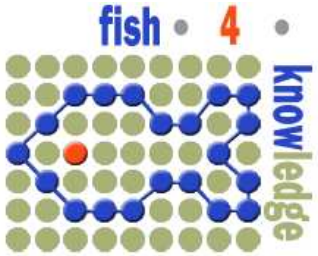
Welcome



Jenny Benois-Pineau LABRI - Laboratoire Bordelais de Recherche en Informatique

Rafael Garcia University of Girona

Stefano Bertolo European Commission

Project Teams

- **Univ of Edinburgh admin**: Fisher
- **Univ of Edinburgh vision**: Boom
- **Univ of Edinburgh workflow**: Chen-Burger, Yang
- **Univ of Catania**: Giordano, Spampinato, Palazzo
- **National Applied Research Laboratories**: Lin
- **Centrum voor Wiskunde en Informatica**: Hardman, Ossenbruggen, Beauxis-Aussalet

Review: Overall Project Goals

1. Acquire, process and store **massive** video datasets: Proposal: 10 cameras, 2 years, 365 days, 12 hours/day, 3600 seconds/hour, 5-10 frames/second

Item	Plan	Actual
Frame rate	10	5
Frames	3×10^9	1.5×10^9
Bytes of compressed raw video	2×10^{14}	1×10^{14}
Detected Fish	10^{10}	1.4×10^9
Result data (bytes)	$10^{11} - 10^{12}$	4×10^{11}
Processing	-	400 core-years

2. Develop methods based on **ontologies and semantic web** concepts for allowing non-programming specialists access to massive datasets.

   Achievements: see User Interface demonstration

3. Build a **working prototype** by month 24, leaving last 12 months for evaluation and developing additional query answering capabilities.

   Achievements: prototype built and evaluated, second version built, to be demonstrated here

4. Work with marine biologists to produce **useful answers to biological questions**.

   Achievements: Many discussions and demonstrations, some early stage technology transfer, no exciting discoveries from the data (so far)

Biologist Empowerment: Questions we can answer

1. What species and numbers of fish appear, filtered by the user's choice of time of day, week of year, by year, camera, location?
2. What is the relative abundance of different species?
3. Show examples of videos, detections and classifications.
4. What other species were also present when species X was seen?
5. Are the observed numbers of species X increasing in the past 3 years?

Key Scientific Achievements

1. Computer vision methods for target detection and tracking in difficult environments, and for accurate species recognition in greatly unbalanced datasets.
2. New ways to present large amounts of time varying information through a complex facet-based user interface
3. Technological development of memory, processor, and task control systems suited for management of large amounts of computation.
4. Ontologies and vocabularies for goals and image processing to implement a virtual workflow model.

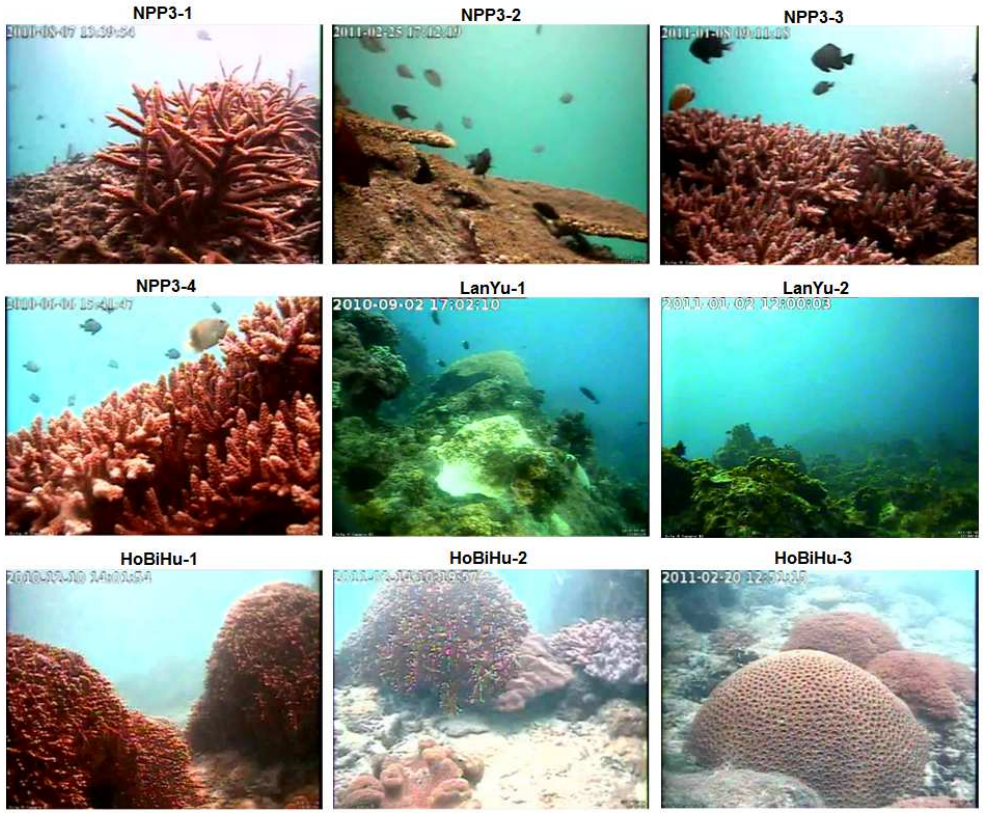
5. SQL based complete and abstracted storage representations for access in massive datasets
6. Control algorithms that allow on-demand as well as routine processing in a complex multi-processor context, where tasks fail occasionally (software, hardware, communications, etc).
7. Massive public fish database, 600 Gb, 23 species, 1.4 billion fish images, possibly the largest analysed public image database in the world.

Year 3 Goals

- Enhance the detection and tracking algorithms: DONE
- Extend the species recognition algorithm to more species and higher accuracy: now 23 most common species, up from 10 last year
- Complete system integration (workflow and user interface): DONE
- Evaluate system performance: DONE
- Enhance system to increase data analysis and query answering speed: DONE, but could use more
- Evaluate usability by marine biologists: DONE
- Catch up with all previously recorded videos: detection 100%, recognition 50%, est 2 months more processing

Technical Overview

Camera Views



Processing through 30 June 2013.

Video Quality Classification

Algorithm to classify video quality, for detector selection

93% accurate

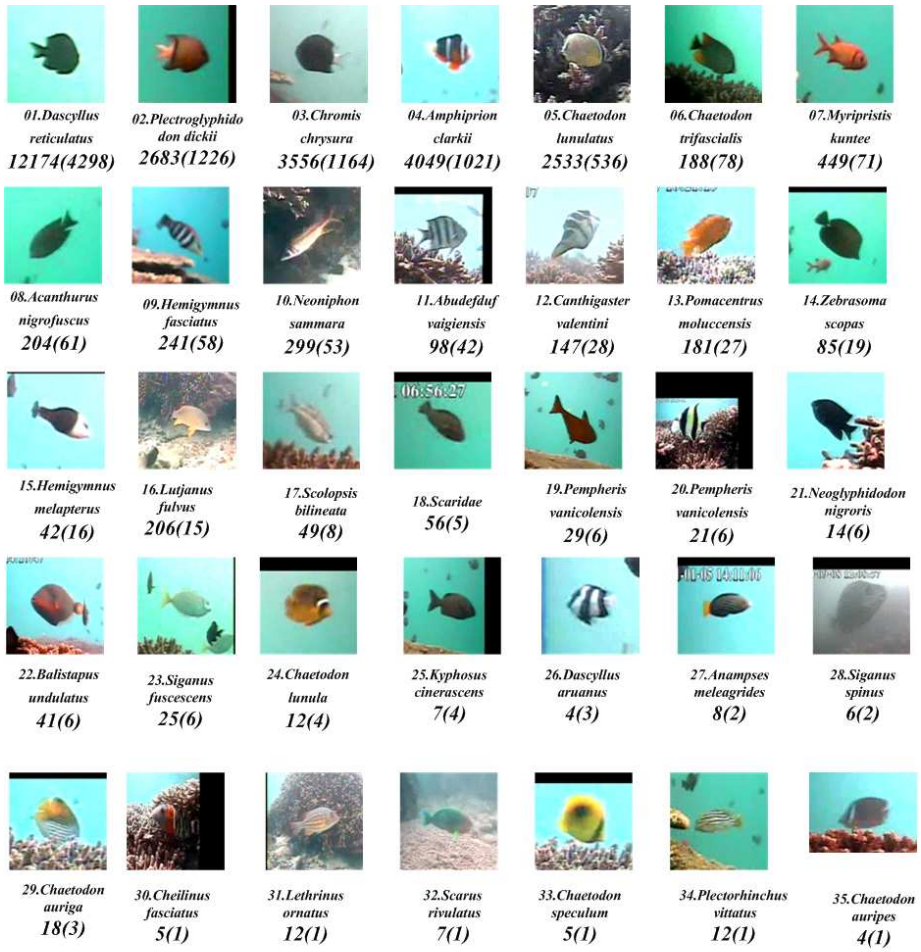
524086 videos, 87K hours:

- Normal (14%)
- Complex Background (7%)
- Algae on Lens (9%)
- Blurred Water (35%)
- Highly Blurred Water (12%)
- Encoding Errors (21%)
- Unknown (1%)

Fish Detection and Tracking

- Difficult environment: algae, moving plants, changing lighting, caustics; small image size 320x240, low frame rate
- But: fixed cameras and background models
- 1.44 billion fish detections, 145 million tracks
- Detection F_1 rate: 0.81
- Frame-to-frame Correct Tracking Decision Rate: 82%

Current Species Ground Truth



35 species 27470 fish (8780 trajectory)

23 Species Recognised

23 most common species (99.7+% of observed fish)

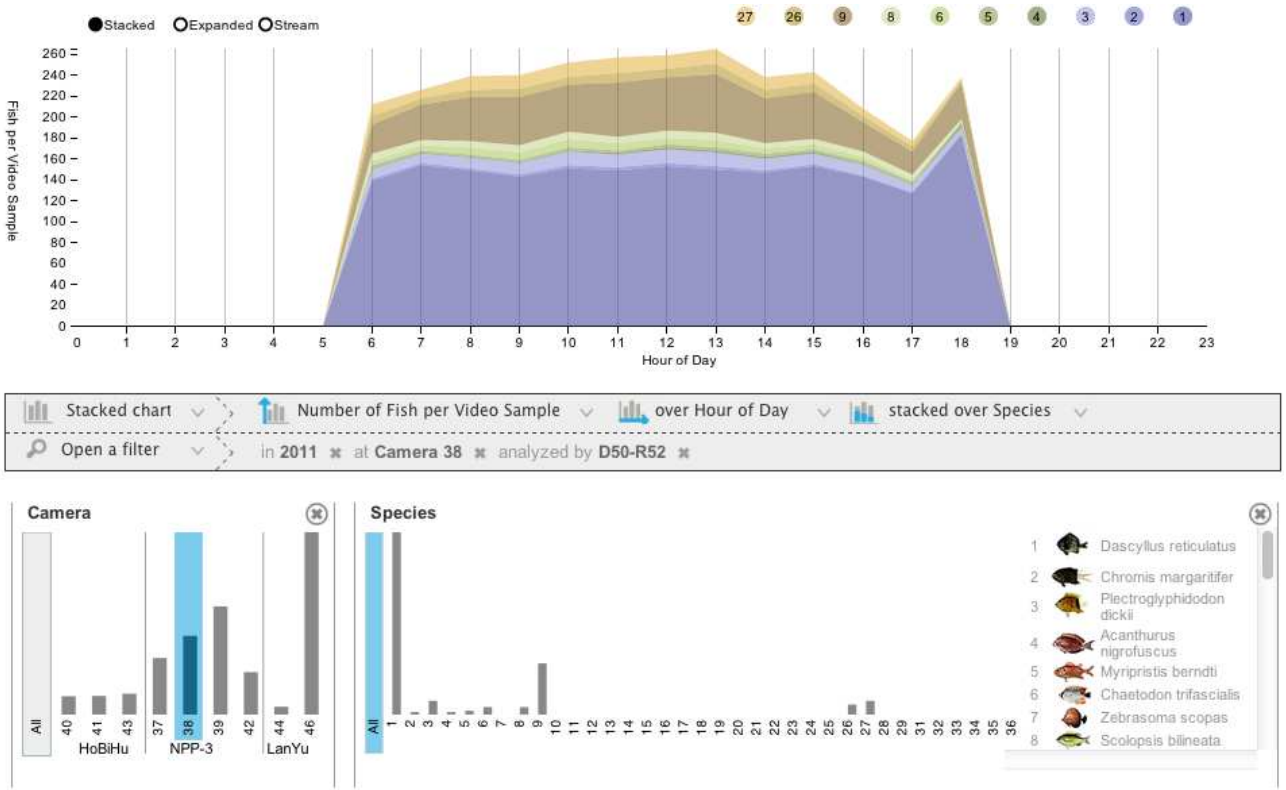
Accuracy on Ground Truth averaged over species: 75%

Accuracy on Ground Truth averaged over fish: 97%

Recognition Processing Summary

Type	Processed	Total	Percent
Algae	18993	49370	38%
Blurred	68711	181965	38%
ComplexScenes	36804	37404	98%
EncodingProblem	39626	108140	37%
HighlyBlurred	27163	65024	42%
Normal	75424	75806	99%
Unknown	6111	6171	99%
TOTAL	272916	524086	52%

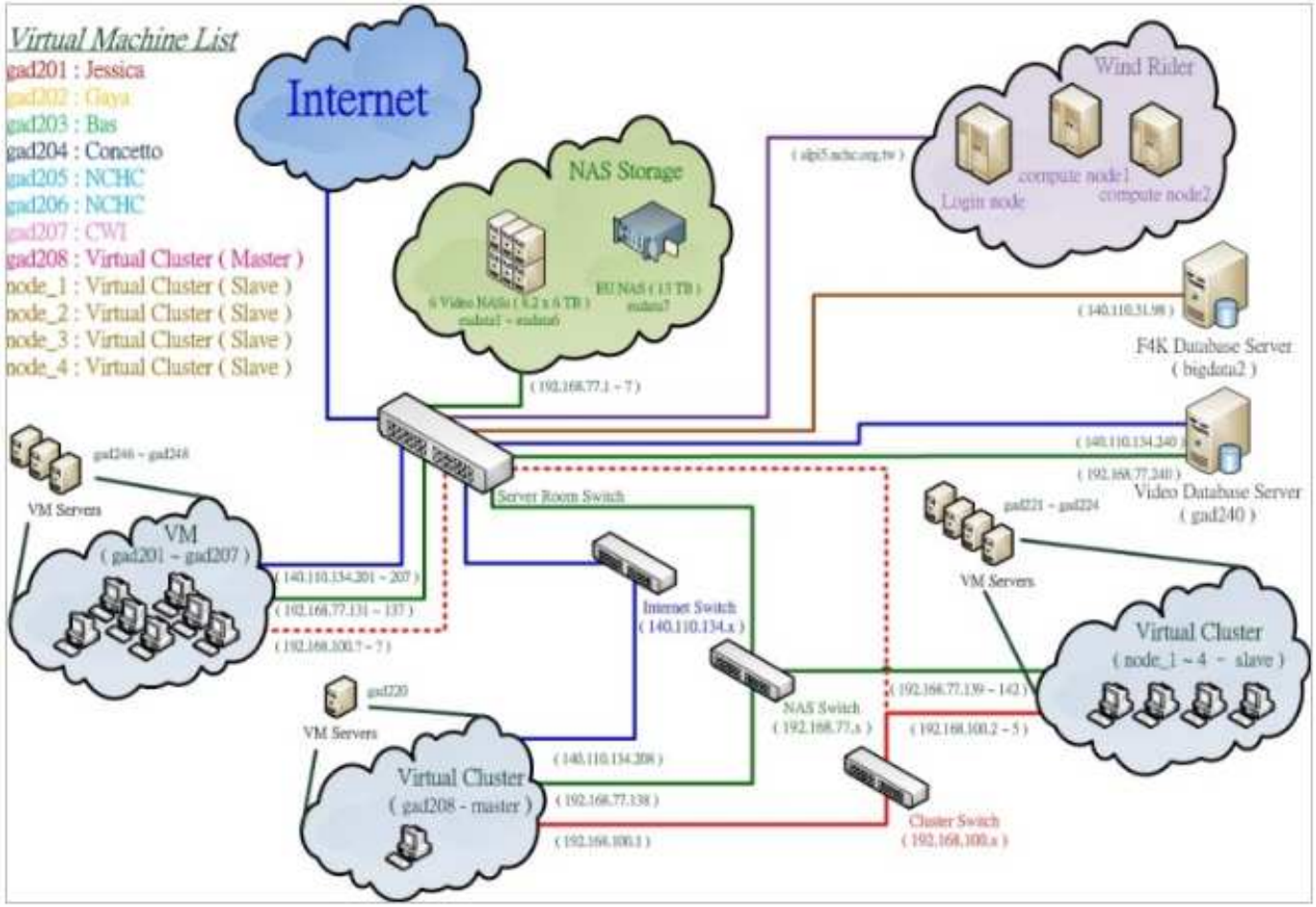
Faceted/Filtered User Interface



96 Dedicated Processor Host + Supercomputer

Also doing 1000+ processor runs

Heterogenous Processor Network

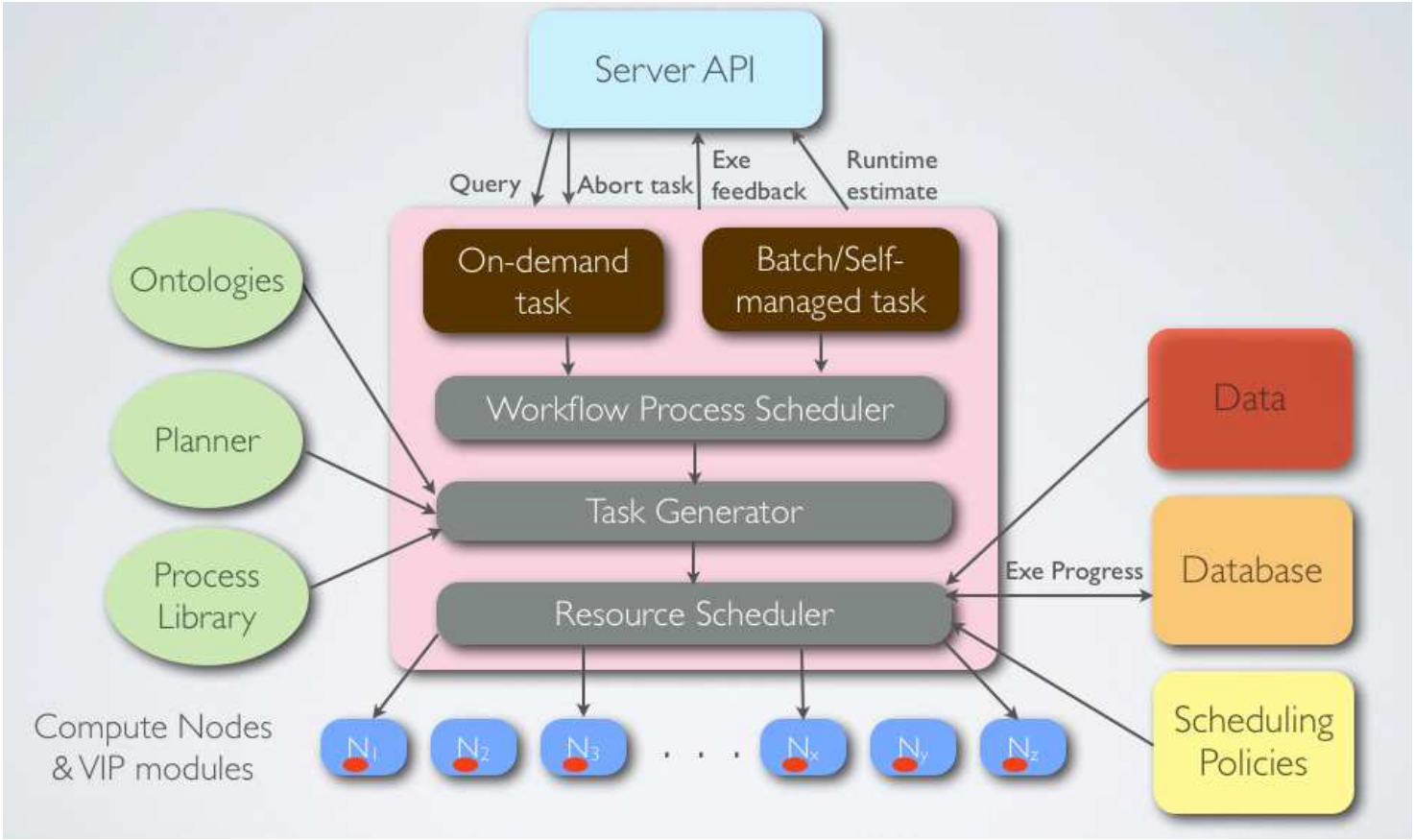


Includes extended NAS storage to 206 Tb video, 400 Gb SQL

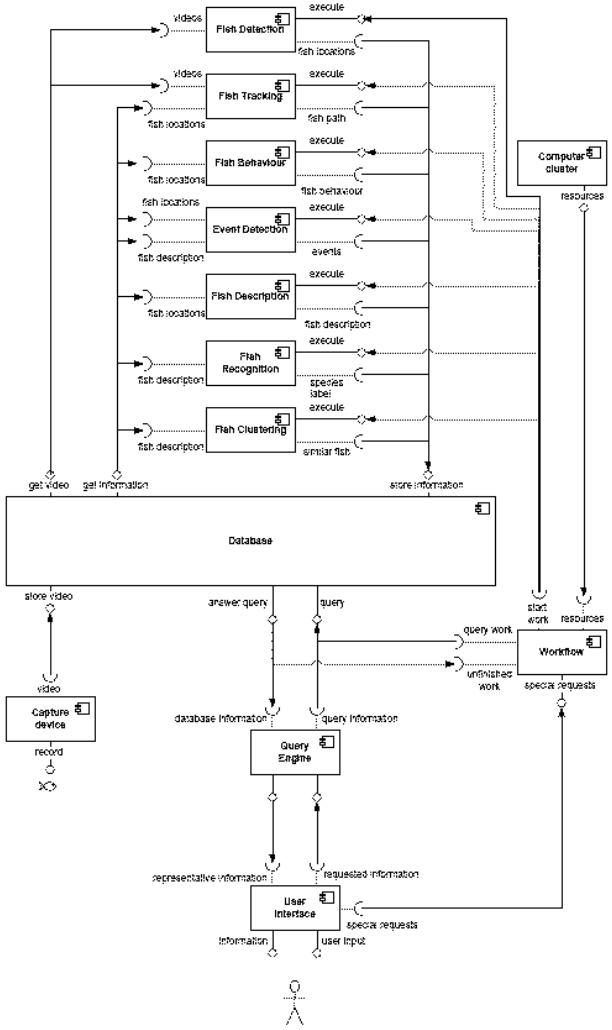
Distributed Process Workflow Control

Heterogenous computing environment

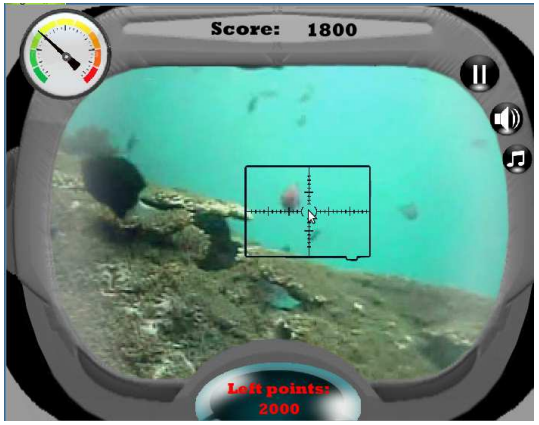
SQL database based monitoring of jobs and queries

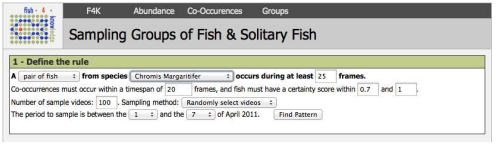


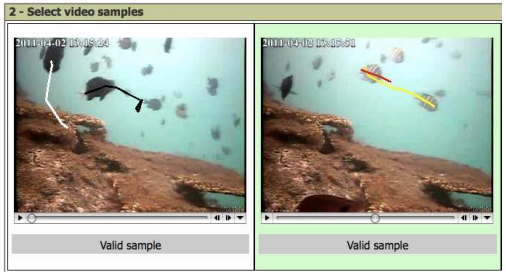
System Integration

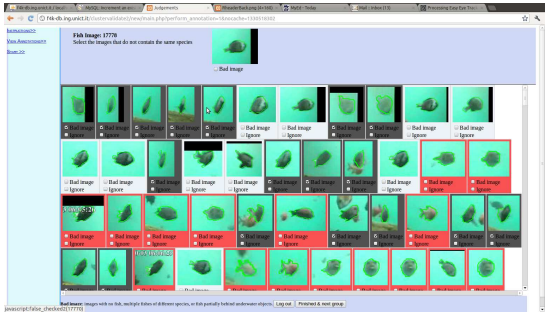


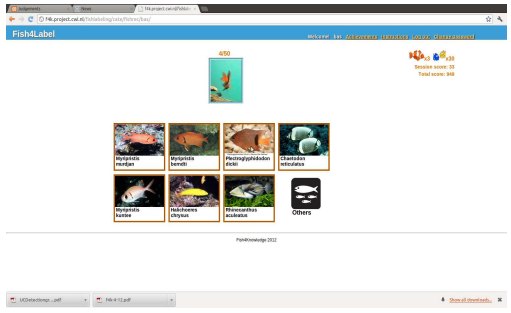
Interfaces for Ground Truth











Dissemination

- Organised scientific workshops and special sessions: 5
- Journal publications: 6
- Conference publications: 41
- Journal special issues: 6
- Invited Talks, Posters and Exhibitions at Scientific Conferences: 15
- Potential collaborations: 10
- MSc theses: 6, PDEng: 1, PhD in progress: 4
- F4K web site: 7800+ hits

Public Available Resources

- Shared source code: SourceForge
- 24 hour and 3 year data samples
- SecondLife Exhibition
- User Interface to full dataset
- Detection and Recognition ground truth
- ImageCLEF dataset

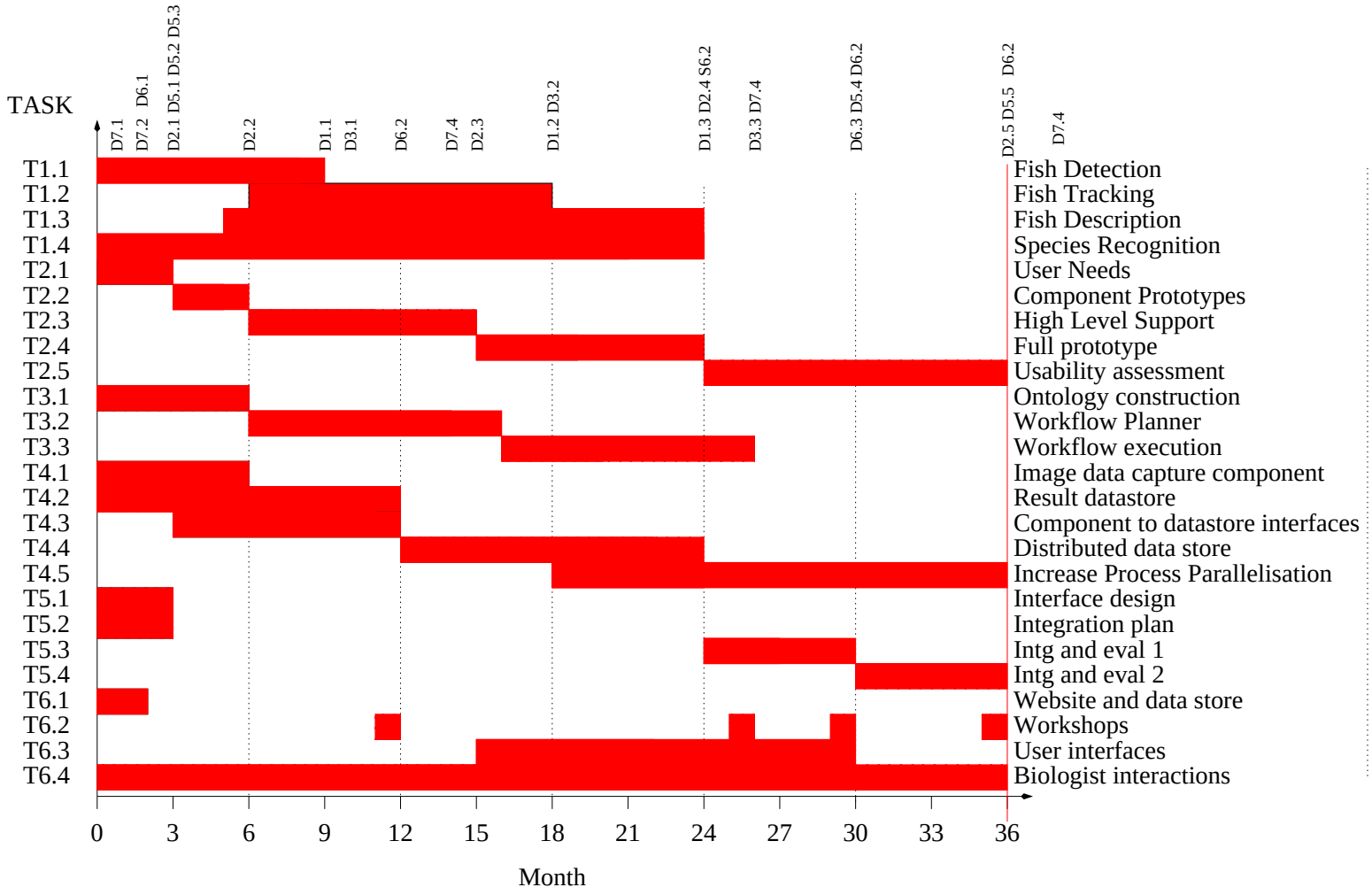
Management

- Person-months Proposed: 261 Actual: 441
- Unspent budget: 201077 Euros

Completed Milestones

All Milestones met

System Progress



Deliverables

All Deliverables complete

Full Consortium Meetings

1. Luxembourg December 2012
2. Taiwan April 2013
3. Catania September 2013
4. Luxembourg November 2013

Additional technical working meetings: Amsterdam, Paris, Taiwan

Post-project Plans

- Finish recognition processing
- Copy all recognition results to project website, for public use
- Fish4Knowledge book
- Collaborations for exploitation, especially Aquacam (Caribbean fish detection, recognition and 3D stereo fish sizes), Walailak (Taiwan) using F4K streaming techology
- Extending methods for evaluating and representing uncertainty.