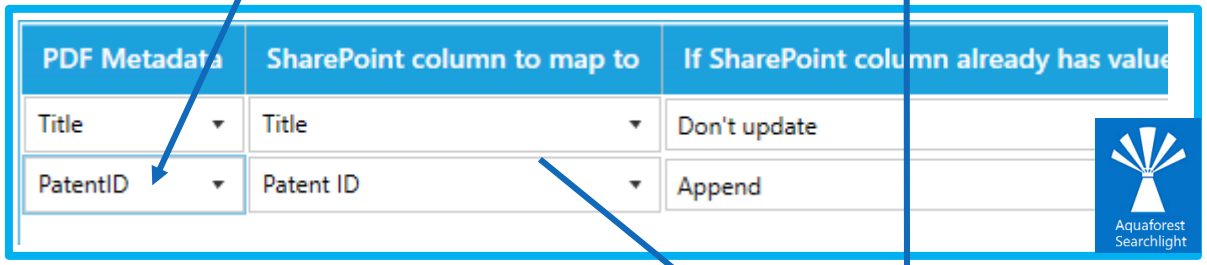
Aquaforest

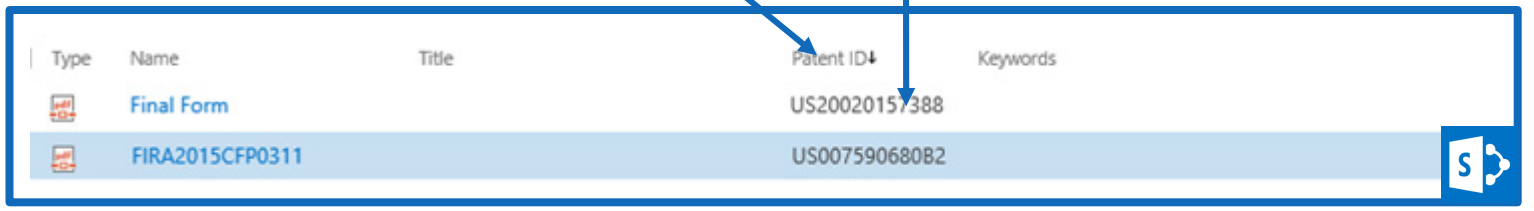
PDF Metadata Extraction





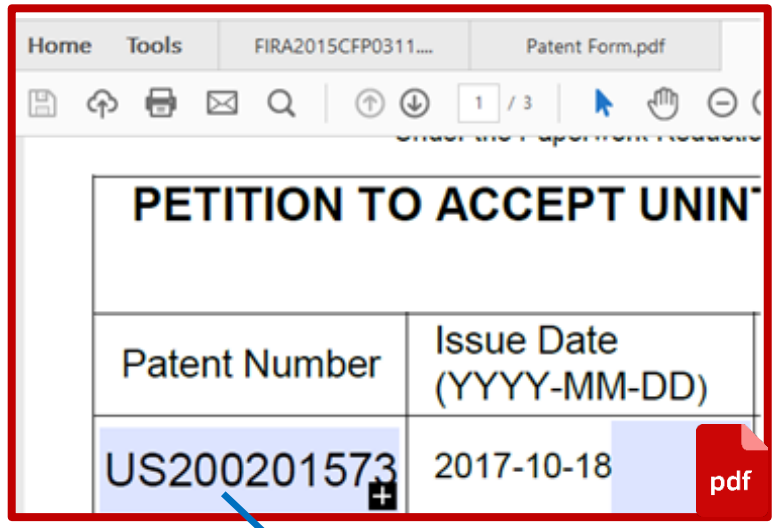
Searchlight can map PDF Metadata (Default/Custom) to SharePoint Metadata columns.



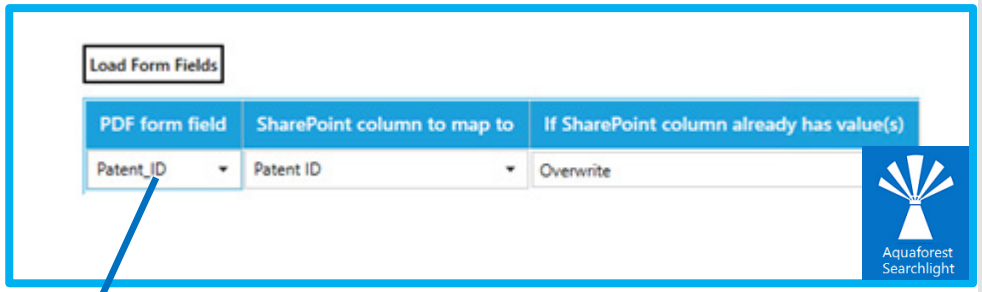


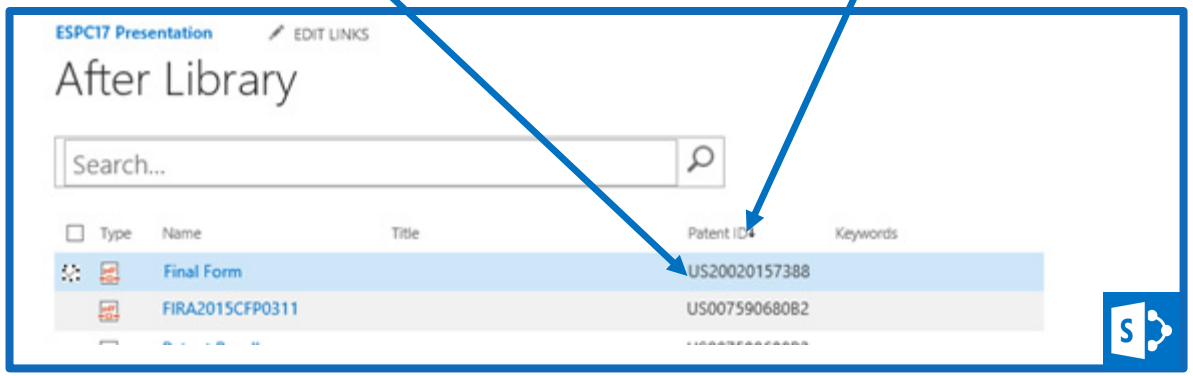
Aquaforest

PDF Forms



Searchlight can map PDF form fields to SharePoint Metadata columns.



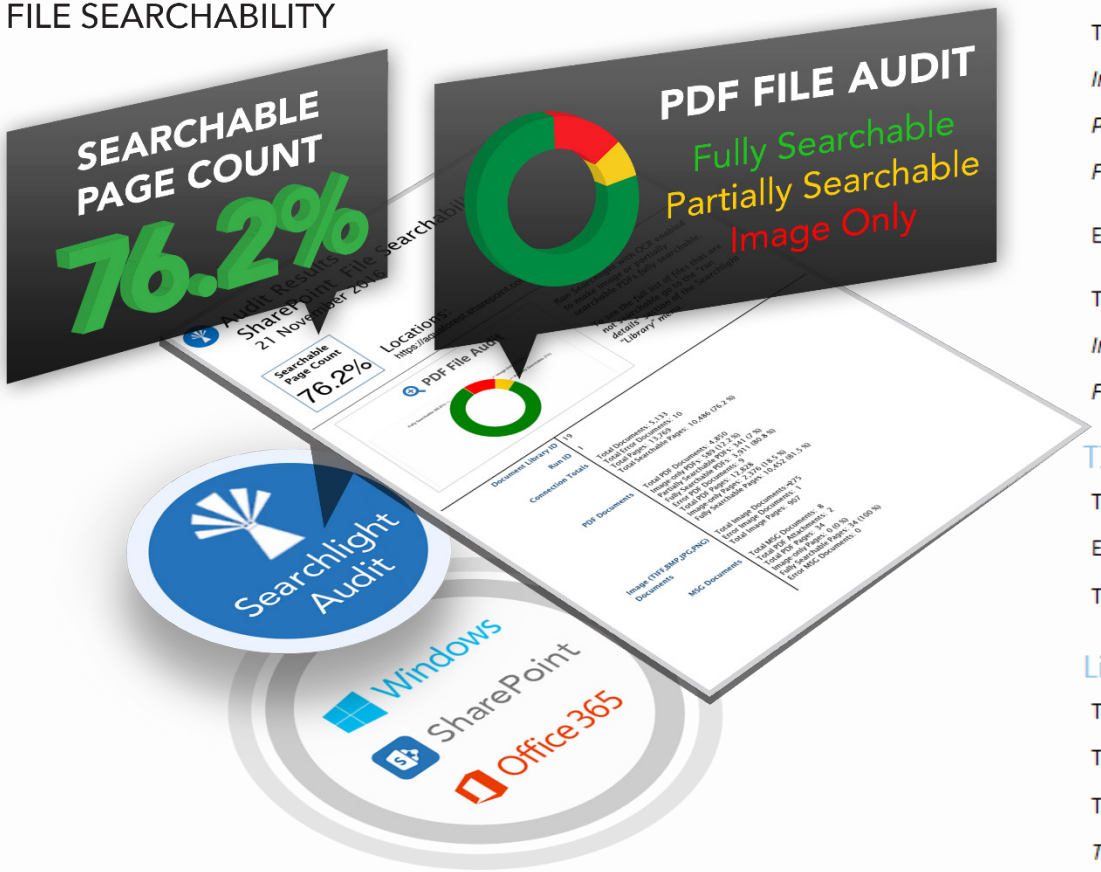


Aquaforest

Image PDFs

SHAREPOINT AUDIT RESULTS

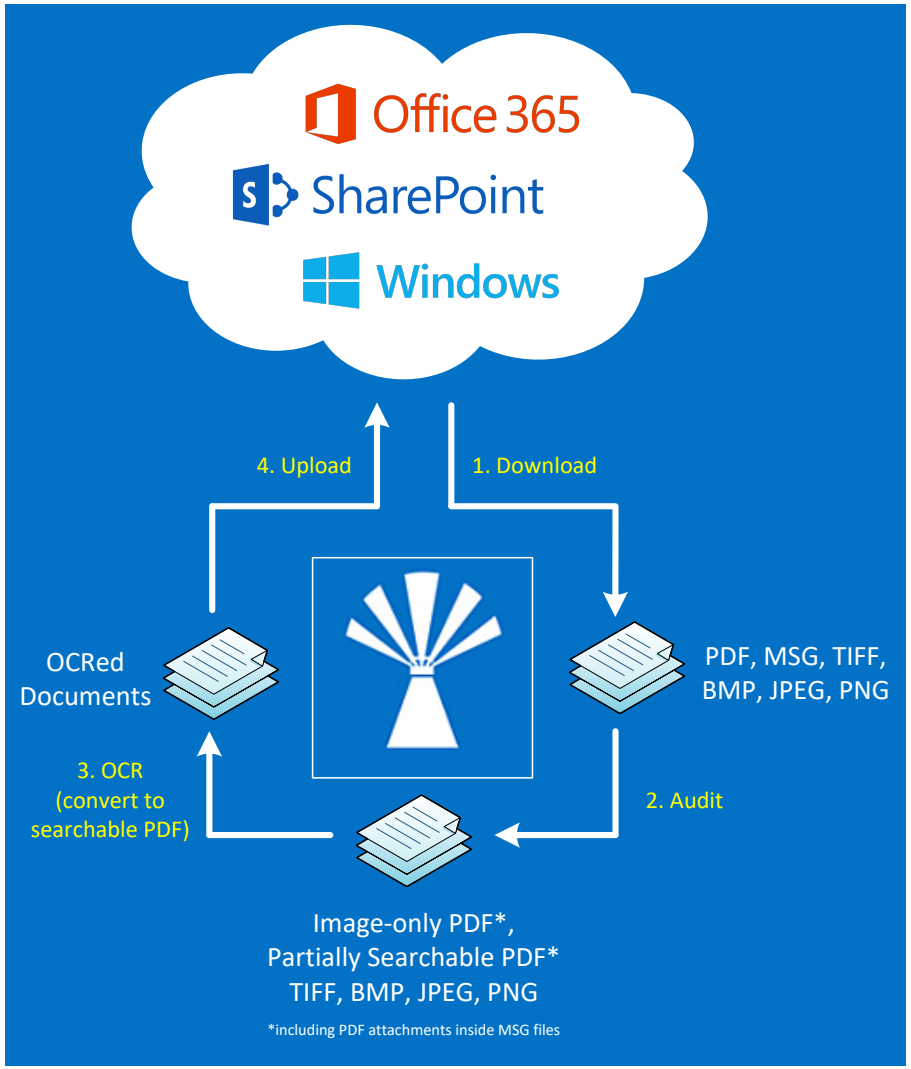
FILE SEARCHABILITY



STATISTICS	
PDF Documents	
Total PDF Documents:	2,421
Image-only PDFs:	*80 (3.3 %)*
Partially Searchable PDFs:	*11 (0.5 %)*
Fully Searchable PDFs:	*2,300 (96.2 %)*
Error PDF Documents:	30
Total PDF Pages:	18,027
Image-only Pages:	*3,327 (18.5 %)*
Fully Searchable Pages:	*14,700 (81.5 %)*
TIFF Documents	
Total TIFF Documents:	10,956
Error TIFF Documents:	83
Total TIFF Pages:	21,620
Library Totals	
Total Documents:	13,377
Total Error Documents:	113
Total Pages:	39,647
Total Searchable Pages:	*14,700 (37.1 %)*

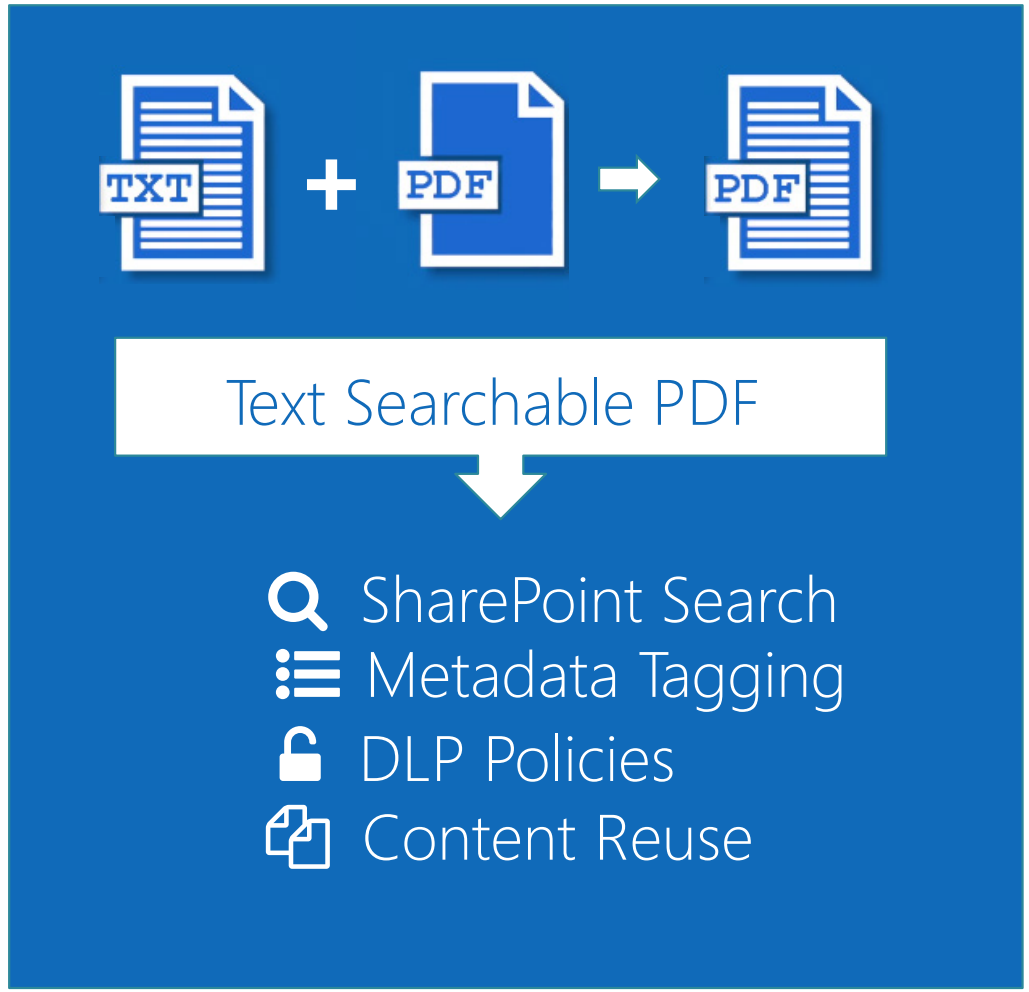
Aquaforest

Image PDF Processing



Aquaforest

Searchable PDF Benefits



Aquaforest

Ask us for More Information

info@aquaforest.com

Aquaforest