



Fish4Knowledge Deliverable D7.7

First 6-monthly report to EC

Principal Author: UEDIN
Contributors: All partners
Dissemination: PU

Abstract: This document summarises the project activities in months 1-6.

Deliverable due: Month 6

1 Administrative Progress

The main administrative activities were 2 project consortium meetings and the eventual recruitment of all of the core project researchers. The first meeting was in Edinburgh in November 2010 and consisted of mainly the PIs, because few of the young researchers had been recruited by that point.

The second meeting was in Taiwan in April 2011. The team visited the NARL laboratories at Hsinchu City and Taichung, and the data collection sites at Hobihu Harbour, Kenting and Lanyu Island. Eight researchers from Europe joined 4-8 researchers from the NARL team in the site visits and presentations. Prof. Shao, one of the Scientific Advisory Board, joined us for the main technical meeting day.

Minutes and PDFs of all meeting presentations are available on the project Members Only pages.

The project web site has been updated with the project membership.

There is about a 3 month delay in project progress as a consequence of the rate at which the research staff were hired.

2 Technical Progress

This section summarises the progress made by each of the different teams during project months 1-6.

2.1 CWI team achievements

- Draft list of 20 questions : after initial consultations with local biologists, and some skype/email conversations with Professors Shao and Stergiou, a set of 20 'marine-biology' questions have been identified as a focus for designing the technical capabilities of the project. Results to appear soon in Deliverable D2.1 and D5.3.
- Analysis of 20 questions in progress (needs, resources, results): The 20 questions have been analysed in terms of the data and algorithmic requirements needed to answer the questions. Results to appear soon in Deliverable D2.1.
- Draft mock up of user interface in progress: Sketches illustrating the appearance and functionality of the user interface have been developed. Results to appear soon in D2.2.
- Review of dataspace for data integration: preliminary sketches for this have been developed, but further work on this is needed. UEDIN is currently discussing designs with their Database research group.
- Review of data searching tasks: examples of these were presented, and will eventually appear in D5.2.
- Search and exploration interface for agrovoc and FAO fish ontologies: these have been assessed and developed.

2.2 NARL team achievements

- The team has maintained and enhanced the three underwater data capture systems at Kenting, Hobihu and Lanyu Island, resulting in 11 operational cameras. The improvements include adding hybrid coaxial cable/optical fibre to allowing longer distances from the cameras to observatory units.
- An internet accessible dedicated 64 processor cluster and tera-scale storage system is nearing completion. This will be the project's main computational resource. On top of this, a virtual machine method has been developed to allow different teams to run their own software components more conveniently.
- A hierarchical video cache system has been developed to improve web-based data access to the historical videos (0.01 sec access time).
- Prepared arrangements for the second F4K workshop at Taiwan, including hotels, meeting rooms and travel arrangements (train, bus, boat, taxi) for visiting the three data capture sites.

Here is the team after visiting the data capture site at Kenting's National Museum of Marine Biology and Aquarium:



2.3 UCATANIA team achievements

- The UCATANIA team has developed three new fish detection algorithms for use in varying circumstances (luminosity, non-stationary background, adaptive). Examples of the detected fish can be seen at:

<http://homepages.inf.ed.ac.uk/rbf/Fish4Knowledge/results.htm>.

- They are adapting a covariance matrix based tracking algorithm to include shape and texture as well as colour. An example of the tracking results superimposed on the background image is:



- They are investigating an algorithm for fish schooling and group detection.
- Using their algorithms on the Taiwan datasets, we have an initial dataset with about 10^6 detected and tracked fish. This is available in a SQL-family database at: <http://151.97.9.184/phpMyAdmin/>, (Username: f4k, Password: f4kpwd) which can also be accessed using a C++ interface.
- A C++ interface to an **Entity-Relation Model** database has been developed.
- Detection and tracking ground truth has been developed 80 minutes of videos: **8 videos * 10 min at 10 fps from different sites gives about 40K frames, 50K detections over 6K fish, using VIPER**. This will be used for detection and tracking ground truthing.

2.4 UEDIN workflow team achievements

- The team is investigating new research directions for workflow usage within F4K: parallelism, massive and dynamic data. In particular, the team is investigating how to a) construct workflows automatically in response to non-standard user queries and b) dynamically allocate processing tasks across the 64 processors and different task components (for both normal data stream and query processing).
- A Java GUI and demo of Semantics-based Workflow system for Automatic Video Analysis (SWAV) has been developed. This is the fish detection workflow system developed in Nadarajan's PhD research that was a prototype that led to the Fish4Knowledge project.

- A Second Life exhibition gallery has been started, now including media screens & a “talking fish” demo. It can be seen at:
[http://maps.secondlife.com/secondlife/...
Edinburgh%20University/70/198/26](http://maps.secondlife.com/secondlife/...Edinburgh%20University/70/198/26).
The team is currently investigating sample questions and answers targeted towards 13-17 year audience for the main gallery and extending the upstairs to create a more academic space showing project results.
- The team is organising special session at 5th Int. KES Conf. on Agents and Multi-agent Systems, Technologies and Applications (KES-AMSTA), Manchester, June 29-July 1, 2011: <http://amsta-11.kesinternational.org/>. The Special Session is on Intelligent Workflow, Cloud Computing and Systems (IS03), Friday 1:00 - 3:00, chaired by: Yun-Heh (Jessica) Chen-Burger, Ching-Long Yeh and Fang-Pang Lin, and will have 5 papers.
- The team has initiated a project wiki for recording project working details at:
<https://wiki.inf.ed.ac.uk/F4K/>.

2.5 UEDIN vision team achievements

- Completing a literature review of fish recognition: this will form part of Xuan Huang’s (project research staff) PhD thesis proposal.
- Reasonably clean list of Taiwan fish species with URNs, image samples: we have extracted the fish from the Fish Database of Taiwan website, removed fish unlikely to be seen (freshwater, deep sea, too small) and found a unique URN for each (mainly by cross-referencing with the Catalog of Life). We have located several example images for most species, to be used as initial training data for species recognition. Of the initial 2894 species, about 1500 species remain after pruning. The web site that summarises this data is at: [http://homepages.inf.ed.ac.uk/bboom/...
species_urn_catalogue_of_life.html](http://homepages.inf.ed.ac.uk/bboom/...species_urn_catalogue_of_life.html).
- Top level project integration plan: through consultation with all teams, a top level integration plan has been developed, identifying the components, dataflows and responsible teams. This is documented in deliverable D5.1. What the team are now developing is a specific set of program calling interfaces and data file/structure conventions.
- Effective web site: The project has developed a project web site at: fish4knowledge.eu, which is being updated as new information is acquired.

3 Plan for technical working meetings

We have identified a number of technical working meetings that are to take place.

- CWI + UCATANIA + UEDIN vision: database content plan (held April 2011 in Milan)
- UEDIN vision + UCATANIA: visual description + data processes (held June 2011 in Edinburgh)

- NCHC + UCATANIA + CWI + UEDIN workflow: visual data processing modularisation + HPC workflow **early year 2 - eg. Dec 2011**
- NCHC + CWI: HPC query interfacing: **early year 3**
- ALL: integration and testing: **early year 3**

4 Deliverable Summary

Below is a summary of the deliverables that are due in the first 6 months. Most are approaching completion, but have been delayed by about 3 months due to the lag in recruitment of the project researchers.

Each Deliverable document, dataset and software component now has an identified person from another project partner who is responsible for reviewing the content and performance of the deliverable. The Quality Control person is identified in the table below.

Num	Team	Title	Mth	Done	Quality Control
D7.1	UEDIN	Consortium Agreement	1	Y	all
D7.2	UEDIN	Organisation and minutes of Project Start-Up Meeting	1	Y	all
D7.3	UEDIN	Project Fact Sheet	1		Jacco
D6.1a	UEDIN	Project web site	2	Y	Emma
D6.1b	NCHC	Project data repository (merge with D4.1)	2		Bas
D6.7	UEDIN	Public Press Release	2		Jessica
D7.10	UEDIN	First annual public report	2	Y	-
D2.1	CWI	User information needs	3		F-P Lin
D5.1	UEDIN	Component Interface and Integration Plan	3	Y	All
D5.2	CWI +UCAT	Datastore Definition	3		UEDIN
D5.3	CWI	Scientific Question and Exp't Plan	3	close	Bob
D2.2	CWI	Identified user scenarios and impl plan & UI design	6	in prog	Jessica
D7.7	UEDIN	First 6-monthly report to EC	6	Y	All PIs