About me-

- Trevor Grant
- PMC Apache Mahout/Apache Streams
- Open Source Evangelist and AI Engineer @IBM

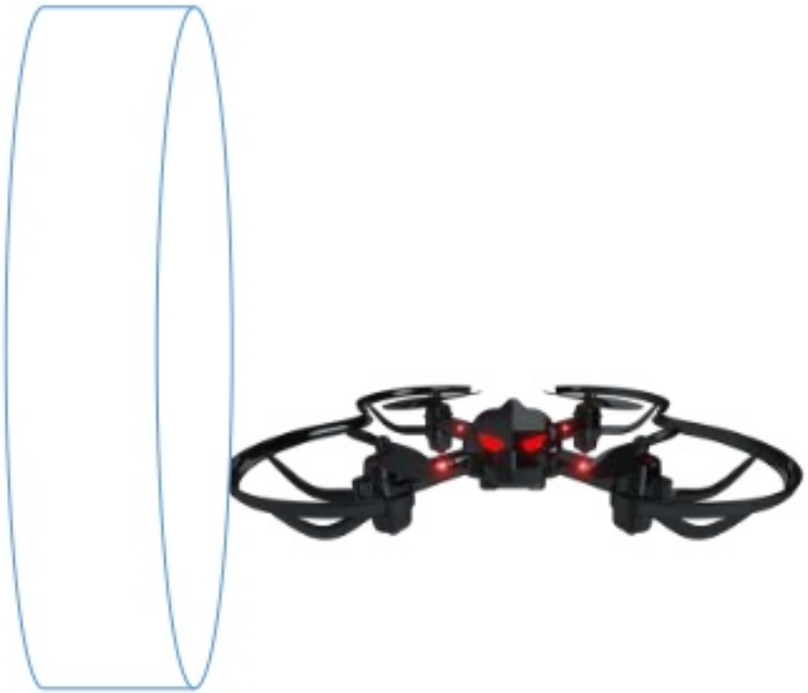
Twitter: @rawkintrevo

Simple Cylon

- OpenCV for Facial Detection
- Mahout / Eigenfaces for Facial Recognition
- Problems

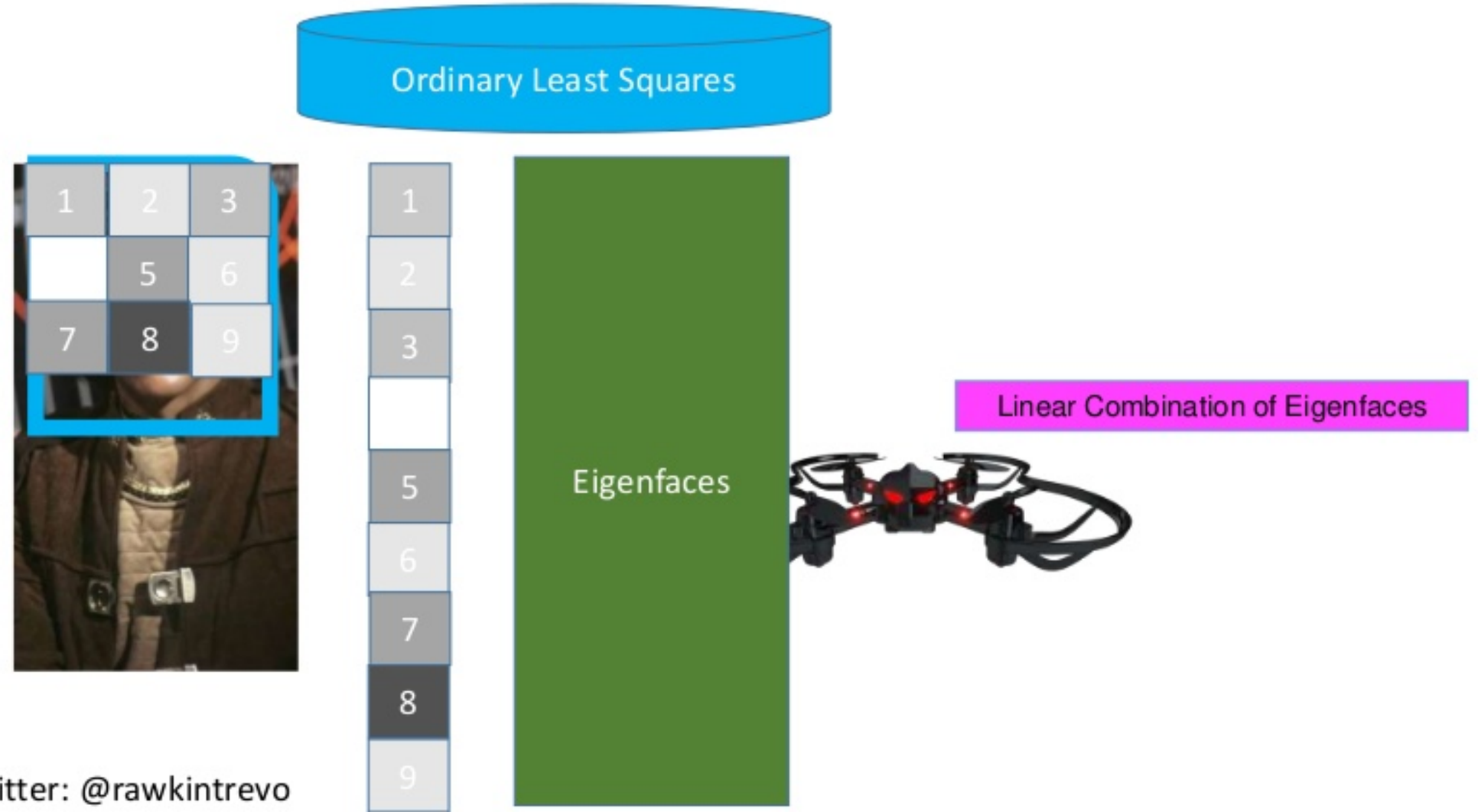
Image Credit: Electric Alivia
https://i.pinimg.com/736x/6a/be/0d/6abe0da5907a5f4923b0c94389a02ddc--pretty-tattoos-pin-up-tattoos.jpg

Open CV *Detects* Faces in Video Frame



Twitter: @rawkintrevo

Mahout Decompose FaceRect into Linear Comb Vec



Twitter: @rawkintrevo

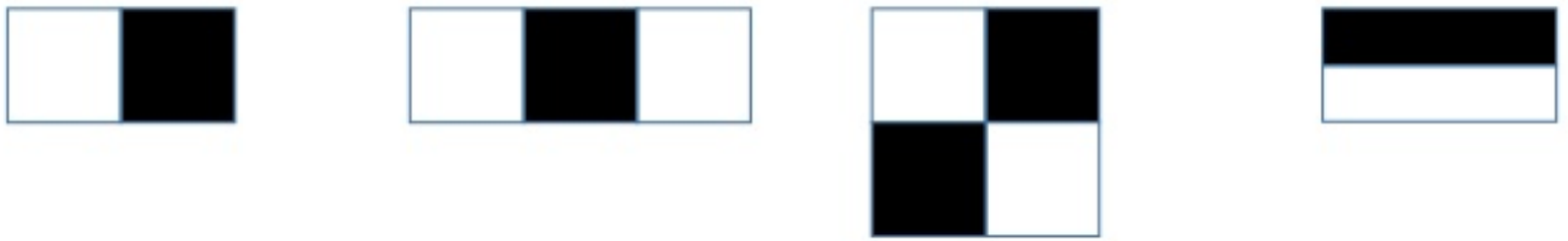
Search k-Length Vector
(k is number of Eigenfaces ~ 130)

Solr Returns Closest Match
Drone says “hi” or shoots to kill

Twitter: @rawkintrevo

Cascade Filter Overview

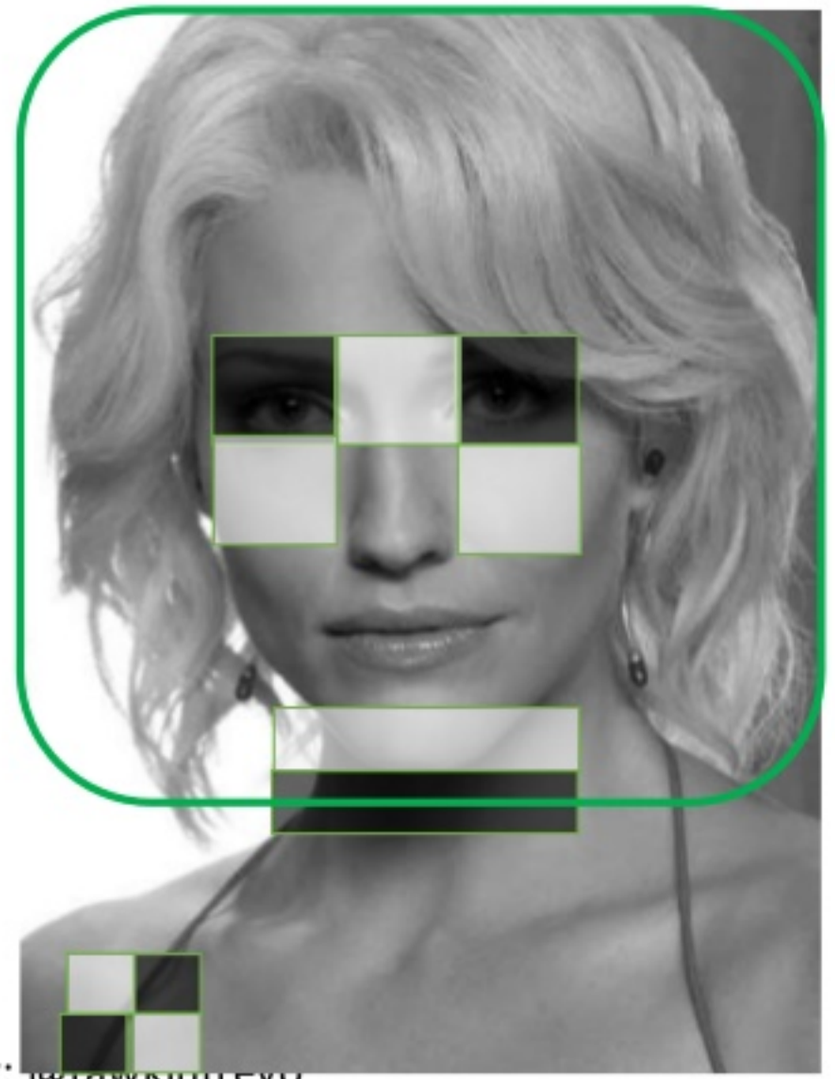
- Not a Neural Net (NNets are too slow)
- OpenCV -> Haar Cascades are much more efficient
- Scans for areas that match certain patterns.



Cascade Filter Overview

Cascade Filter

Cascade Filter (Areawise)

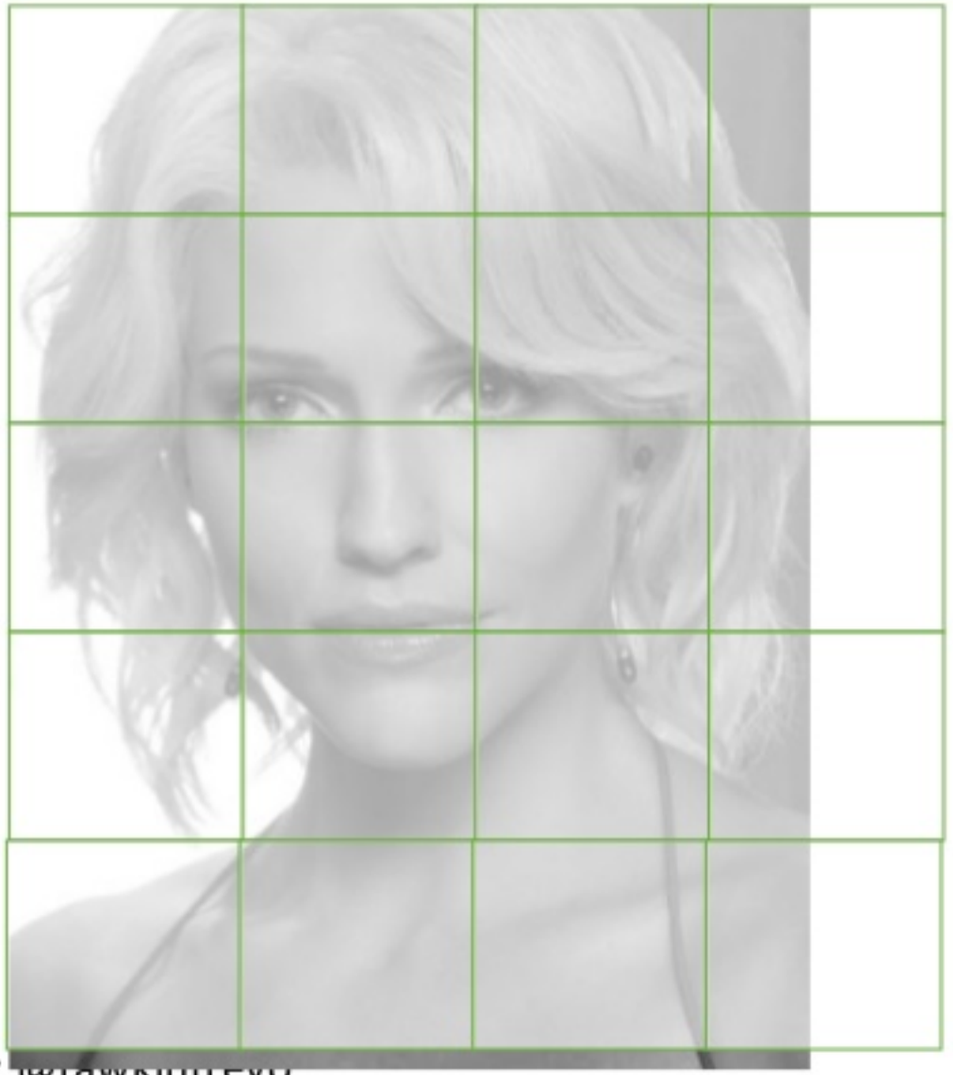


OpenCV

Eigenfaces

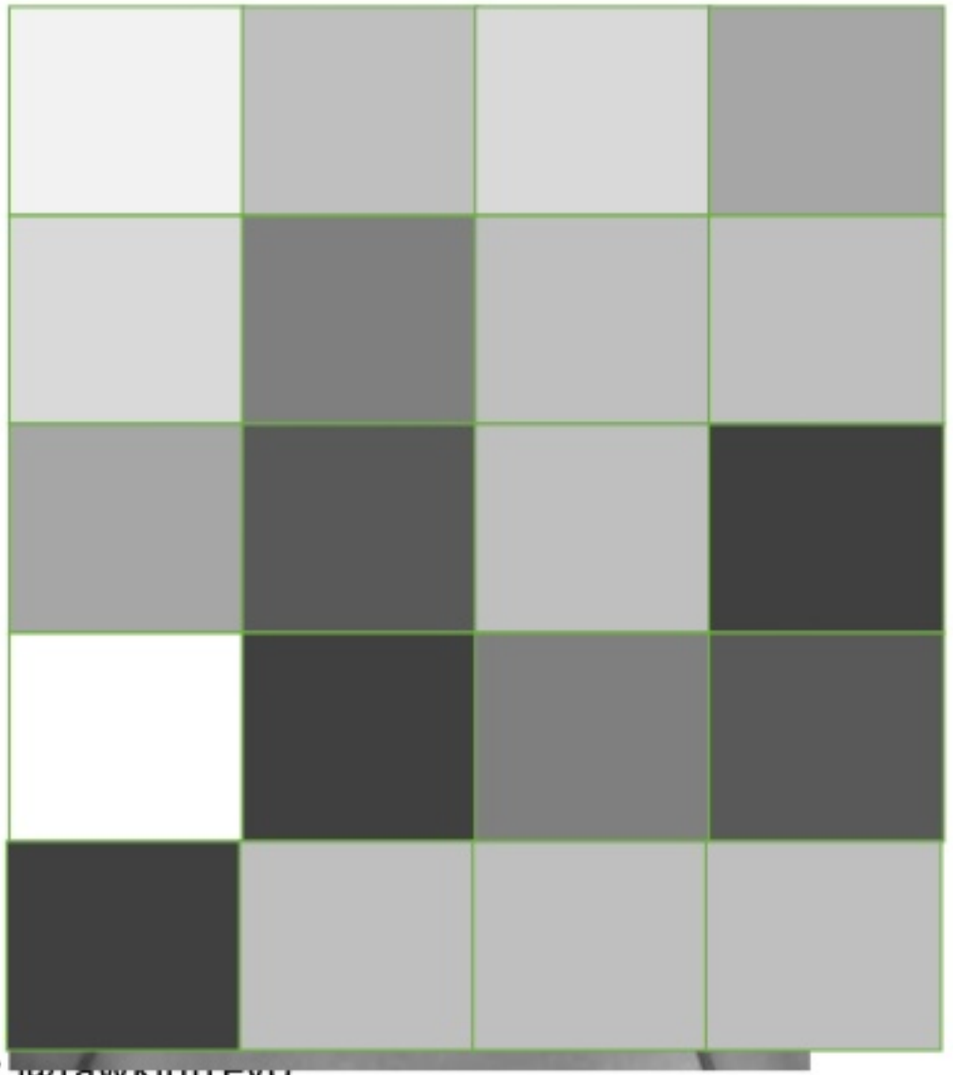
Eigenfaces

Eigenfaces (Pixelwise)

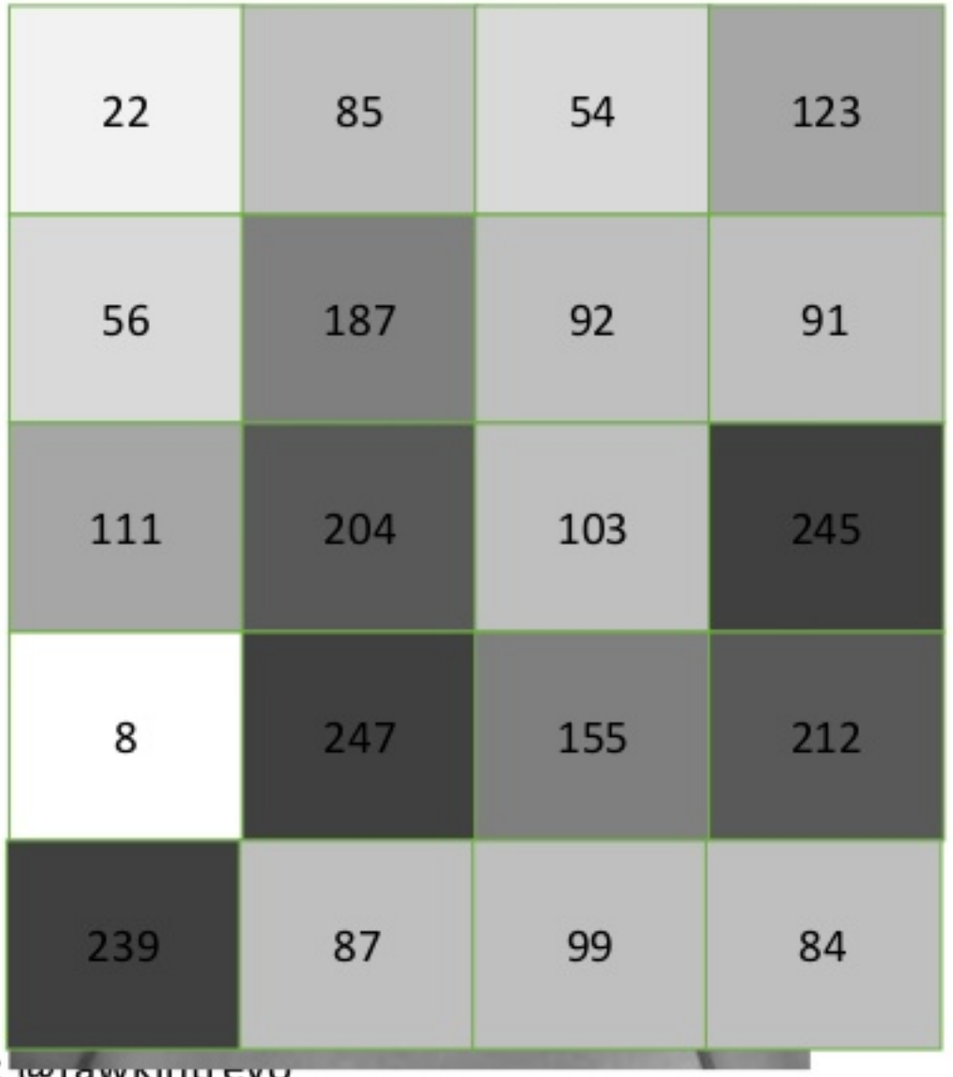


Eigenfaces (Pixelwise)

Eigenfaces (Pixelwise)



Eigenfaces (Pixelwise)



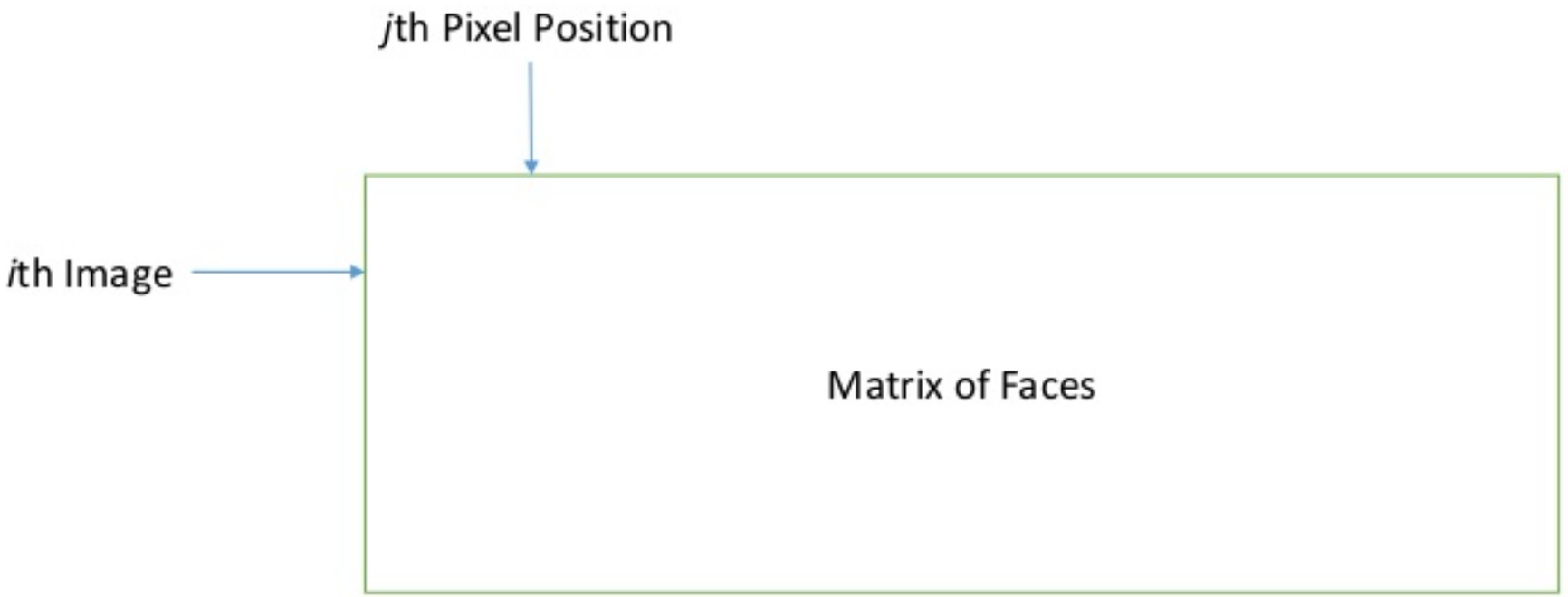
Eigenfaces (Pixelwise)

22	85	54	123
56	187	92	91
111	204	103	245
8	247	155	212
239	87	99	84

Eigenfaces

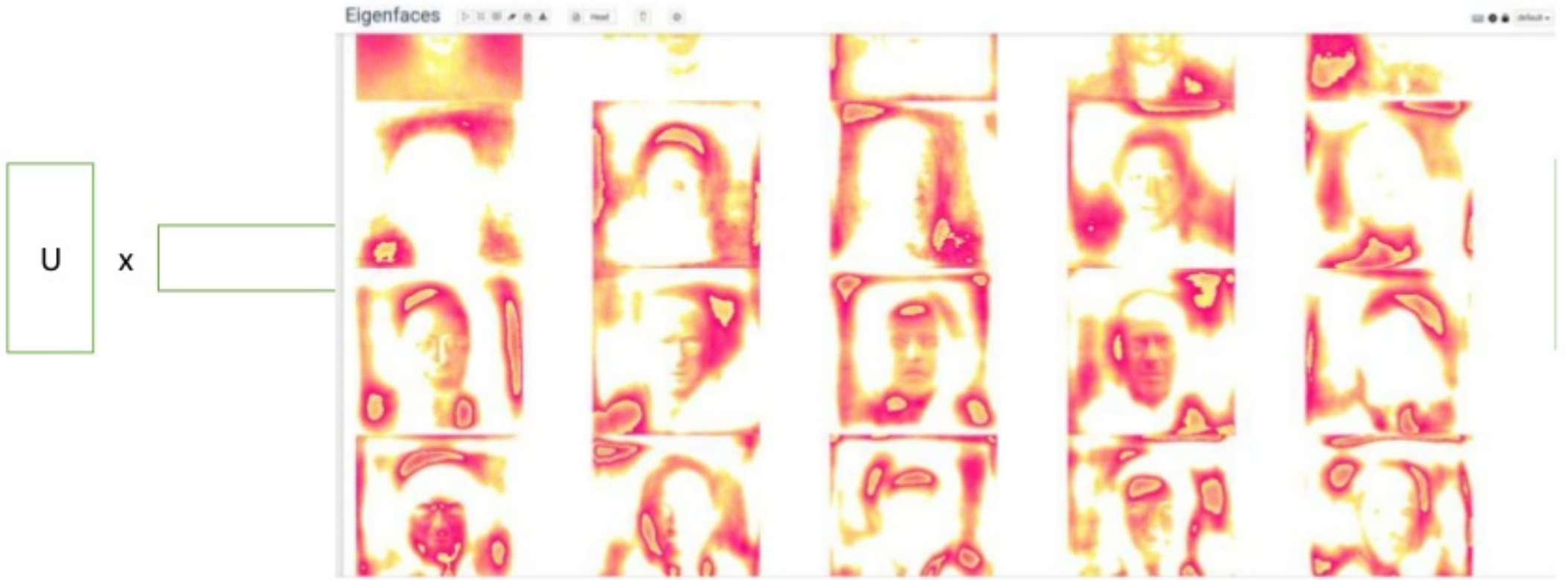
Eigenfaces

Eigenfaces



Eigenfaces: Singular Value Decomposition

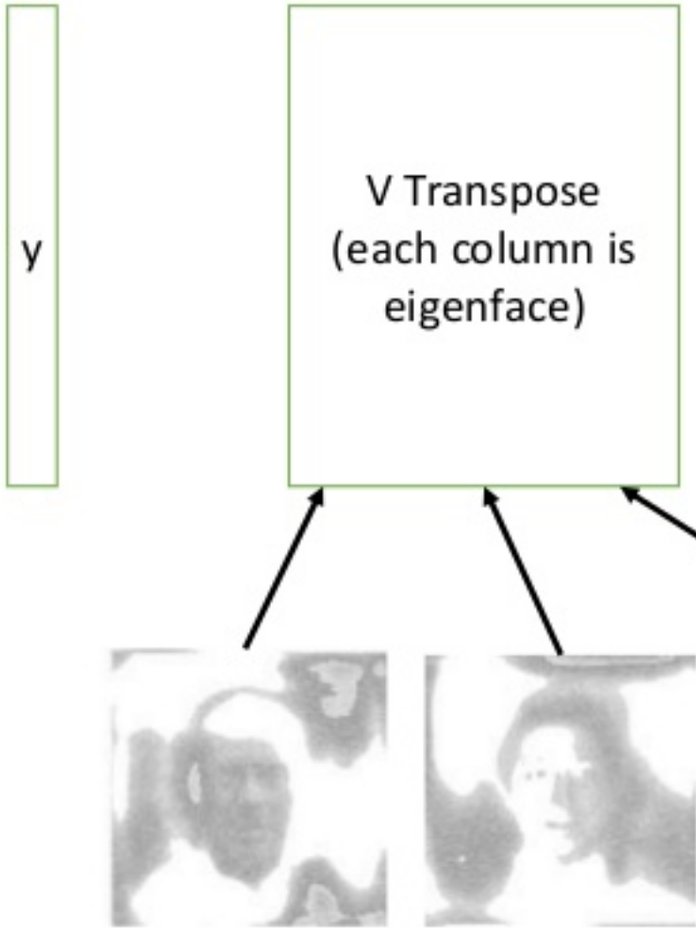
Eigenfaces: Matrix V



Eigenfaces: Matrix U

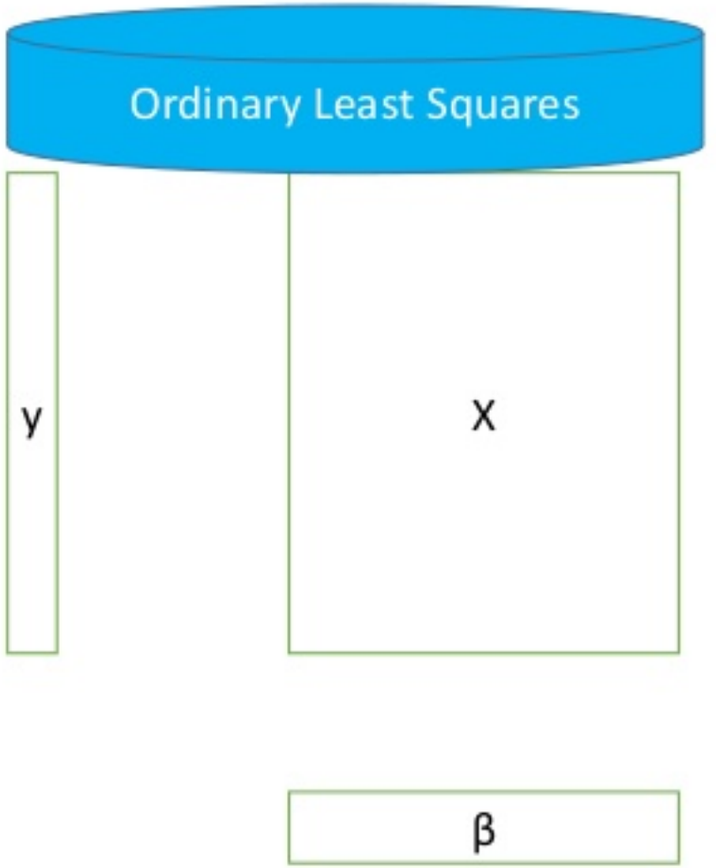
Linear combinations of Eigenfaces required to form the Nth Face

New Faces



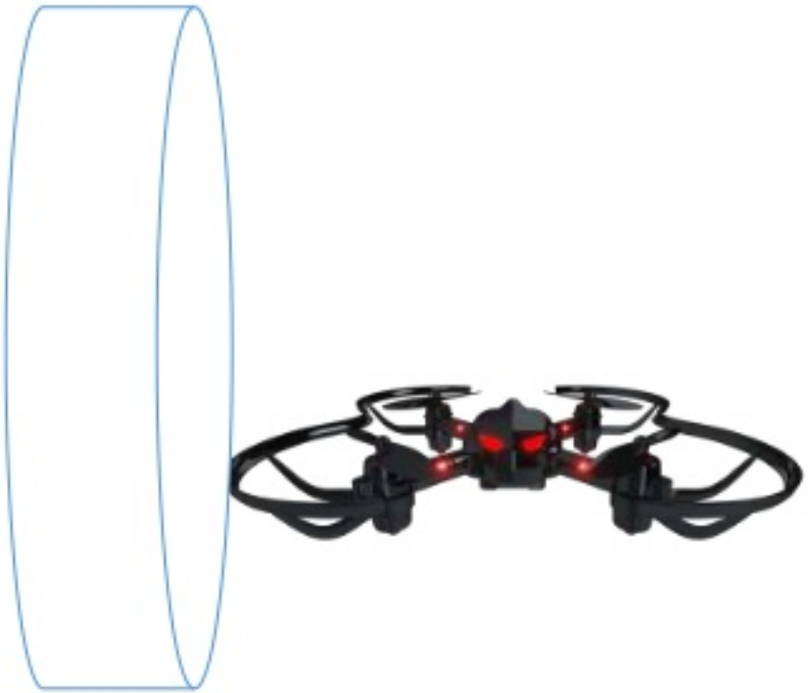
Twitter: @rawkintrevo

New Faces



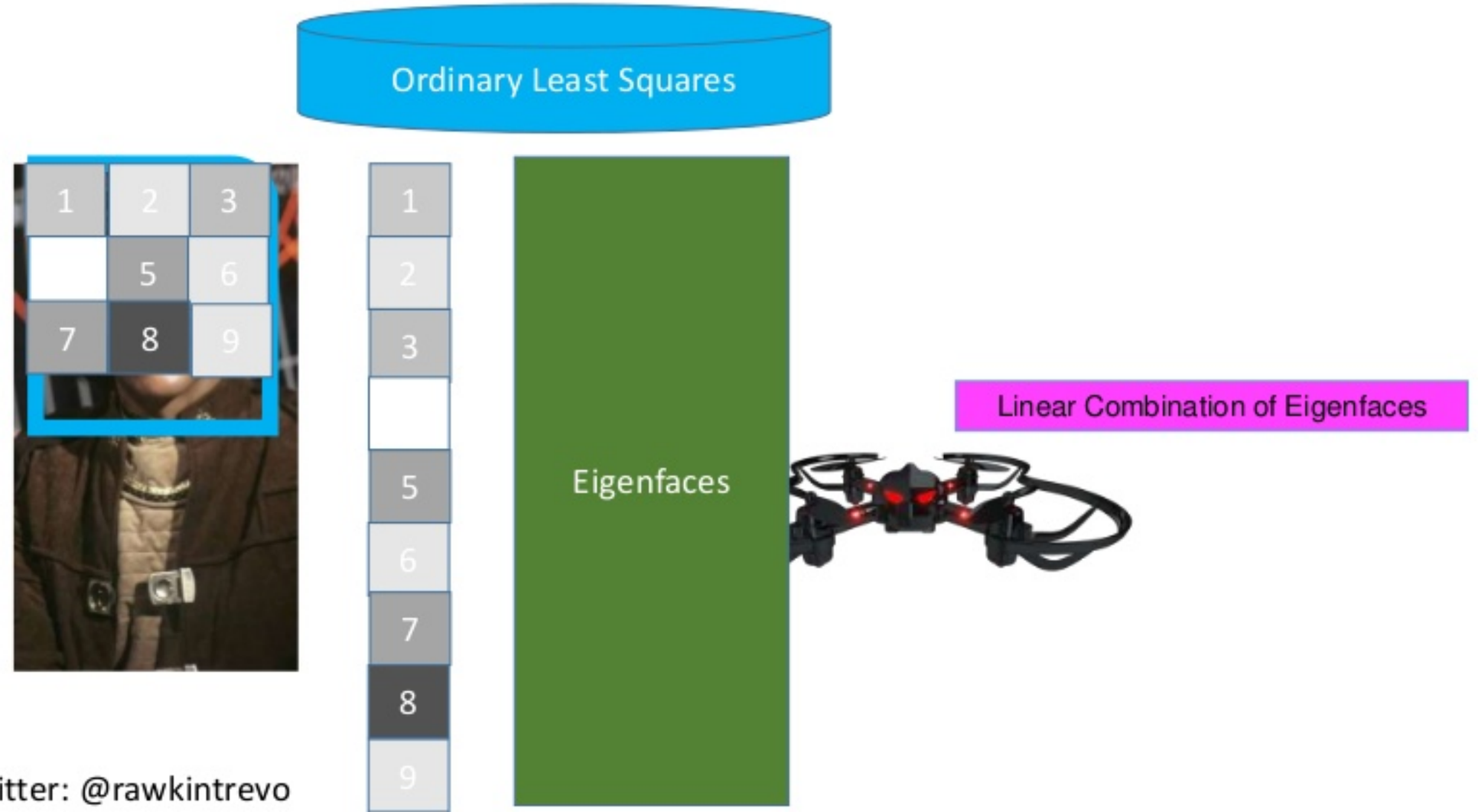
Simple Regression (OLS)

Open CV Detects Faces in Video Frame



Twitter: @rawkintrevo

Mahout Decompose FaceRect into Linear Comb Vec



Twitter: @rawkintrevo

Search k-Length Vector
(k is number of Eigenfaces ~ 130)

Solr Returns Closest Match
Drone says “hi” or shoots to kill

Twitter: @rawkintrevo

Problems with Simple Cylon

- OpenCV/Haar Cascades inconsistently boxes face
  - (Eigenfaces requires images be centered and same size as training)
- Adds face every time angle changes (ideally we want multiple combinations assosciated with each person)
- Will insert “bad” faces, and “not” faces.
- To much ”radio chatter” with Kafka (esp with “who else is around” context clues)

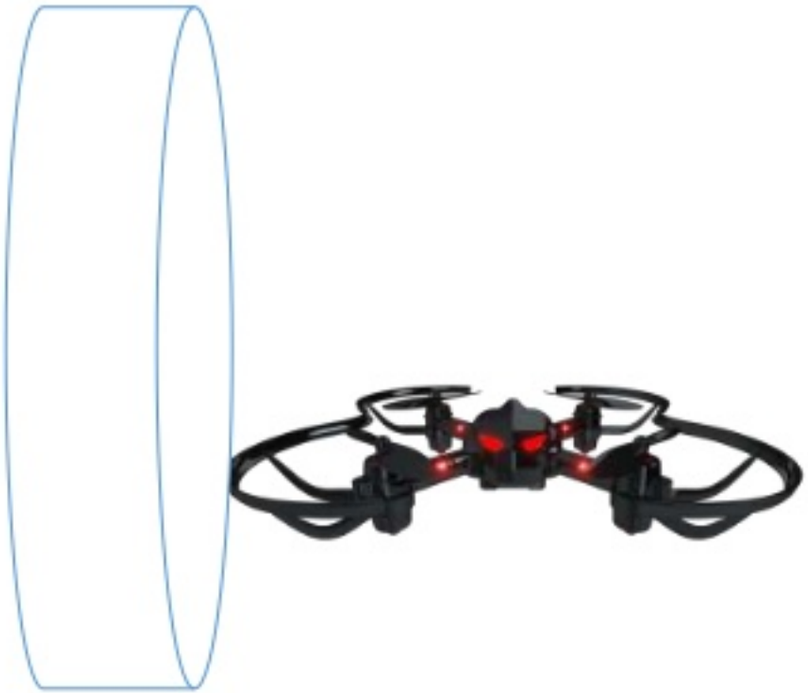
Open CV Detects Faces in Video Frame

Twitter: @rawkintrevo

Open CV Detects Faces in Video Frame

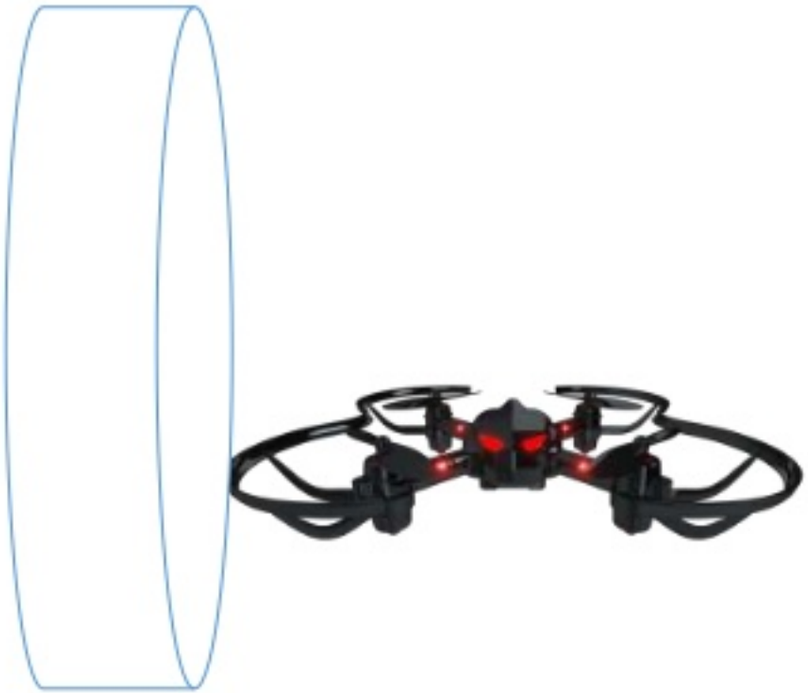
Twitter: @rawkintrevo

Open CV Detects Faces in Video Frame



Twitter: @rawkintrevo

Open CV Detects Faces in Video Frame



Twitter: @rawkintrevo

Open CV Detects Faces in Video Frame

Twitter: @rawkintrevo

OpenCV/ Haar Cascade Filter inconsistent

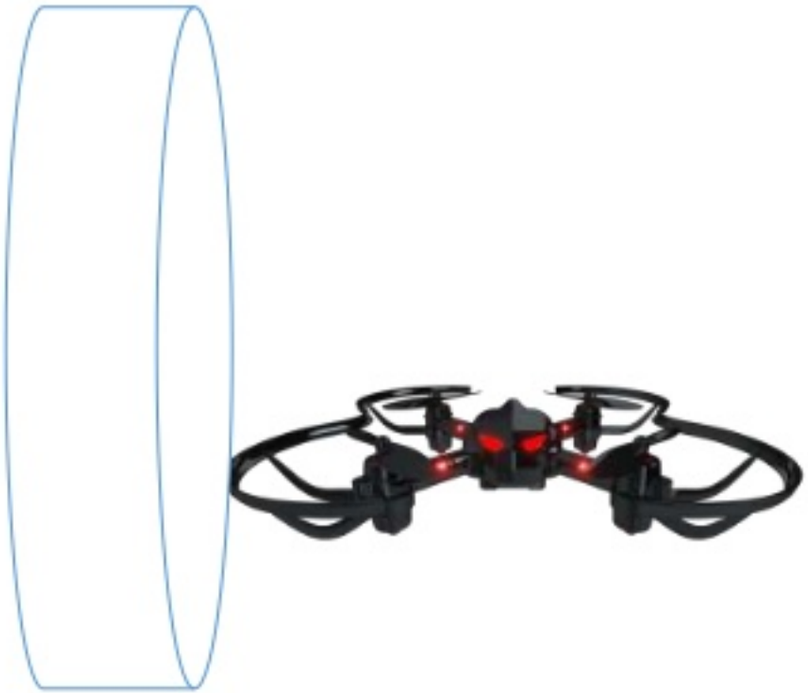
Twitter: @rawkintrevo

Canopy Clustering

- Create N Second Window
- Cluster Faces in Window
- Quick dirty clustering- but effective.
  - First point is “center”
  - All points within distance *t2* are “in that cluster.
  - If a point is not within *t2* of any cluster- it becomes a new cluster center.

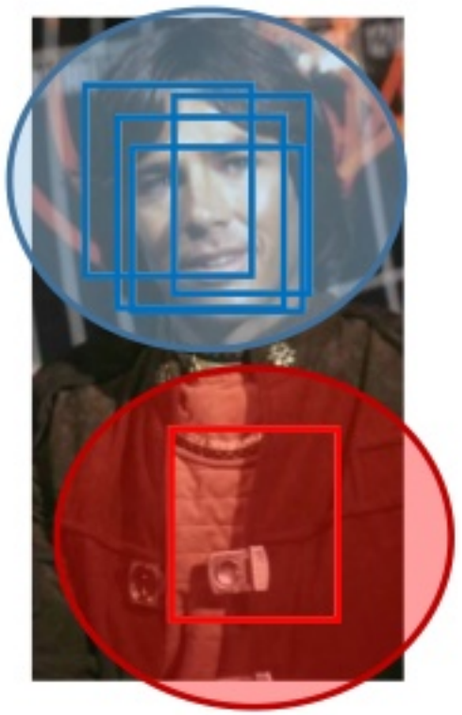
Open CV Detects Faces in Video Frame

t2= max square width



Twitter: @rawkintrevo

Canopy Clustering To Remove “Ghost” Faces



First rect – new cluster

t2= max square width

Second Rect- within one width of first rect (same cluster)

Third Rect- within one width of first rect (same cluster)

Forth Rect- **NOT** within one width of first rect (new cluster)

Fifth Rect- within one width of first rect (same cluster)

Finally- any cluster with less than two entities in windows gets filtered out.

Twitter: @rawkintrevo

Canopy Clustering To Remove “Ghost” Faces

Twitter: @rawkintrevo

Remaining Faces (Stupid Way)

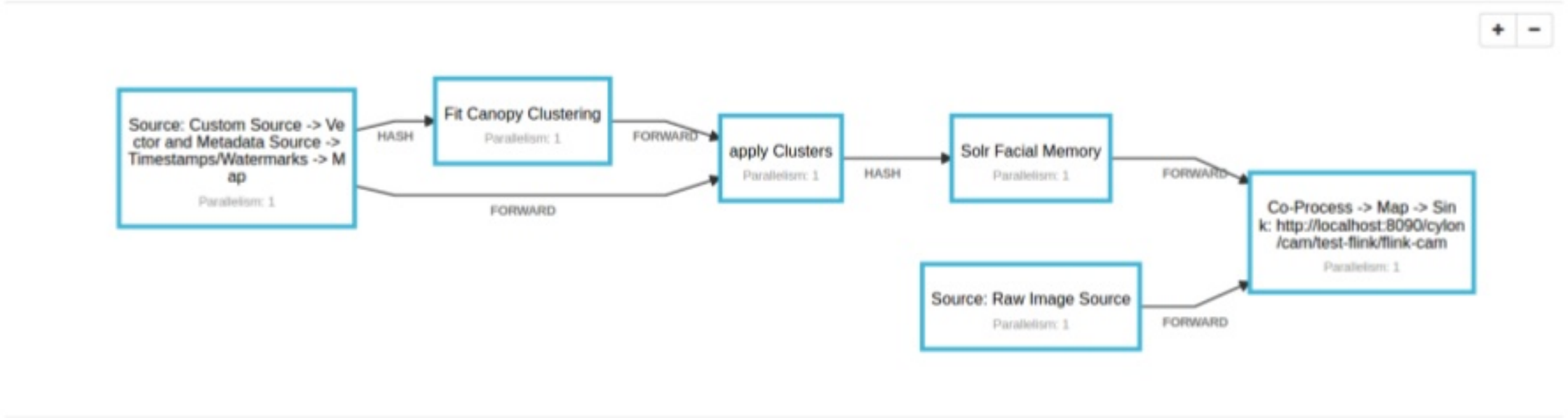
Remaining Faces (Smarter Way)

All faces from a given cluster are likely from the same entity

Flink Enabled

- Windowing lets us use “Intelligence”
  - Removing Ghost Faces
  - Using Context Clues (Person in one part of frame doesn’t magically become someone else)

Flink Plan


Source: Custom Source -> Vector and Metadata Source -> Timestamps/Watermarks -> Map
Parallelism: 1
HASH
Fit Canopy Clustering
Parallelism: 1
FORWARD
FORWARD
apply Clusters
Parallelism: 1
HASH
Solr Facial Memory
Parallelism: 1
FORWARD
Source: Raw Image Source
Parallelism: 1
FORWARD
Co-Process -> Map -> Sink: http://localhost:8090/cylon/cam/test-flink/flink-cam
Parallelism: 1

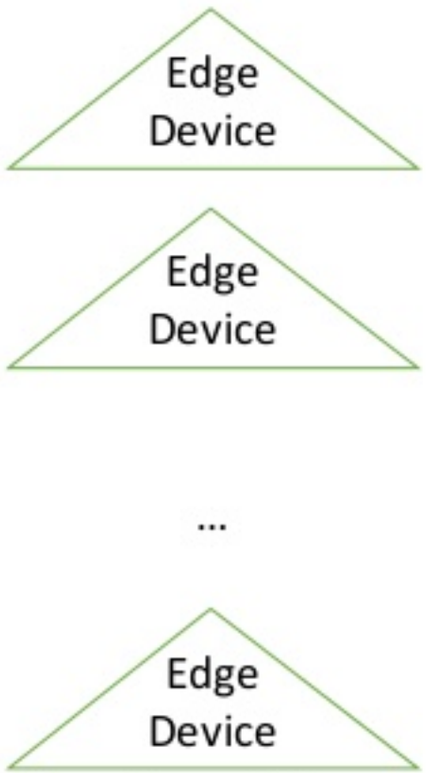
Definition of Intelligence

- Merriam Webster:
  - a (1) : the **ability to learn or understand** or to deal with new or trying situations : reason; also : the skilled use of reason
  - (2) : the ability **to apply knowledge to manipulate one's environment** or to think abstractly as measured by objective criteria (such as tests)
- Oxford
  - The ability to **acquire and apply knowledge and skills**.
- Contested subject among researches…

Artificial Intelligence

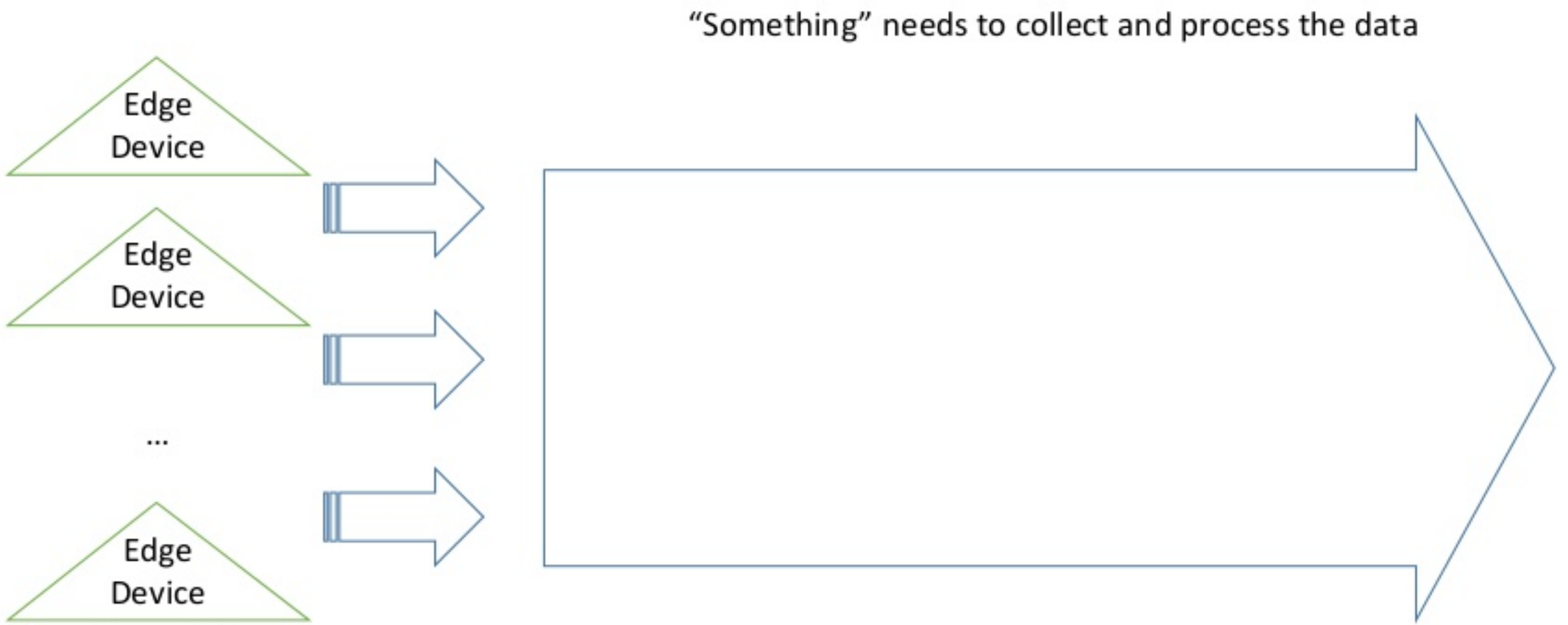
- Collect information from environment
- Adapt to changing environment
- Learn and understand / gain knowledge / improve skills
- Apply knowledge

IoT AI - A Diagram



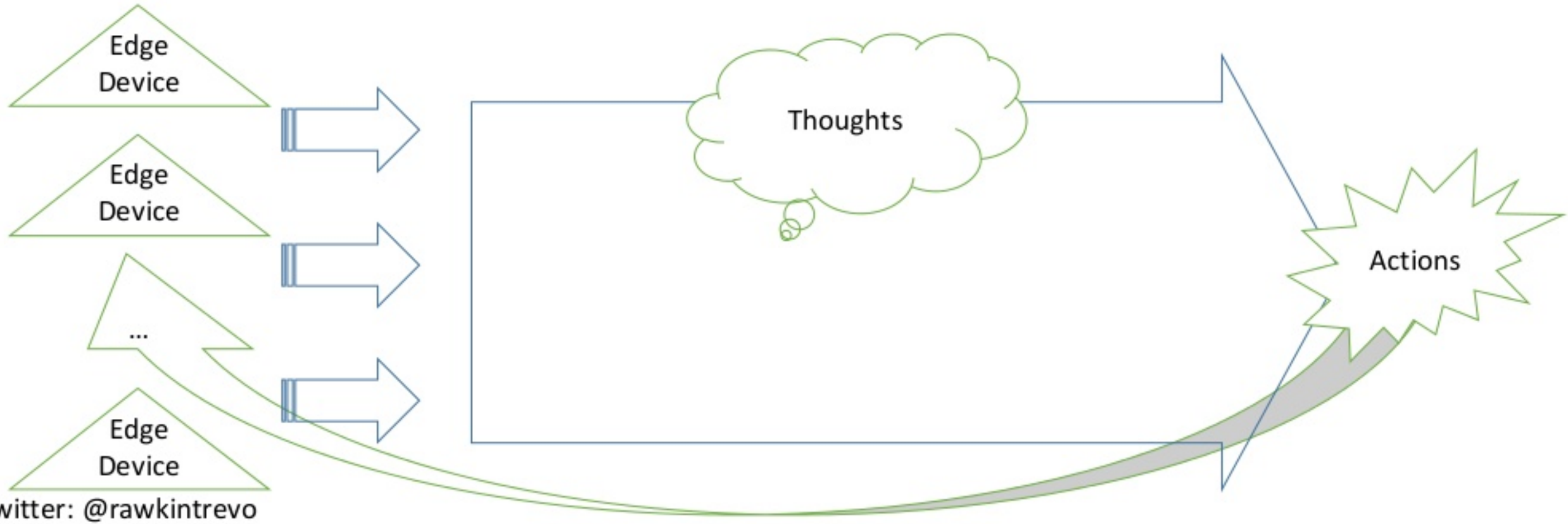
Edge devices produce data (sensors, drones, cars, etc.)

A Diagram: Collecting from environment



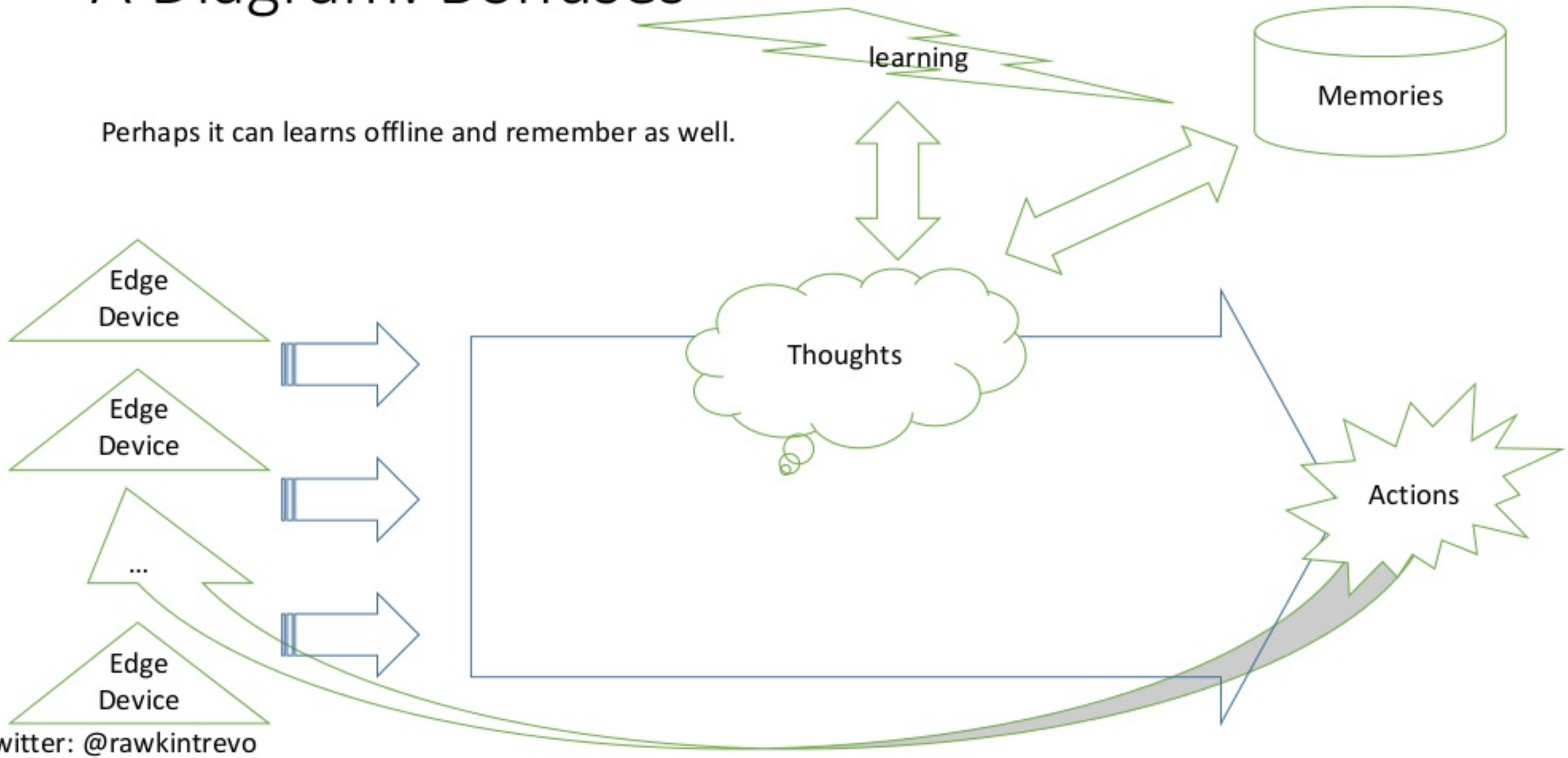
A Diagram: Learn, understand, adapt, affect

That “Something” needs to support learning and decision frame works



A Diagram: Bonuses

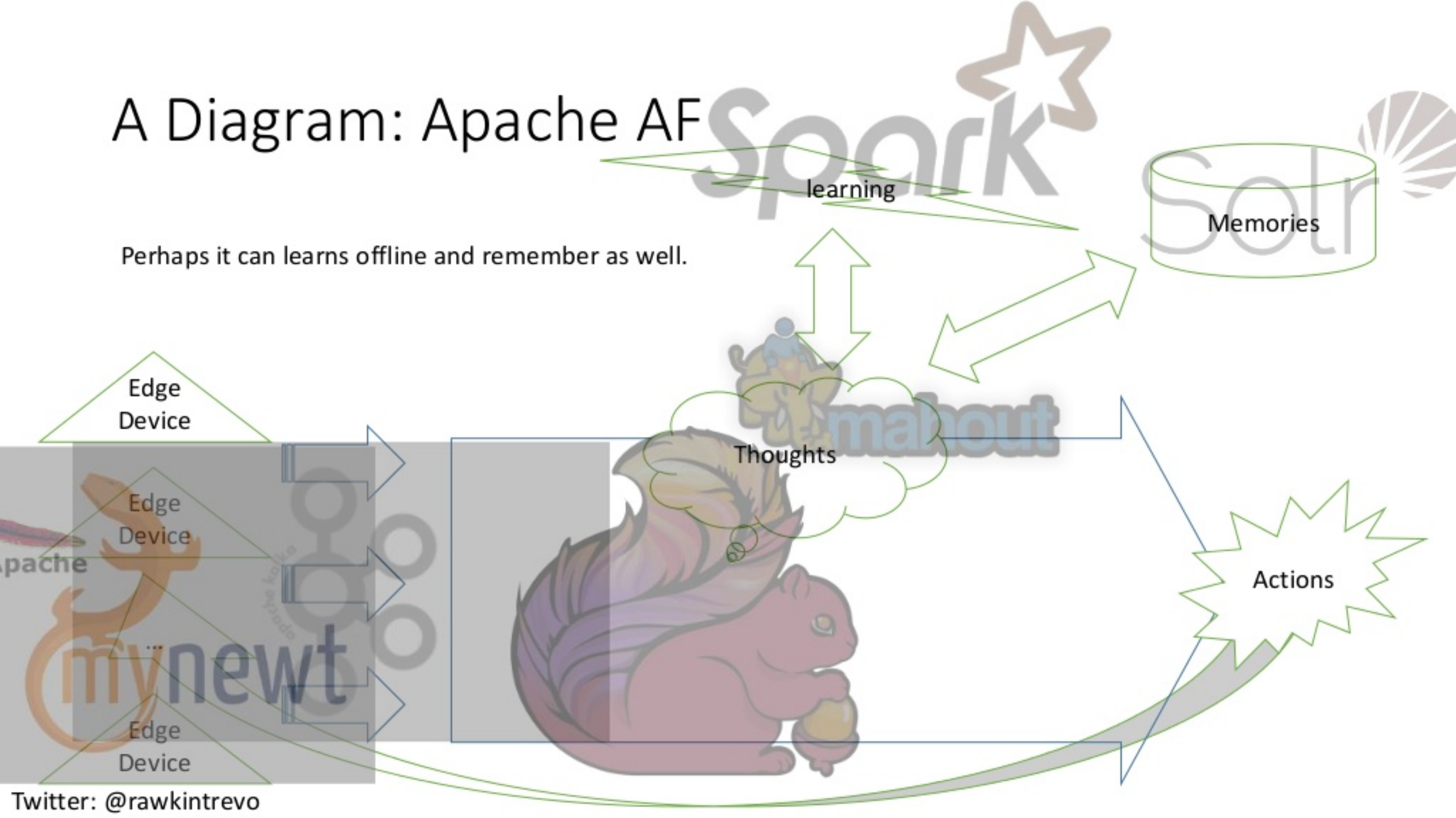
Perhaps it can learns offline and remember as well.



Twitter: @rawkintrevo

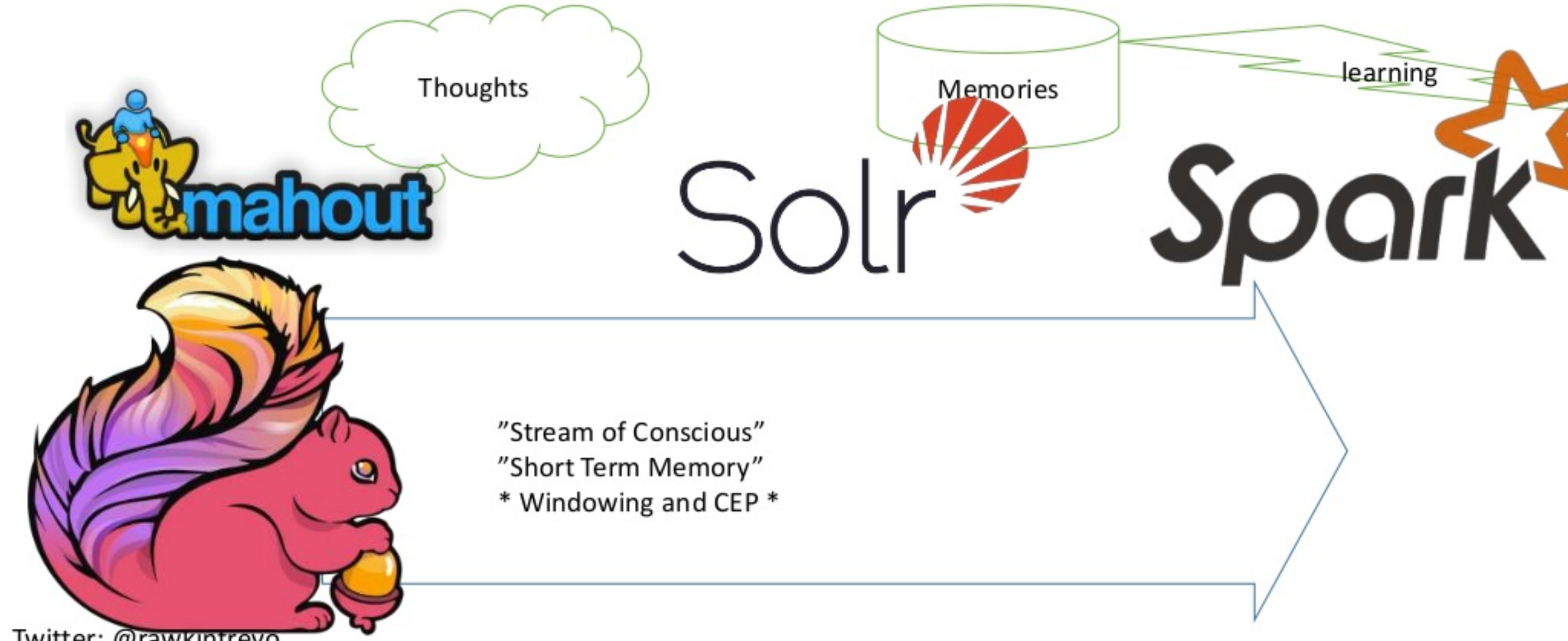
A Diagram: Apache AF

Perhaps it can learns offline and remember as well.

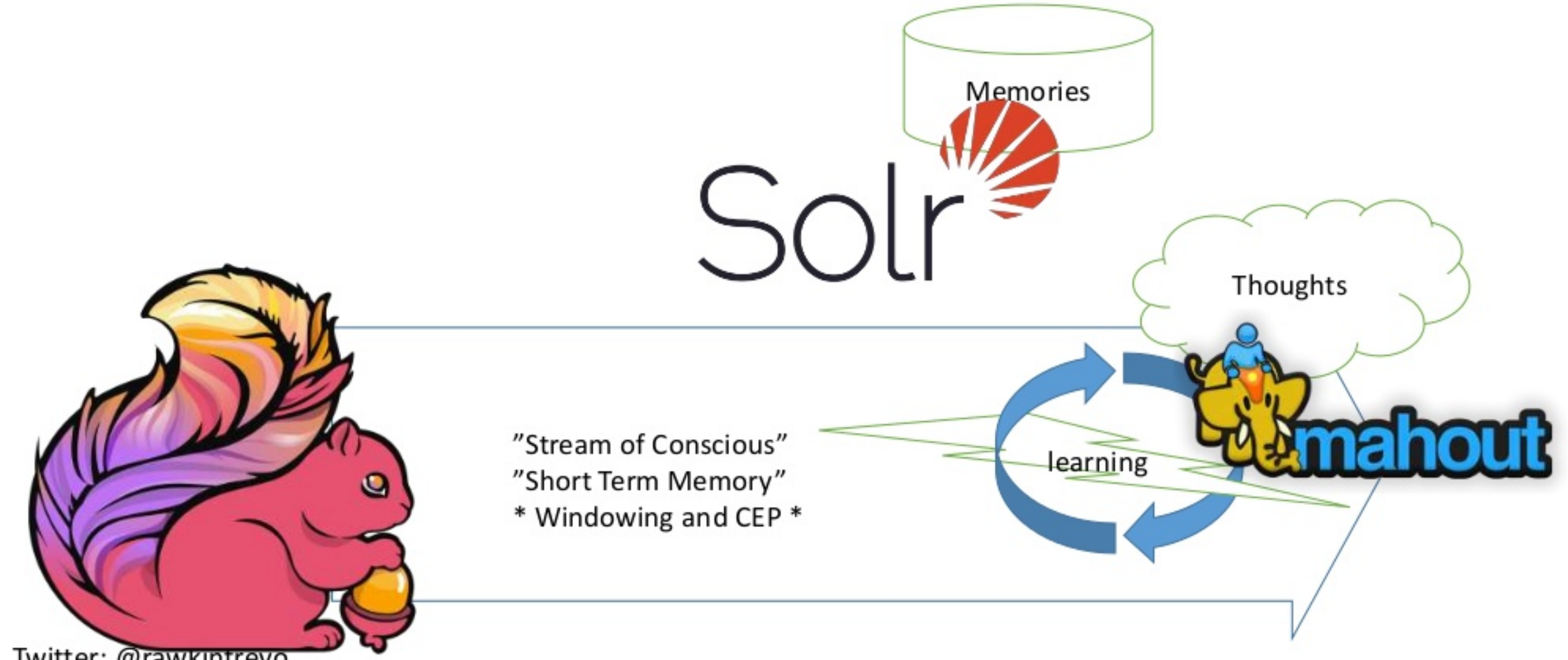


Twitter: @rawkintrevo

A Diagram: Apache AF (Lambda)



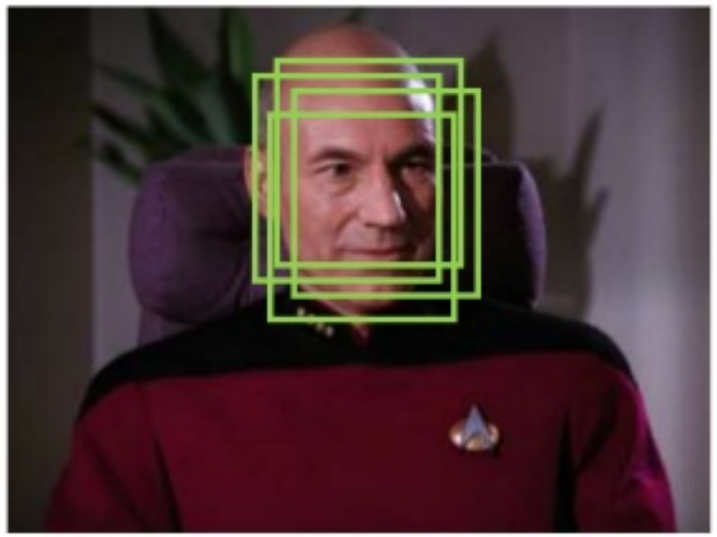
A Diagram: Apache AF (Kappa)

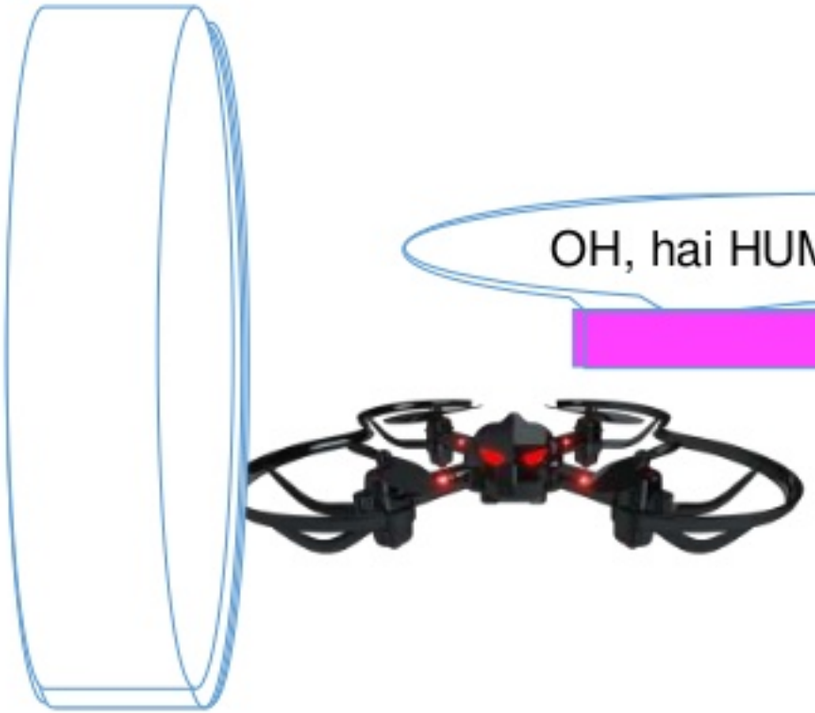


Technically “Borg-style” AI, not Cylons

- A finer technical point for those familiar with the Cylons and the Borg
- “Hive Mind” Architecture

Solr Returns Closest Match
Drone says “hi” or shoots to kill





Shape of things to come.

”Science Fiction” of 10 years ago, today is domain of hobbyists

Demo presented here is “Science Fair” grade AI.

Vlad Putin’s recently talking about “it is undesirable for anyone to monopolize AI”. (Yay Apache!)

Twitter: @rawkintrevo