Combining Semantic Tagging and Support Vector Machines to Streamline the Analysis of Animal Accelerometry Data




Lianli Gao[1], Hamish Campbell[2], Craig Franklin[2], Jane Hunter[1]

[1]eResearch Lab, The University of Queensland, Brisbane, Australia

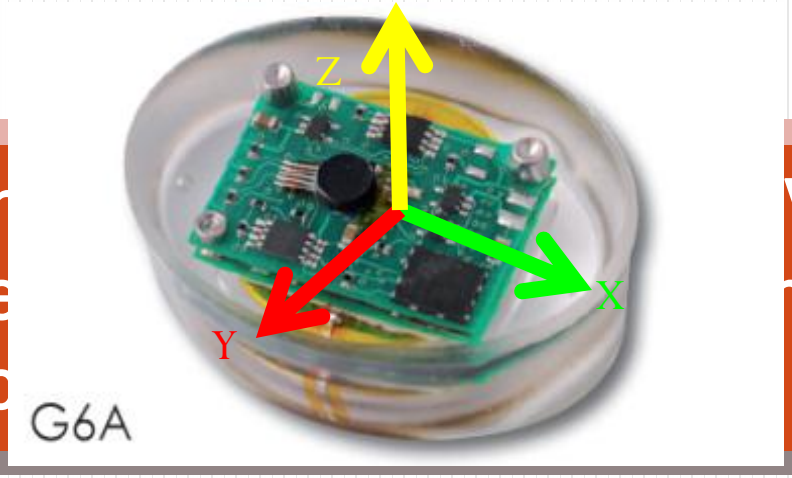
[2]ECO-Lab, The University of Queensland, Brisbane, Australia

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

http://oztrack.org

AUSTRALIA

Combining Semar Vector Machines to Stre himal Acc






Lianli Gao[1], Hamish Campbell[2], Craig Franklin[2], Jane Hunter[1]

[1]eResearch Lab, The University of Queensland, Brisbane, Australia

[2]ECO-Lab, The University of Queensland, Brisbane, Australia

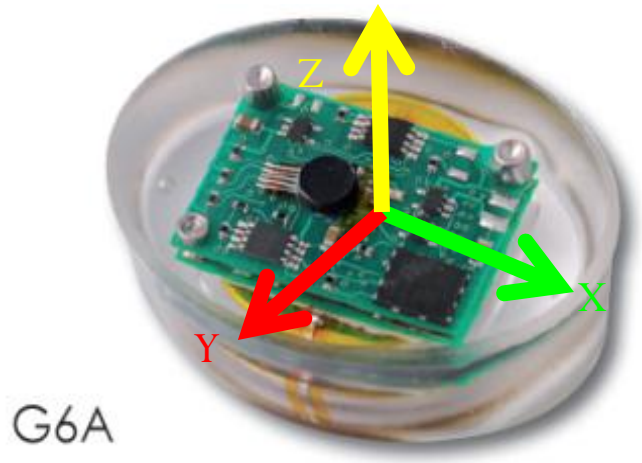
THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Overview

- Background
- Objectives
- Methodology
  - System Architecture
  - User Interface
  - Data Collection
  - Evaluation Result
- Conclusions & Future work

Data Collection-Device

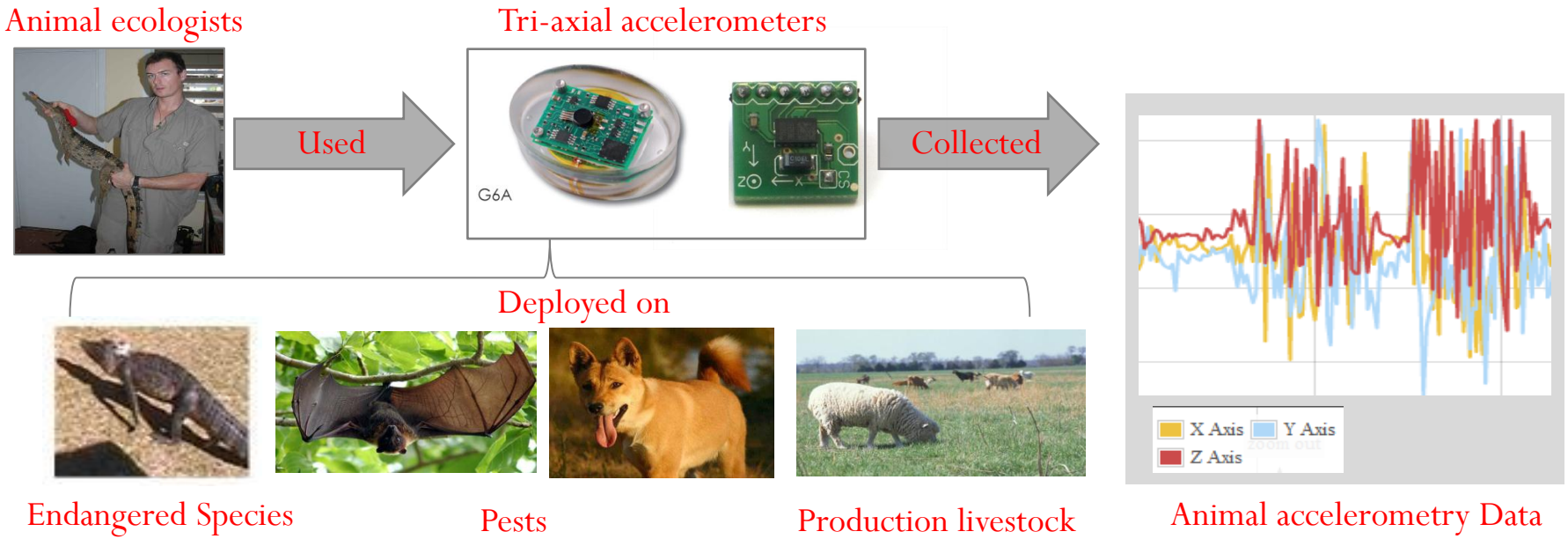
- Devices parameters
  - A tri-axial accelerometer named G6A
  - Produced by CTL(Cefas Technology Limited)
  - 40mm*28mm*16.3mm
  - 16MB memory
  - 7.3g weight in air
  - Sampling rates
    - From 1Hz to 30Hz



G6A

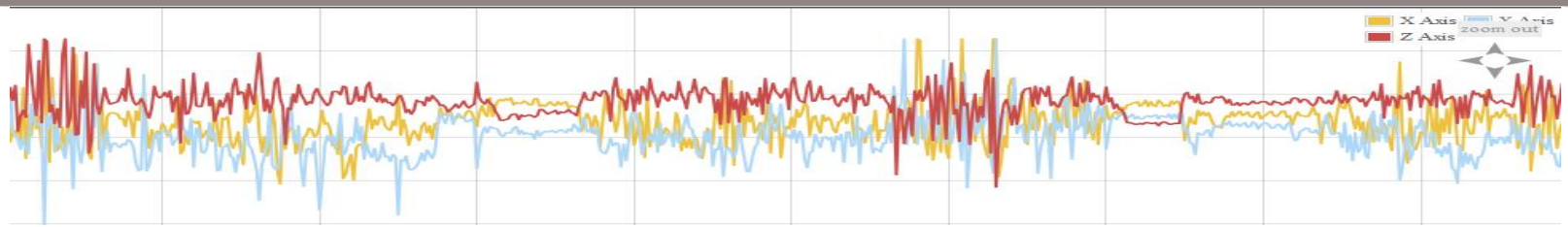
Background: accelerometer data

- Explosion in animal-attached accelerometers
  - To monitor animal movements and behaviour
- Collected an avalanche of raw tri-axial accelerometer data streams
  - Enable the identification of specific animals behaviours

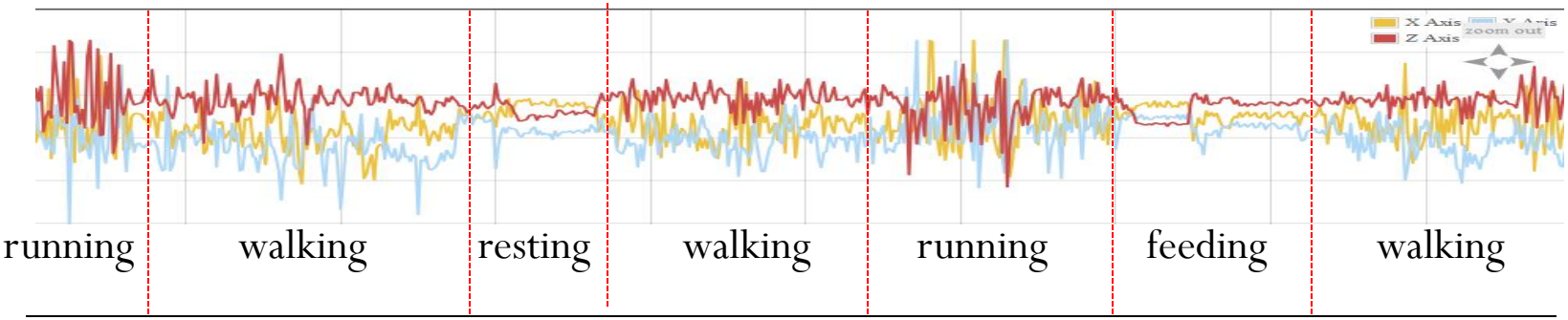


Background: accelerometer data

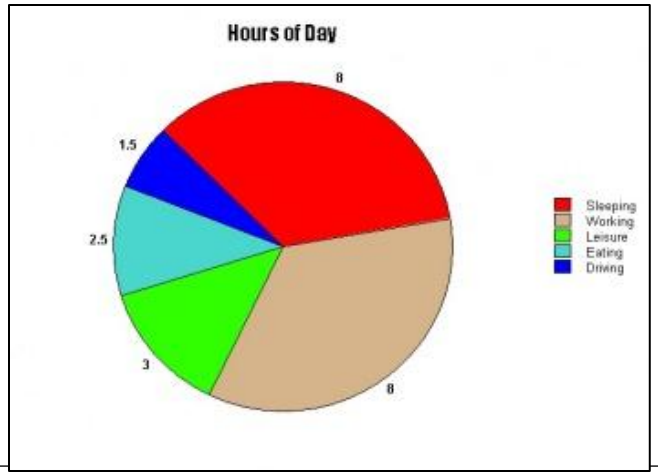
Step 1
Collect data



Step 2
Analyse data



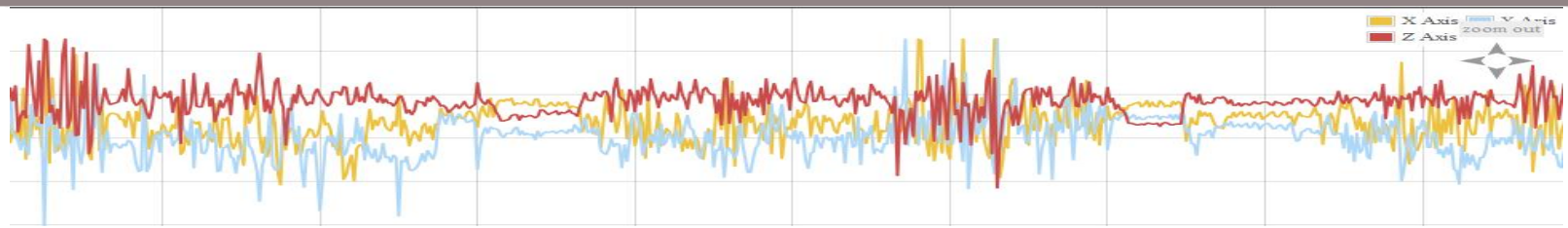
Step 3 –Visualization



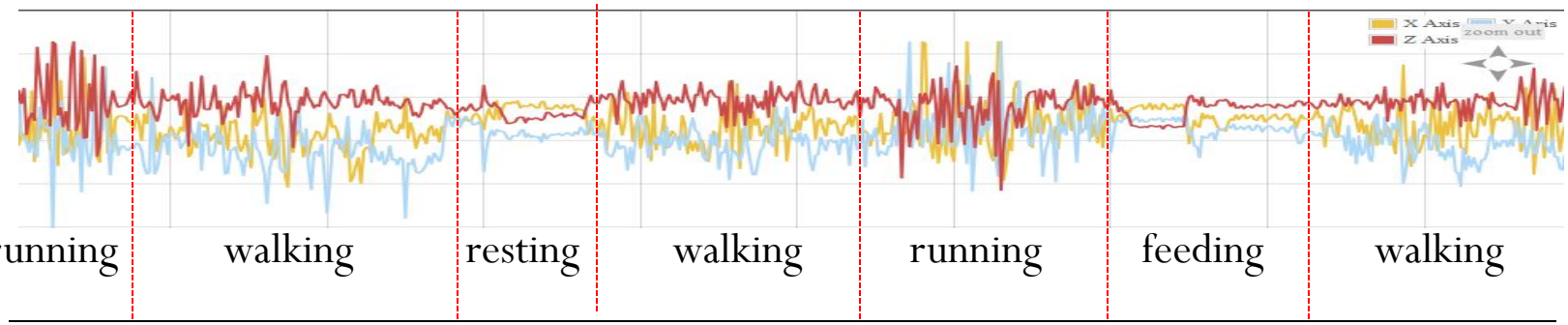
Understand animal health,energy consumption, food/water requirements

Background: accelerometer data

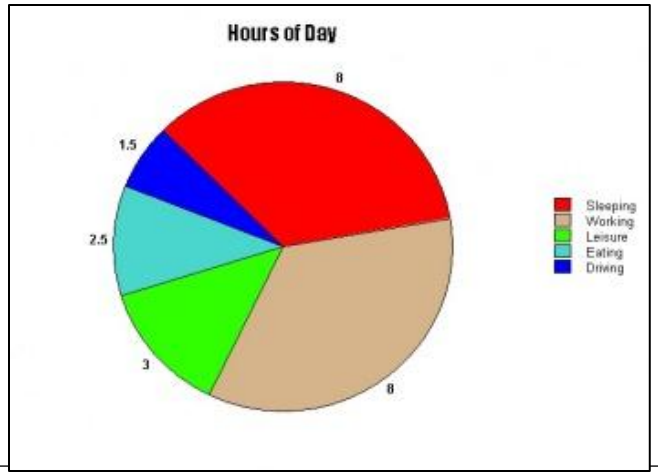
Step 1
Collect data



Step 2
Analyse data



Step 3 –Visualization



Understand animal health,energy consumption, food/water requirements

Background: challenges

- Limitations of raw 3D acceleromtery data streams
  - Numerical, unstructured, complex, imprecise, large volume
  - Poor data representation
- Problems
  - Massive volumes of complex data
  - Lack of automatic analysis
  - Lack of pattern recognition tools
  - Manual analysis
    - Onerous, time consuming, expensive
    - Poor quality, subjective
- Wild animal activities
  - Difficult to monitor and analyze

What is 87???

What is the measurement unit??

	A	B	C	D	E	F	G
1	07.06.07 19:00:00	___	87.001		14.06.07 13:	___	85.001
2	07.06.07 19:00:09	___	86.001	?	14.06.07 13:	___	85.001
3	07.06.07 19:00:18	___	87.001		14.06.07 13:	___	85.001
4	07.06.07 19:00:27	___	87.001		14.06.07 13:	___	85.001
5	07.06.07 19:00:36	___	89.001		14.06.07 13:	___	85.001
6	07.06.07 19:00:45	___	87.001		14.06.07 13:	___	83.001
7	07.06.07 19:00:54	___	87.001		14.06.07 13:	___	85.001
8	07.06.07 19:01:03	___	87.001		14.06.07 13:	___	85.001
9	07.06.07 19:01:12	___	87.001		14.06.07 13:	___	85.001
10	07.06.07 19:01:21	___	87.001		14.06.07 13:	___	83.001
11	07.06.07 19:01:30	___	87.001		14.06.07 13:	___	83.001

How to analyze???

How to improve???

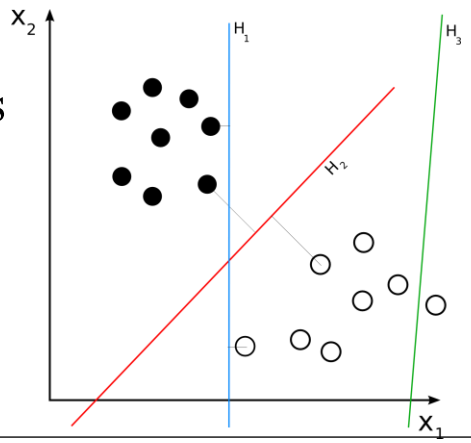
Possible solutions

- Semantic Tagging
  - attach meaningful descriptions
  - mark up events/activities
  - enrich data semantic meaning
  - collect expert domain knowledge

	A	B	C	D	E	F	G
1	07.06.07 19:00:00	___	87.001		14.06.07 13:	___	85.001
2	07.06.07 19:00:09	___	86.001	?	14.06.07 13:	___	85.001
3	07.06.07 19:00:18	___	87.001		14.06.07 13:	___	85.001
4	07.06.07 19:00:27	___	87.001		14.06.07 13:	___	85.001
5	07.06.07 19:00:36	___	89.001		14.06.07 13:	___	85.001
6	07.06.07 19:00:45	___	87.001		14.06.07 13:	___	83.001
7	07.06.07 19:00:54	___	87.001		14.06.07 13:	___	85.001
8	07.06.07 19:01:03	___	87.001		14.06.07 13:	___	85.001
9	07.06.07 19:01:12	___	87.001		14.06.07 13:	___	85.001
10	07.06.07 19:01:21	___	87.001		14.06.07 13:	___	83.001
11	07.06.07 19:01:30	___	87.001		14.06.07 13:	___	83.001

```
<Observation>
        <observes>x-axis acceleration</observes>
        <unit>m/s</unit>
        <hasLocation>Brisbane</hasLocation>
        <hasValue>87.001</hasValue>
        <hasTime>2007-06-07T19:00:54</hasTime>
</Observation>
```

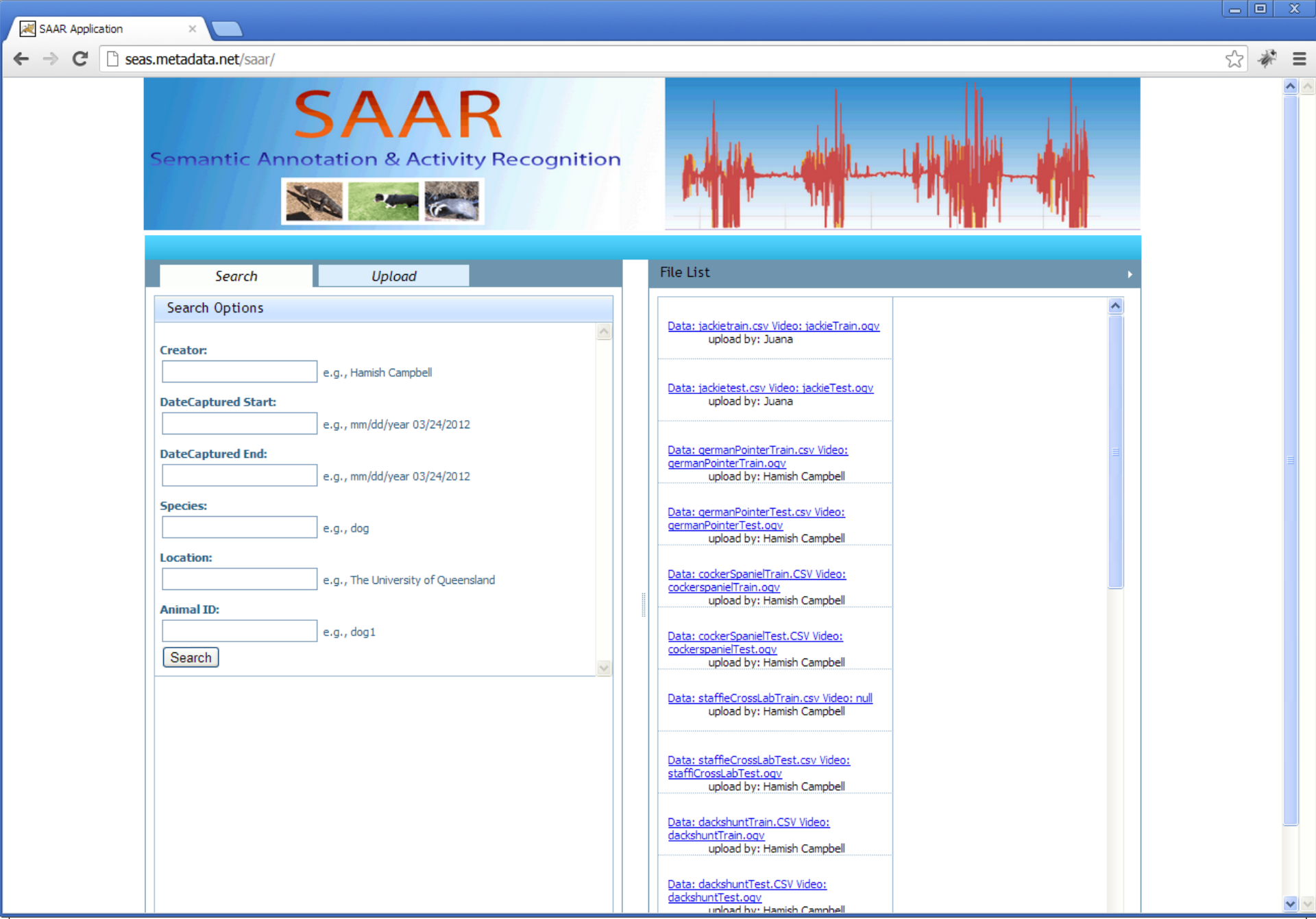
- Support Vector Machines
  - supervised machine learning models
  - support automatic classification



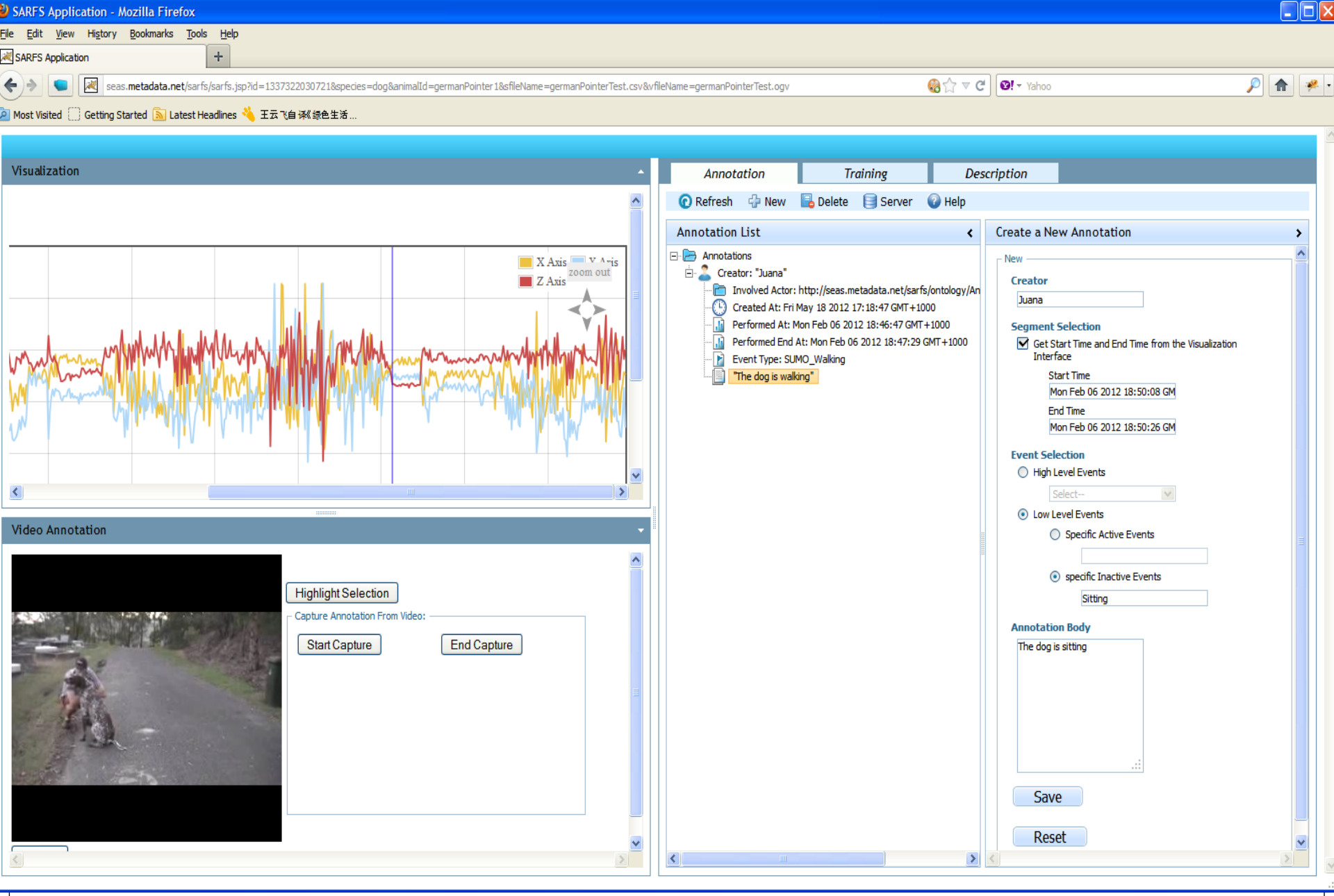
System objectives

- Web-based repository
  - To upload and share tri-axial accelerometer animal datasets
  - To search, retrieve and compare datasets
- Annotation services
  - Record, share and re-use expert knowledge on animal movements within tri-axial accelerometer data streams
  - Using terms from pre-defined ontologies
  - Analyze, tag and visualize 3D accelerometry datasets
  - Synchronized with video to compare with base truth
- Automated analysis services
  - Build activity recognition models by training classifiers using features extracted from pre-annotated training sets
  - To improve the quality of results generated by SVM-based activity recognition classifier
  - To enable the sharing, re-use and refinement of activity recognition classifiers developed for specific species, between scientists.
- Simple statistical visualisation of annotated data streams

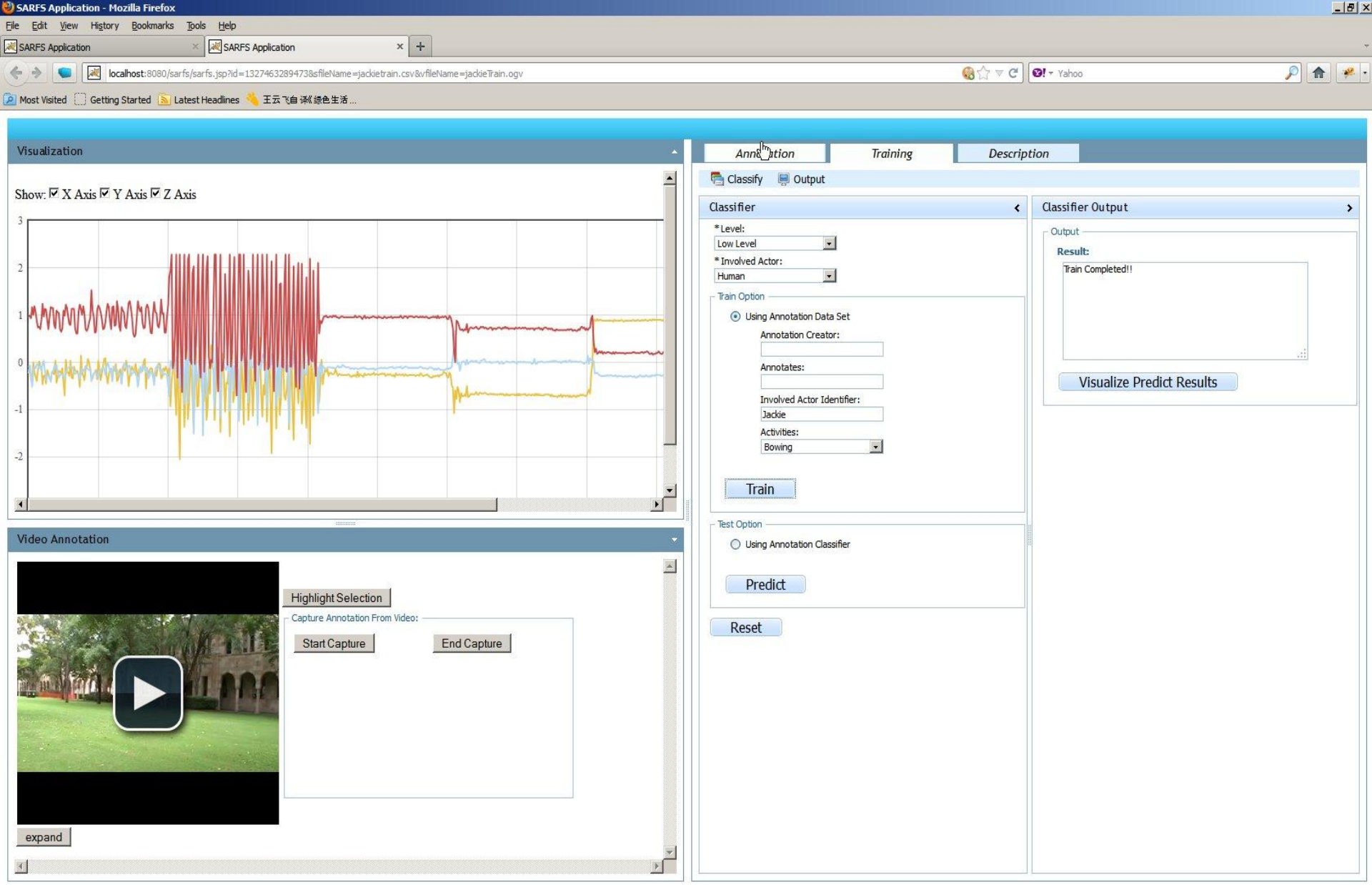
Screenshot of the SAAR upload user interface



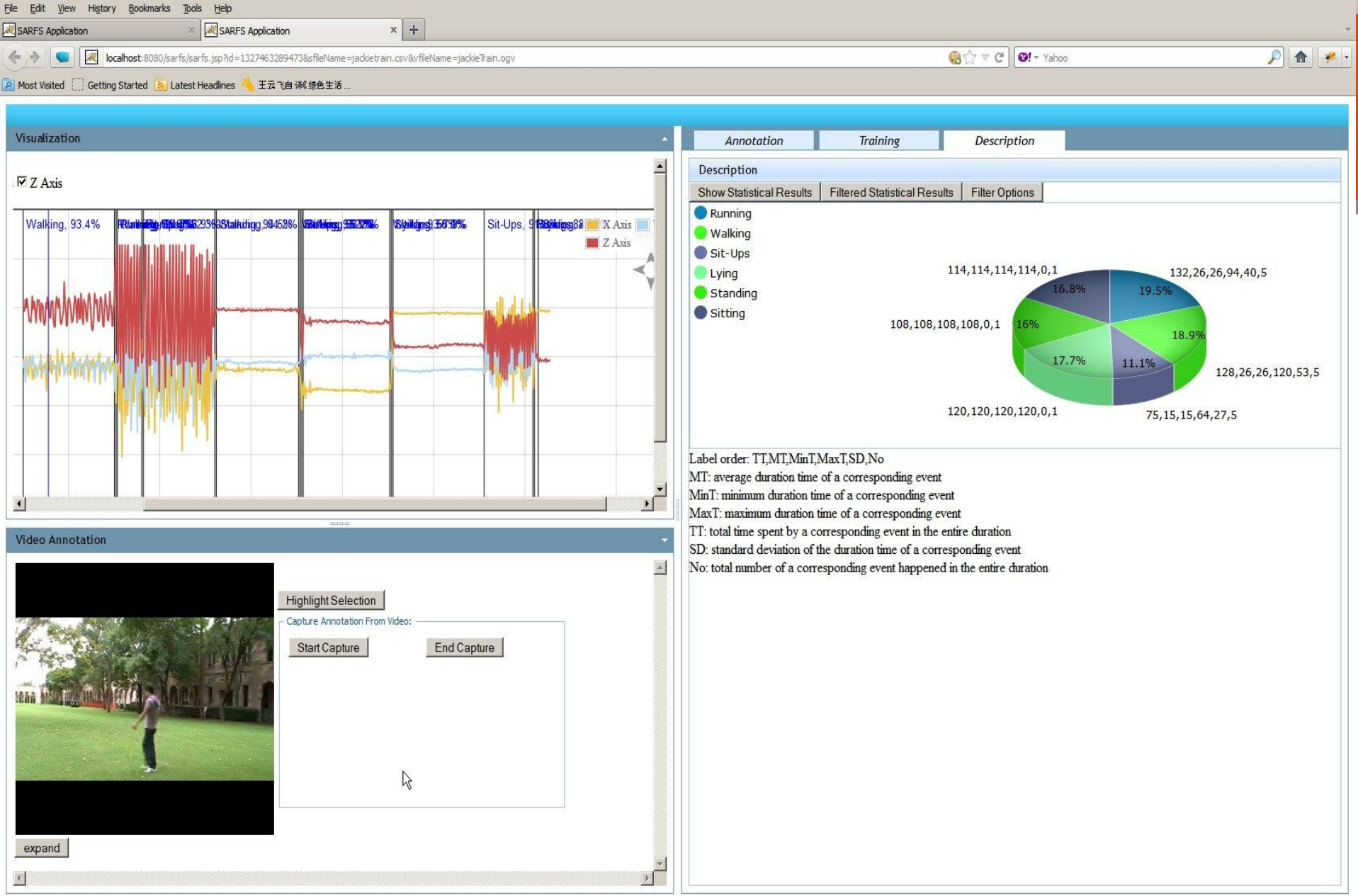
Screenshot of the SAAR search user interface



Screenshot of SAAR Plot-Video visualization interface and the annotation interface

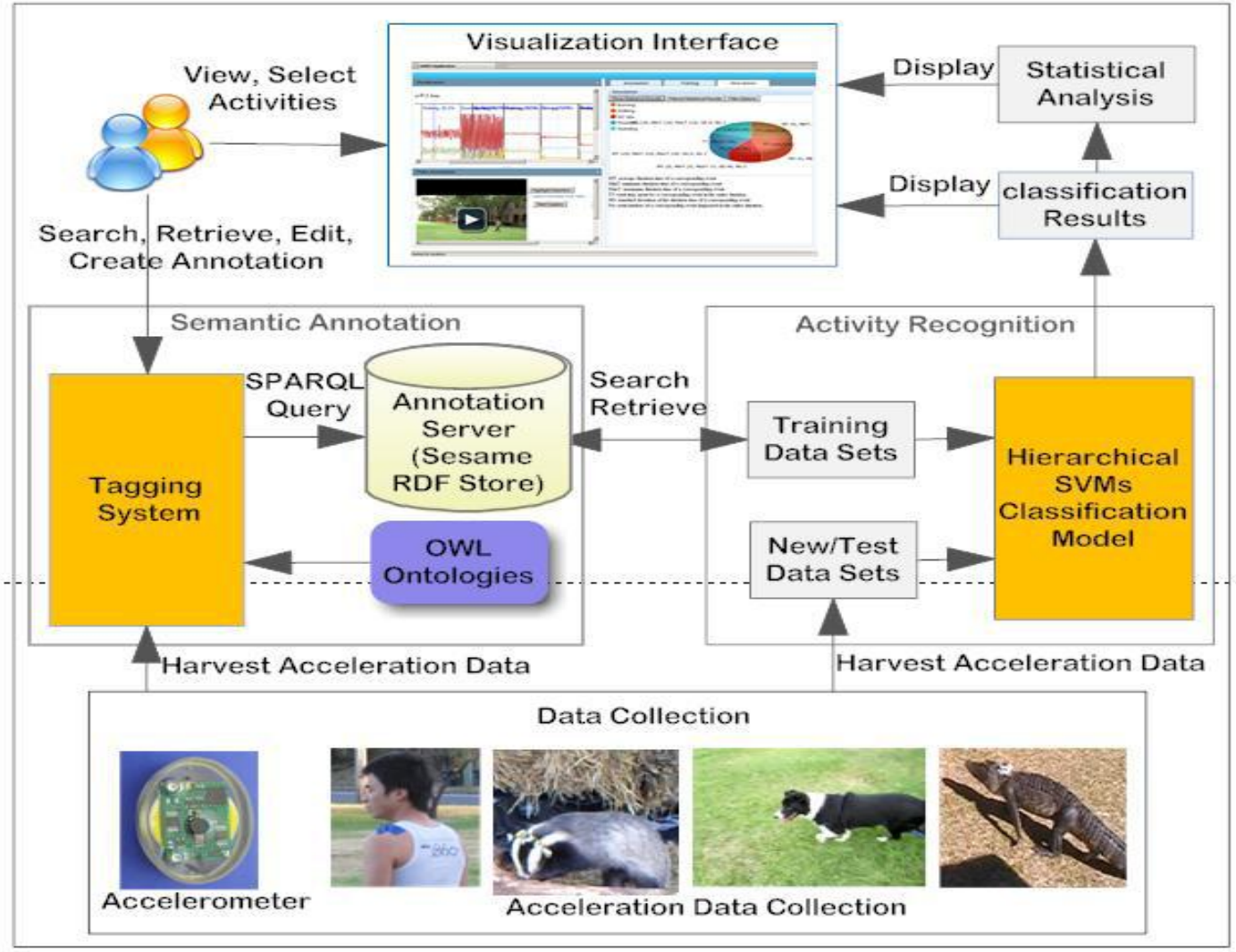


User interface when retrieving all the specific annotation to train a SVM activity classifier



Screenshot of the SAAR interface with human activity identification results

System Architecture

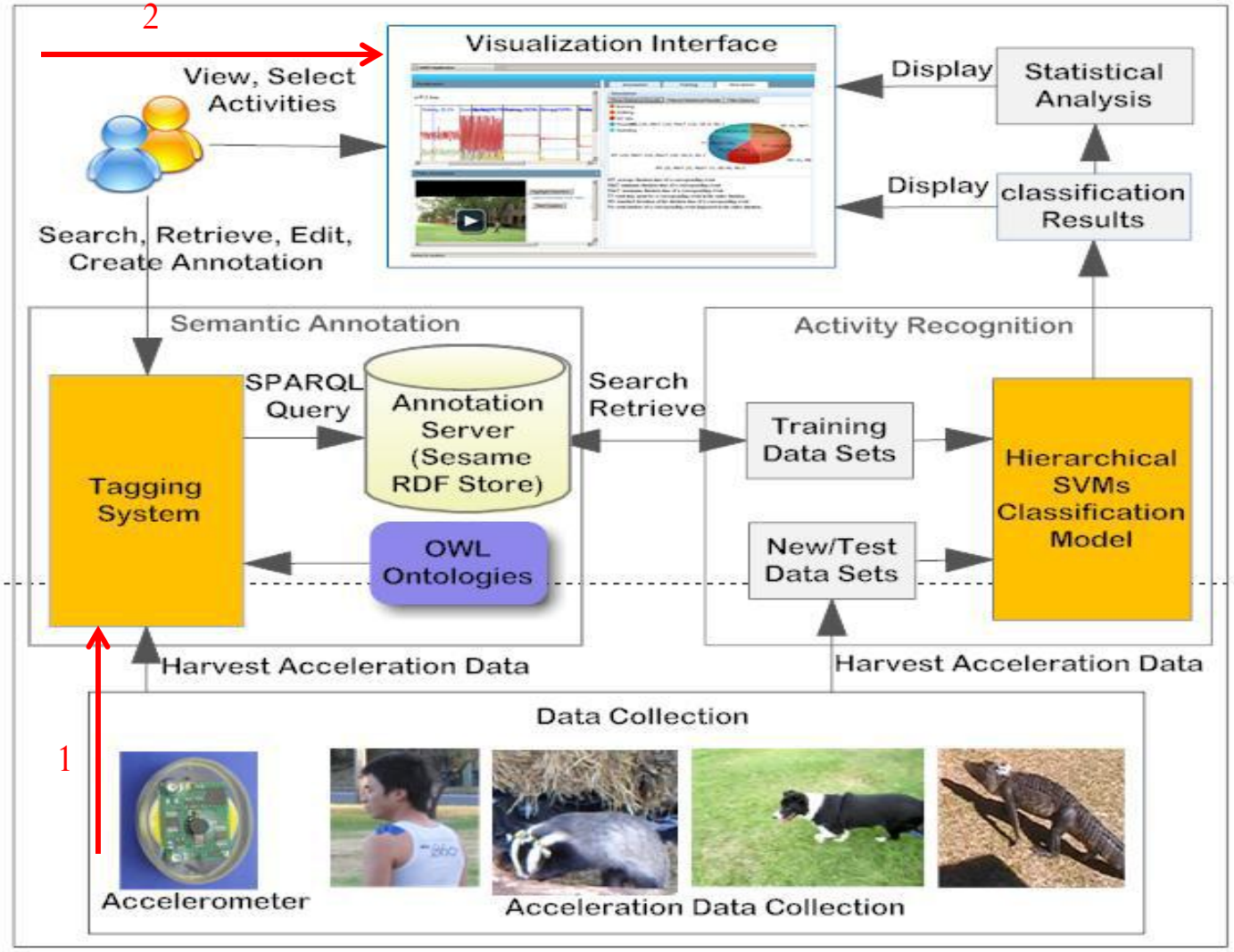


SAAR –Semantic Annotation and Activity Recognition System

Visualisation

Step 2:
display raw data and video in the visualization interface

Step 1:
upload raw data and videos

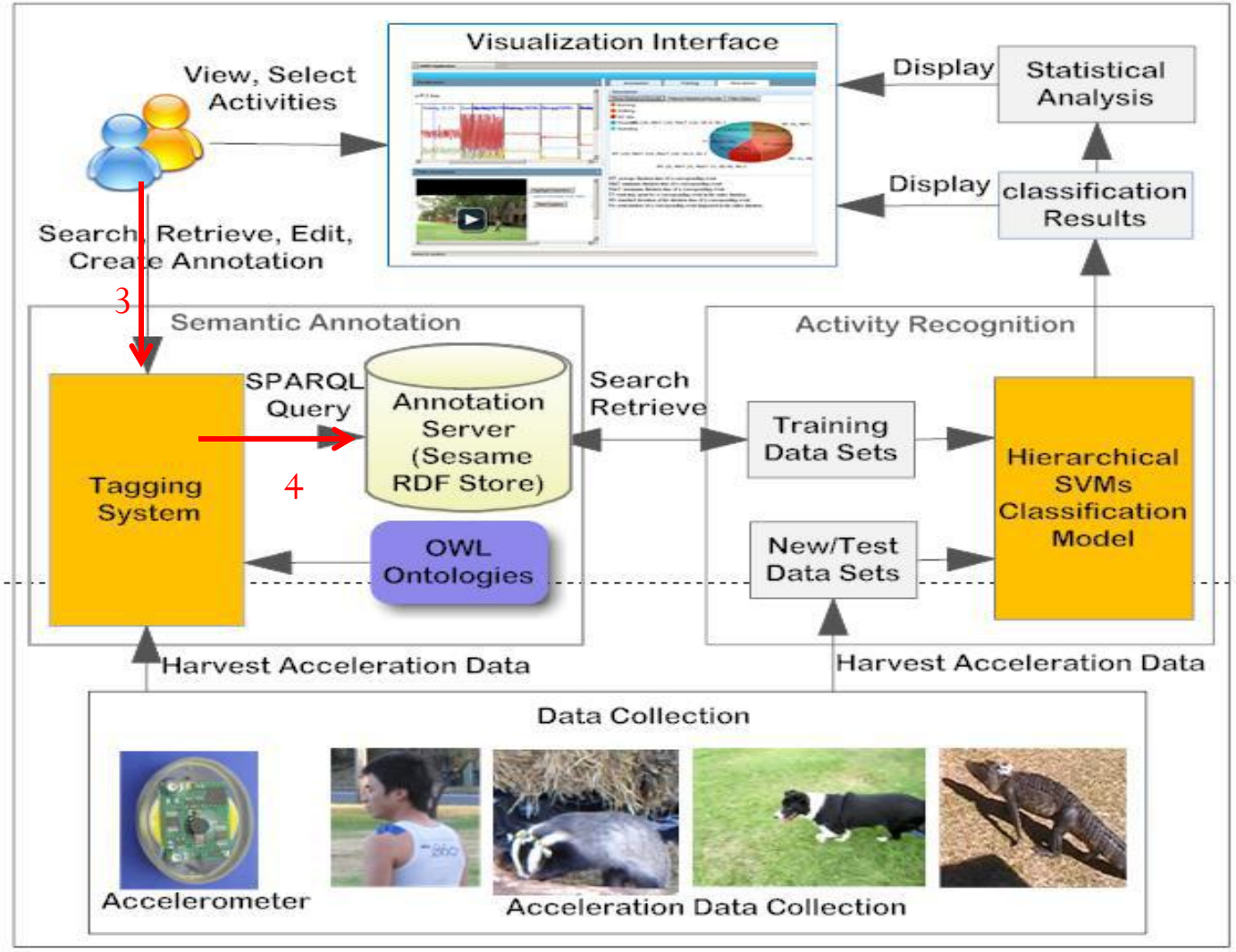


SAAR –Semantic Annotation and Activity Recognition System

Annotation

Step 3:
users create,
search, retrieve,
edit annotations

Step 4: created
annotations marked up
with ontologies,
and then saved in
the annotation server



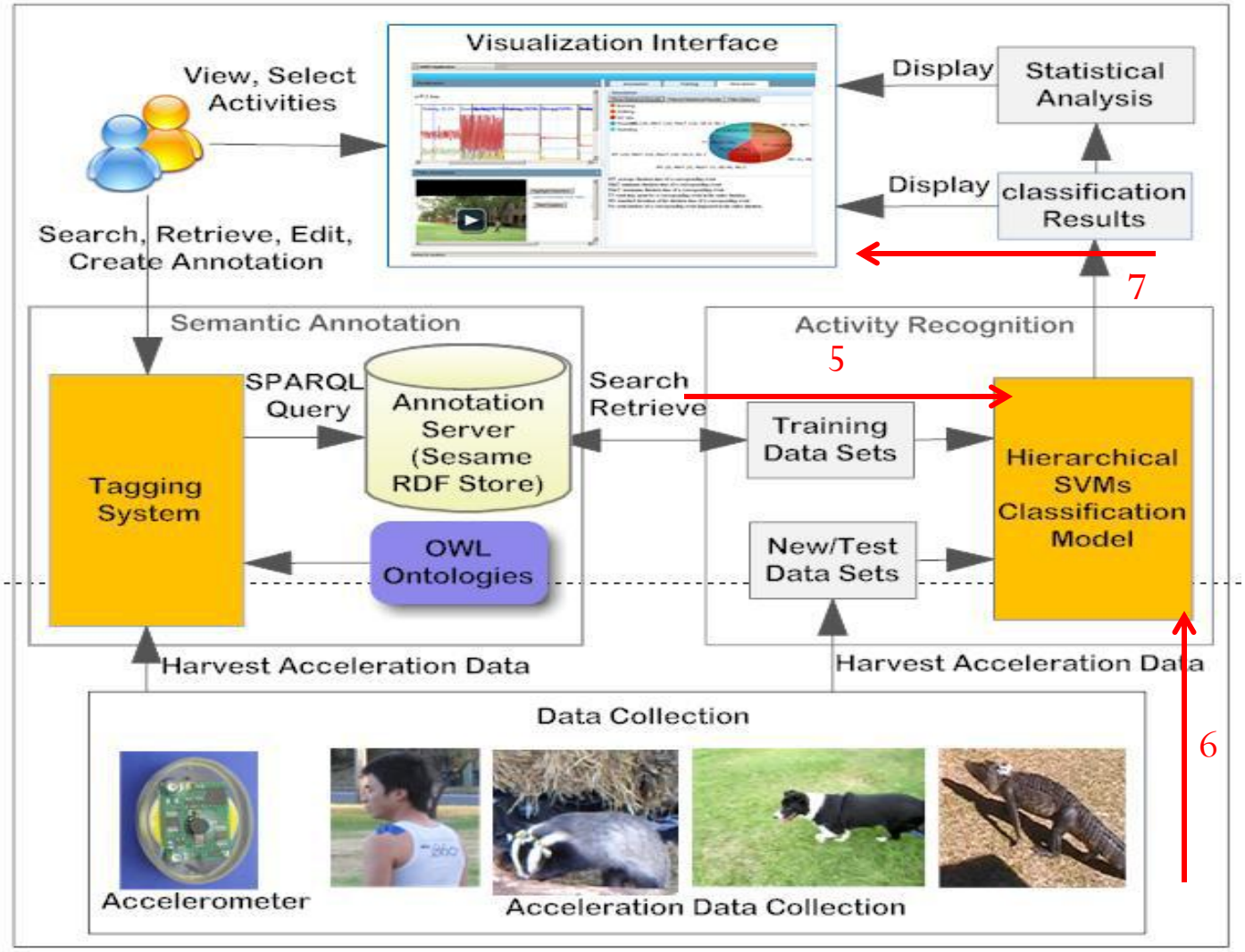
SAAR –Semantic Annotation and Activity Recognition System

Model training and use

Step 5:
search annotations
as training data
set to train SVMs
classification models

Step 6:
trained classifiers
take new acceleration
data as input to classify
animal activities

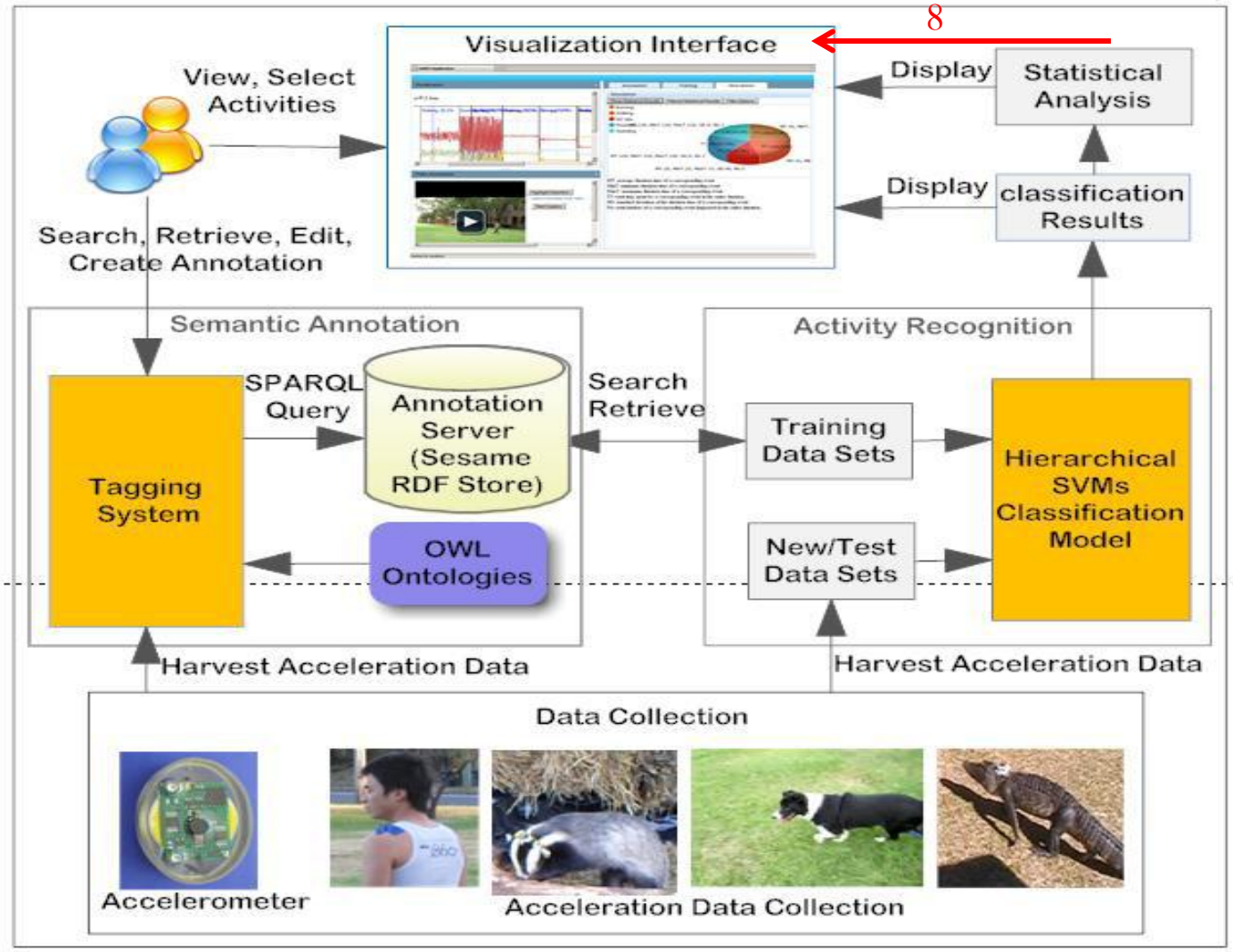
Step 7:
classification results
display in the
visualization interface



SAAR –Semantic Annotation and Activity Recognition System

Statistical summary

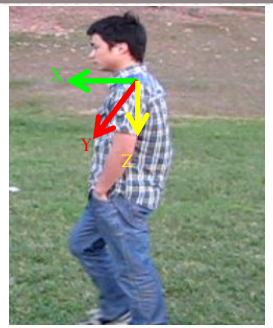
Step 8:
statistical analysis takes classification results as input and displays the result in a pie chart



SAAR –Semantic Annotation and Activity Recognition System

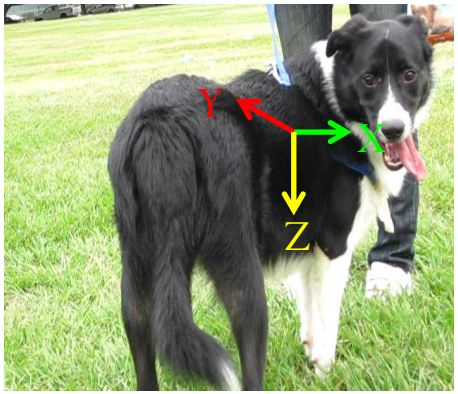
Data Collection-Human

- Human data collection
  - Location : University of Queensland
  - 4 males & 4 females
  - Age ranges from 25 to 38
  - For each voluntary
    - 3 minutes walking
    - 3 minutes running
    - 1 minutes sit-ups
    - 3 minutes standing
    - 3 minutes sitting
    - 3 minutes lying
    - Randomly perform these 6 activities over 15 minutes



Data Collection-Dog

- Dog data collection
  - Location: Brisbane
  - 6 dogs of different breeds
  - Height and weight
    - 20cm-55cm, 8.9kg-25.8kg
  - Each dog led by its owner to perform
    - 2 minutes walking, running, standing,
    - 2 minutes Sitting and lying
    - Randomly perform these 5 activities for 10 minutes
  - One well-trained king charles spaniel
    - 1 minutes foraging/digging
    - 1 minutes climbing



Evaluation-Metrics

- Standard evaluation metrics
  - Accuracy

$$\text{Accuracy} = \frac{\text{number of true positives} + \text{number of true negatives}}{\text{number of true positives} + \text{false positives} + \text{true negatives} + \text{true negatives}}$$

  - Precision

$$\text{Precision} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{false positives}}$$

  - Sensitivity

$$\text{Sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}$$

  - Specificity

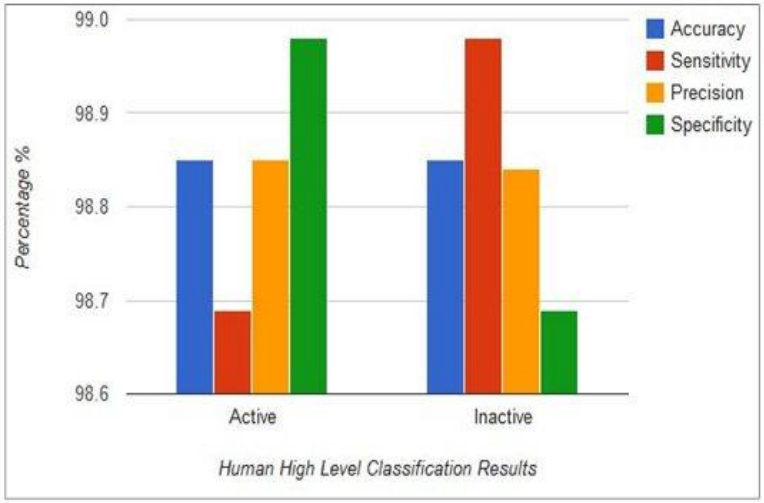
$$\text{Specificity} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false negatives}}$$

		Type as determined by a classifier	
		True	False
Test outcome	+ve	True positive	False positive
	-ve	False negative	True negative

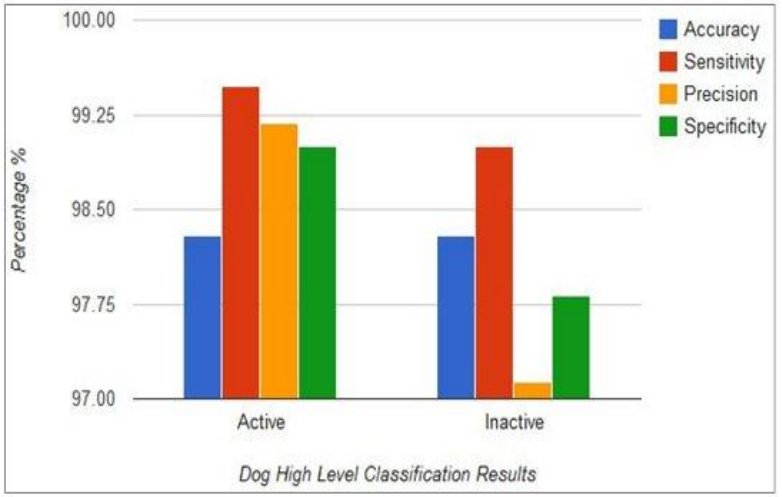
Evaluation-Result-1
Humans and Dogs

Human Data Set

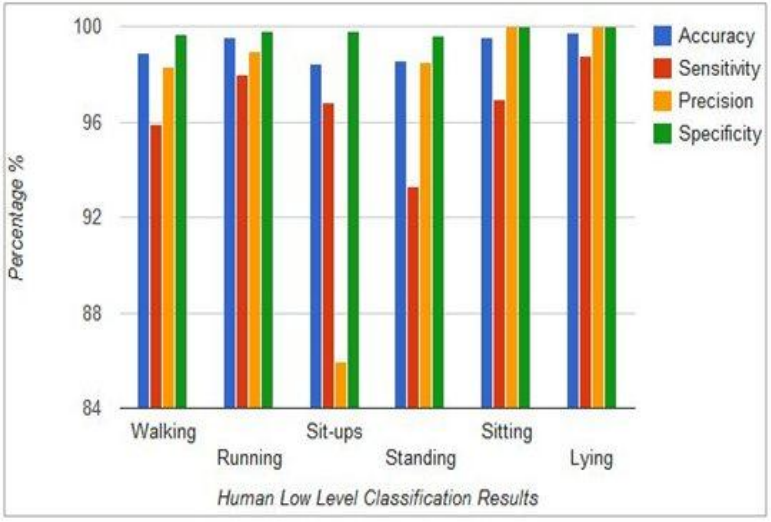
High Level

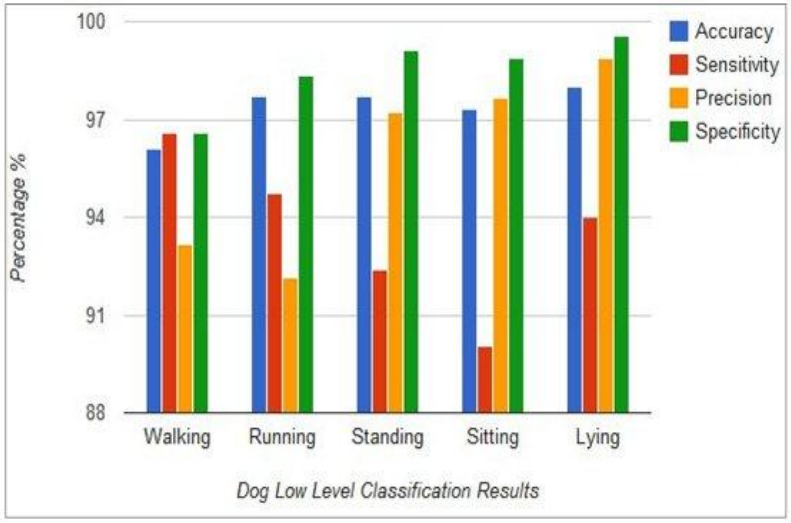


Dog Data Set



Low Level



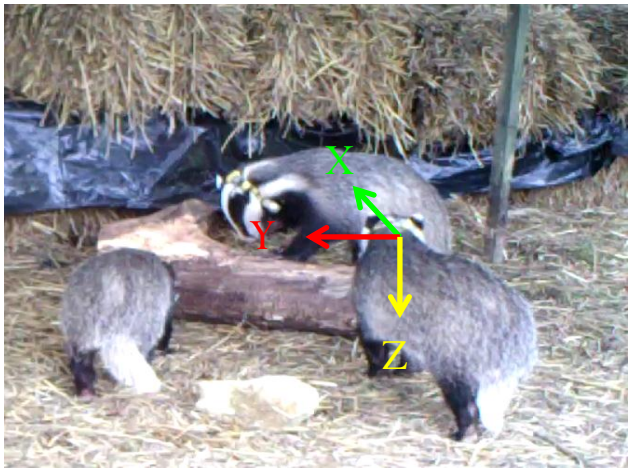


Humans

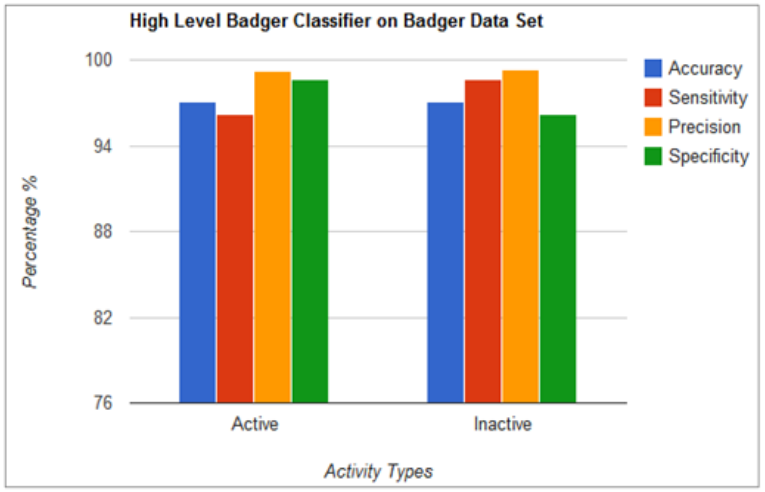
Dogs

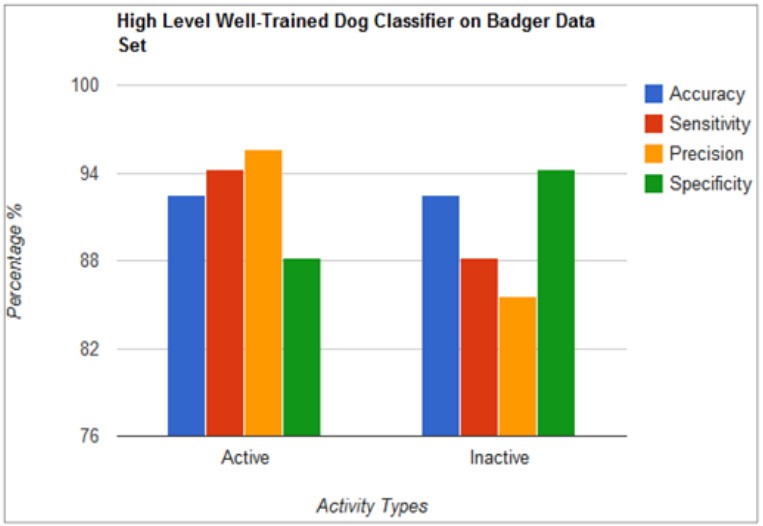
Data Collection-Eurasian Badger

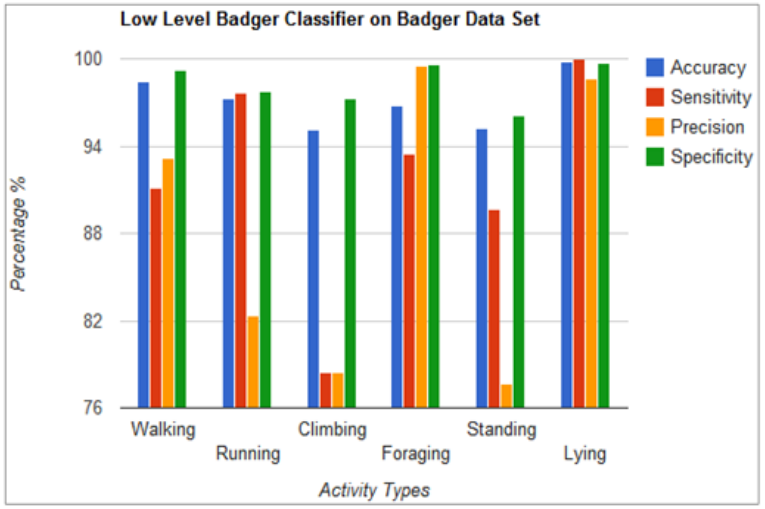
- Eurasian badger collection
  - Location: Somerset, United Kingdom
  - 3 Eurasian badgers
  - Six activities were recorded
    - Walking
    - Running
    - Climbing
    - Foraging
    - Standing
    - Lying

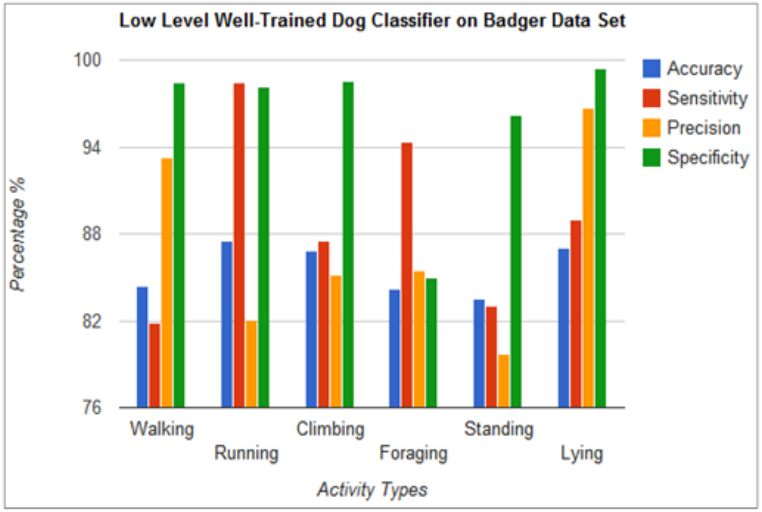


Evaluation-Result 2









Badger[3]

Badger data, dog model

Evaluation - Conclusions

- High level classifiers performs better than low level classifiers
  - High level: accuracy>96%, sensitivity>97%,specificity>96%
  - Low level: accuracy>96%, sensitivity>80%, precision>80%
- Human classifier performs better than dog classifier which performs better than badger classifier
  - More noise the "wilder" the animals
- Species-specific classification models perform better than migrating the classification models across species, but migration still yields reasonable results

Summary

- The Semantic Annotation and Activity Recognitions system delivers
  - An easy-to-use Web-based repository
    - For accelerometer data streams
  - A set of semantic tagging, visualization services
    - For annotation meaning of accelerometer data streams
  - Activity recognition services
    - Accuracy decreases for more "unpredictable" animals
    - Accuracy decreases across species
    - BUT still very useful

Future work

- Future work
  - Integrate GPS data to track animal trajectory + add map visualization
  - Acquire data from wild dingoes, captive tigers, wild birds
  - Apply captive models to wild animals
    - Dogs to foxes, dingoes
    - Birds to bats
    - Horses to camels
  - Evaluate different machine learning methods

Acknowledgement

- Owen R Bidder, from the Swansea Moving Animal Research Team (SMART) at Swansea University College of Science, for providing access to the badger accelerometer data, and for providing comprehensive and useful feed-back
- The China Scholarship Council




Questions?

- Thank you!
- Contact:
  - Lianli Gao: l.gao1@itee.uq.edu.au
  - Hamish Campbell: hamish.campbell@uq.edu.au
  - Craig Franklin: c.franklin@uq.edu.au
  - Jane Hunter: jane@itee.uq.edu.au
- Websites
  - eResearch Group: http://itee.uq.edu.au/~eresearch
  - ECO-Lab: http://www.uq.edu.au/eco-lab/
  - SAAR: http://seas.metadata.net/saar/

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA