AvePoint

Set Your Records Straight

in Office 365

Records Management & Machine Learning:

What you need to know

28 July 2020 **11:00 – 12:00 BST**

In Partnership With:

DATA PROTECTION FORUM

irms

Reserve Your Spot

Meet the Presenters

Alyssa Blackburn

Director, Records & Information Management

AvePoint

@blackburnaj

Bruce Berends

Senior Solution Engineer

AvePoint

@bruce_berends

AvePoint

@avepoint

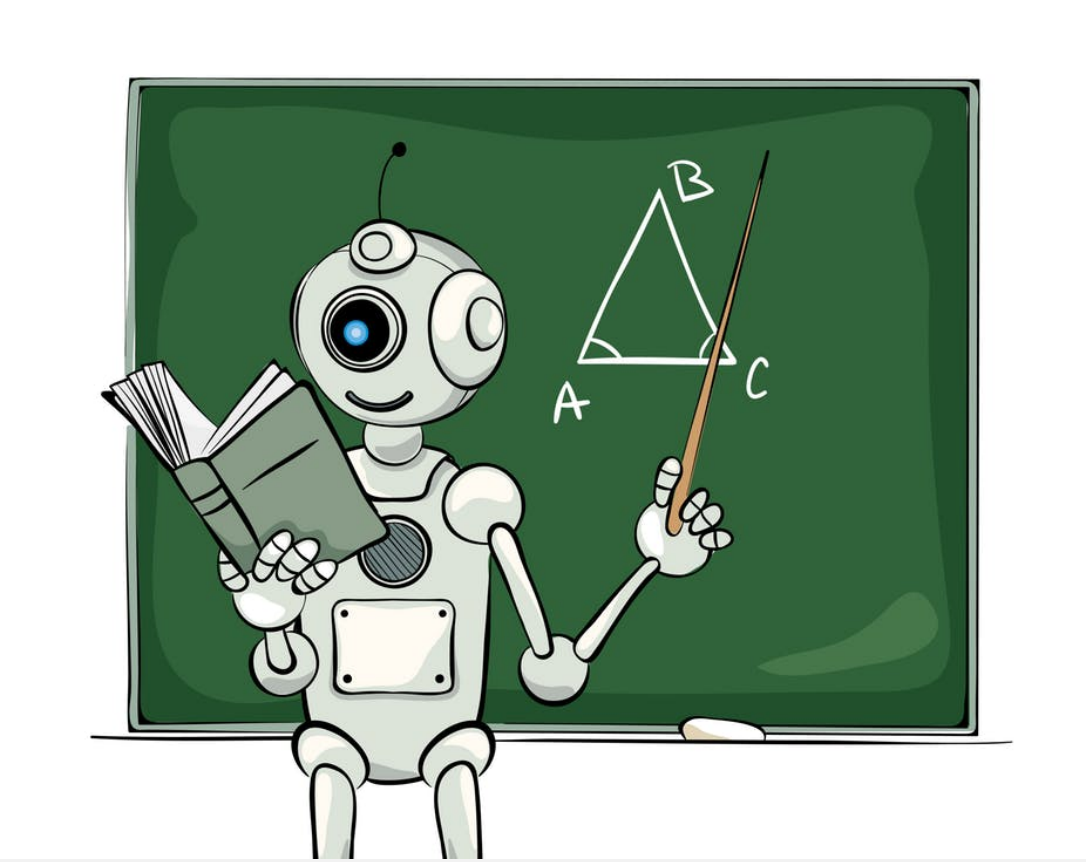
www.avepoint.com

What are we going to talk about?

- What is Machine Learning exactly?
- How do you use it for Records Management?
- What do you need before you start?
- Training the machine
- How does it all fit together?

What is Machine Learning exactly?

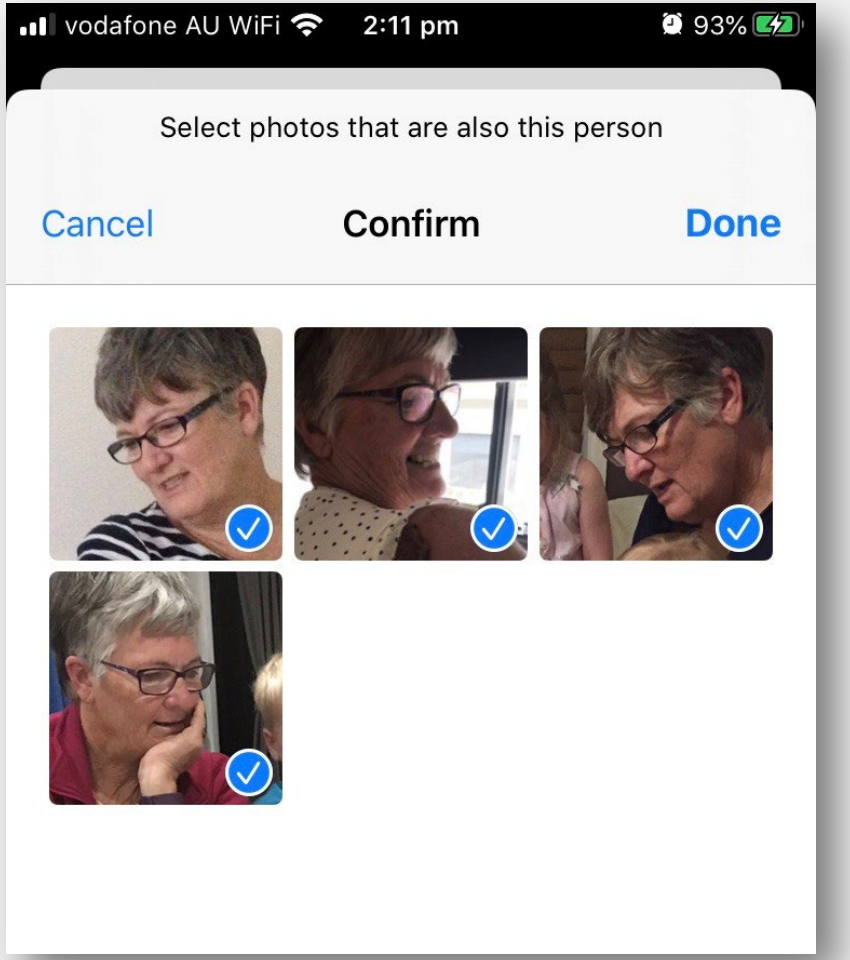
In its most simplistic form…

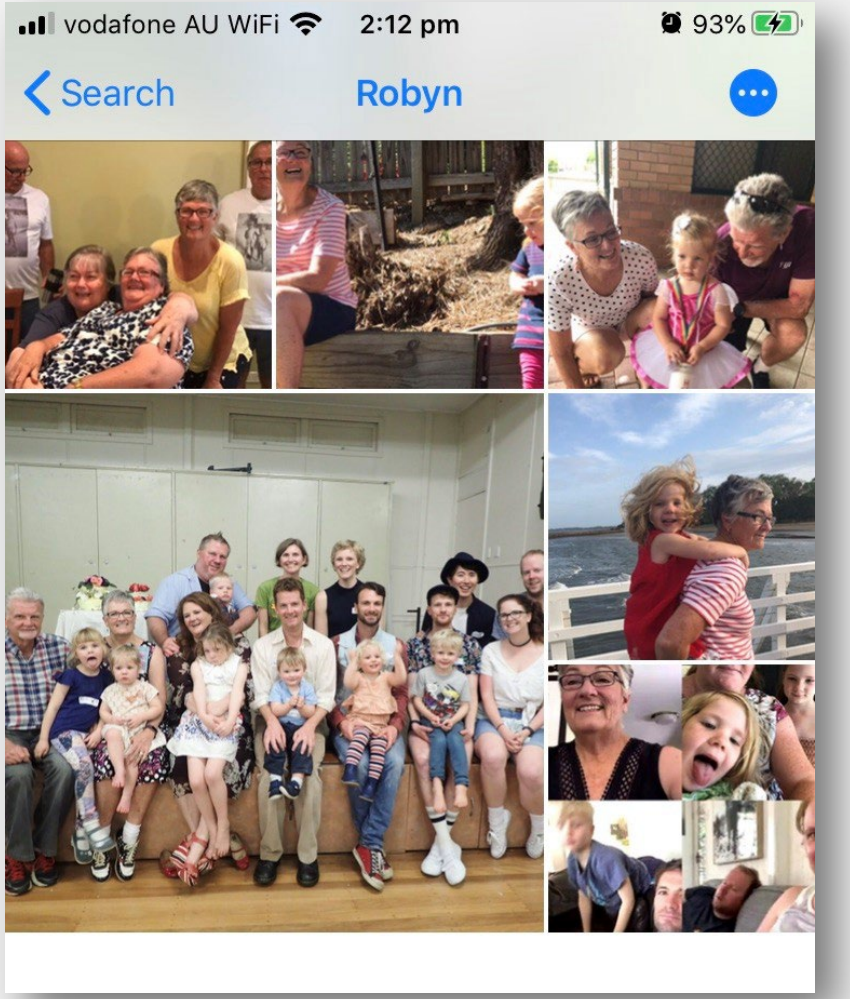


Labelling things using examples.

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Meet my Mum!





©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

MACHINE LEARNING
TECH
REALITY

What is it not?

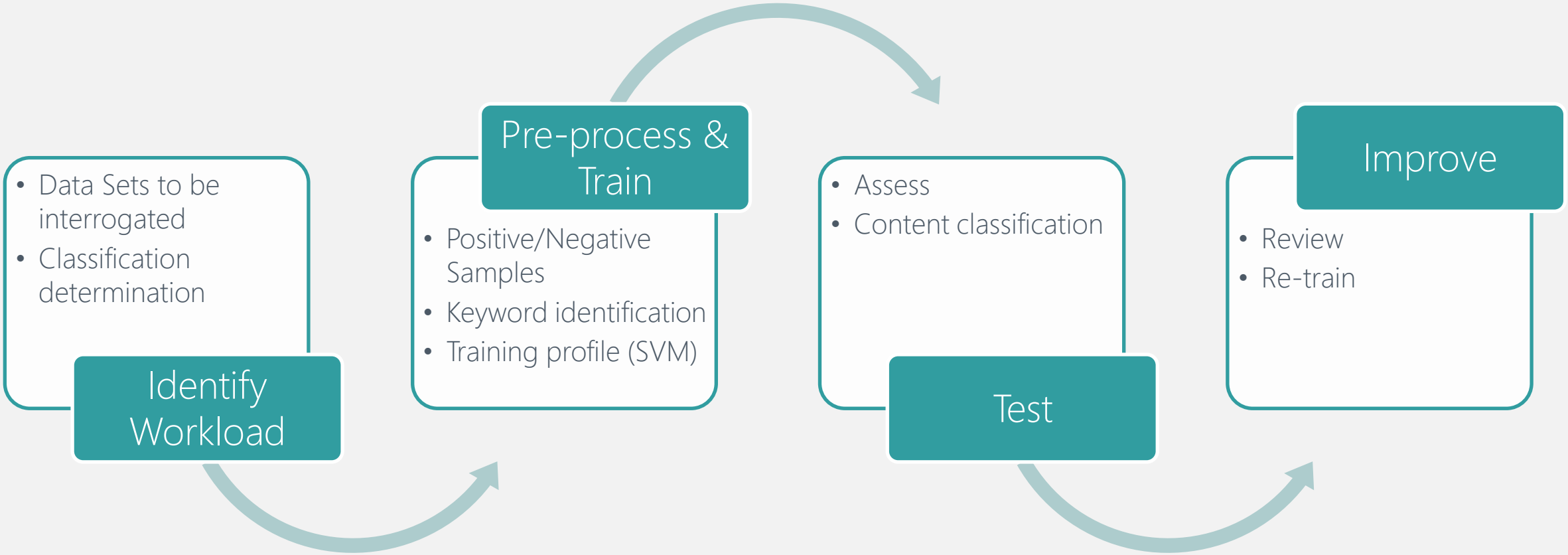
No matter how much we'd like it to be

Machine learning is not a silver bullet.

Like any technology, it's only going to be as good as the effort you put into it.

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Machine Learning Process



©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

How do you use it for Records Management?

Why would we consider Machine Learning?

- Too much content
- Can't rely on users
- Need to automate
- Focus on content that is easily repeatable

Machine learning is great for content that has consistent look and feel, identifiable patterns, high proportion of repeatability.

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

What do you need before you start?

What do you need before you start?

- **Scanning/Machine Learning Software/Solution** (AvePoint Compliance Guardian)
- **Positive samples** are the documents you are interested to find with Compliance Guardian
- **Negative samples** can be any kind of document used to train the model and distinguish between what you are looking for as document types and everything else.

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Ex-samples

- Positive

John Doe
128 First Avenue
Sydney, New South Wales 2251
Australia

Tax Invoice

Bill To Henry Higgins
1 Main St.
Melbourn,
Victoria 3035
Australia

Invoice no. 0000
Date 03/03/2015

Description	Quantity	Unit Price	Amount
Product name	1	$0.00	$0.00
		Subtotal	$0.00
		Balance Due	$0.00

Type your notes here

- Negative

StarShipIT

PACKING SLIP

SHIP TO: John Smith
Company Inc., Building A
30 George Street
Sydney
NSW, 2000
Australia

ORDER ID: S-TM27109
REFERENCE: PO-005101
ORDER DATE: 2019-Jun-11
Shipping Method: EXPRESS
Shipping Desc.: Express Shipping

INSTRUCTIONS: Please place each T-Shirt in a separate plastic bag, and stack them by size first then color.

SKU	DESCRIPTION	LOCATION	COLOR	SIZE	QTY
TS-SL-0129	Short Sleeve T-Shirt	RB12	Black	S	10
TS-SL-0139	Short Sleeve T-Shirt	RB13	Black	M	20
TS-SL-0174	Short Sleeve T-Shirt	RC9	White	M	30

Thank you for your purchase. If you are not satisfied with your purchase or the product is damaged please contact our support team at support@starship1.com

Order Number: S-TM27109

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Training the Machine

Configuring the Content Classification Tool

What is the Content Classification Tool useful for?

The machine learning check helps Compliance Guardian learn to classify files from unstructured data by using the Support Vector Machine (SVM) learning algorithm. It utilizes the Content Classification tool to perform the pre-processing and training.

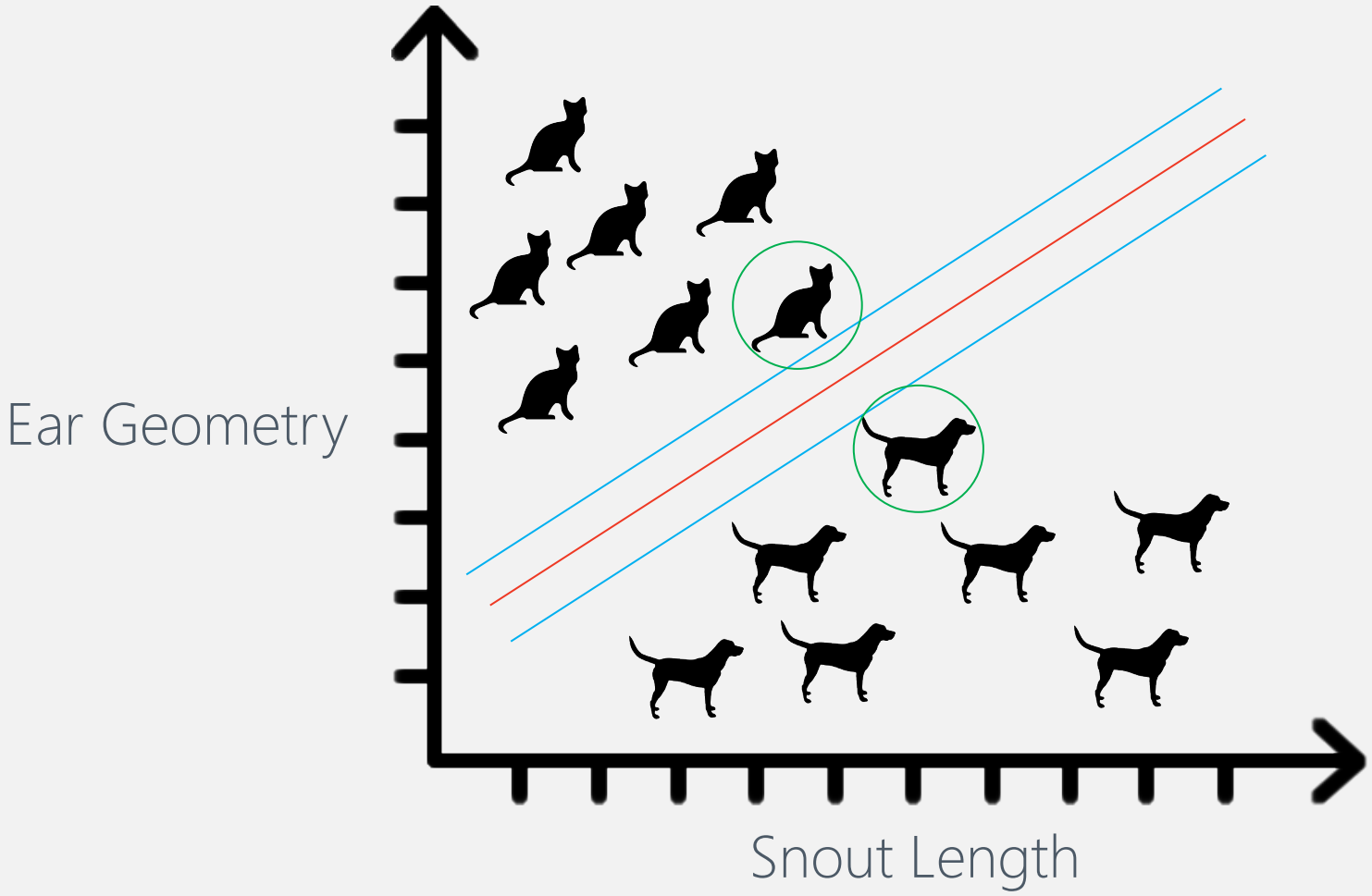
The SVM algorithm essentially takes 2 datasets of items and tries to find as many distinguishing features and differences as possible to determine support vectors – from that, the machine can determine which side of the line (hyperplane) that any item analyzed by the machine falls on and therefore classifies it accordingly.

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

How do we classify a dog that looks like a cat?

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Cats'n'Dogs - Support Vector Machine (SVM)



©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Let’s see it in action

How does this all fit together?

Linking machine learning and records management

Modernising RM for automation

This kind of technology cannot be successful unless we modernise the process we want it replicate.

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Let’s see it in action

Machine Learning Limitations

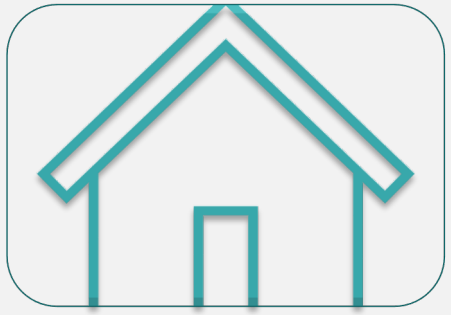
- It won't tell you if a document contains a Credit Card, SSN or Tax File Number. For example, it can tell if a document is a loan application, pay slip, medical form or anything you have it trained for.

- As every technology, it can produce false positives. This is indirectly as a result of a limited model or it requires better training input.

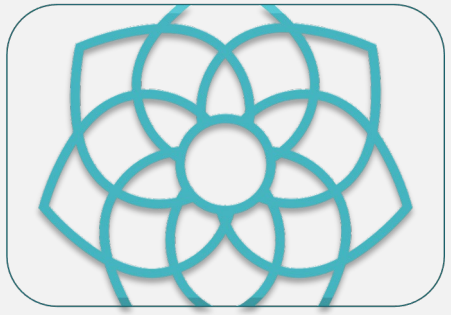
- Models may need to be re-trained as information changes.

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

Arsenal of Auto Classification



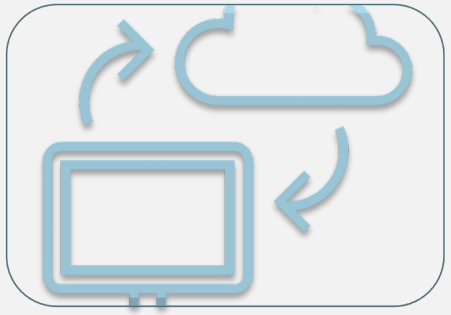
Location



Simple Metadata

Text Analysis

Document Fingerprinting



Machine Learning

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

3 thing to remember

Autoclassification is great!

It should always be linked to an outcome

It will only be as good as the effort you put in

©AvePoint, Inc. All rights reserved. Confidential and proprietary information of AvePoint, Inc.

QUESTIONS

What to do next….

Looking to improve your skills?

Microsoft Office 365 & SharePoint Records Management Training. All you need to know

Are you records ready for Office 365?

Office 365 readiness roadmap Assessment for records management

(led by Alyssa Blackburn)

©Aver oint, Inc. All rights reserved. Confidential and proprietary information of Aver oint, Inc.

thank you

Gracias	ευχαριστώ	Danke	Grazie	благодаря
Hvala	Obrigado	Kiitos	شكراً	Tak
Ahsante	Teşekkürler	متشكرم	Salamat Po	감사합니다
Cám ơn	شكريه	Terima Kasih	Dank u Wel	Děkuji
நன்றி	Köszönöm	ありがとうございます	ขอบคุณครับ	Dziękuję
谢谢	Tack	Mulţumesc	спасибо	Merci
תודה	多謝晒	дякую	Ďakujem	