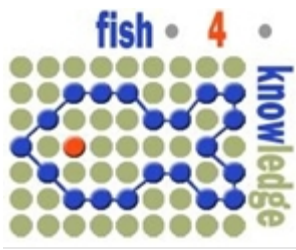# Fish4Knowledge Annual Report

www.fish4knowledge.eu

The Fish4Knowledge project will investigate: information abstraction and storage methods for reducing the massive amount of video data (from 10E+15 pixels to 10E+12 units of information), machine and human vocabularies for describing fish, flexible process architectures to process the data and scientific queries and effective specialised user query interfaces. A combination of computer vision, database storage, workflow and human computer interaction methods will be used to achieve this. Data from 10+ undersea cameras will populate the massive database which will in turn support structured queries by the marine biology user community.

Summary of Activities

The project started October 2010, and we have successfully assembled the team, including recruitment of most staff and held our kick-off meeting. A project web site has been developed, listing the project members and key responsibilities. A collection of resources is starting to be developed, including:

- Examples of the image data from some of the undersea cameras: **http://ecosite.nchc.org.tw/kenting/**

- A collection of background publications

On a more technical note, we have:

- Extracted the 3000+ fish from the Fish Database of Taiwan and linked most of them so far to a URN from the Catalogue of Life resource, thus providing a unique designator for each species.

- Sketched out the major components of project and assigned responsibilities to different teams.

- Started design of the structure of the datastore.

- Have a preliminary fish detection process based on background research. This process will provide data for starting the development of the tracking, recognition and user query processes.

- Are currently developing a 'system design' for the project that will allow all teams to more easily integrate their components into an integrated demonstration system.

Unique Fish Species Nomenclature

The project needed a vocabulary for referring to the different fish species, genera, families, orders and classes, and the vocabulary should be standard, recognizable and computer usable. We decided that the URN (Uniform Resource Name) style should be used. This has required establishing a URN for each species known to be present in Taiwan, and hopefully using a previously established URN. Fortunately, the Catalogue of Life has these, which meant that we mainly needed to look up the URNs. As there are ambiguities and synonyms, as well as species names not in the CoL, this causes problems, but we currently have standard URNs for about 90% of the 3000+ Taiwan species.

Preliminary Fish Detection

A preliminary fish detection process has been developed by combining two previous techniques to give more reliable fish detection, as each algorithm solves some problems of the other. The outputs of the two techniques are combined and further filtered to remove small fish and noise. Examples of detected fish from one of Taiwan underwater cameras are shown here:
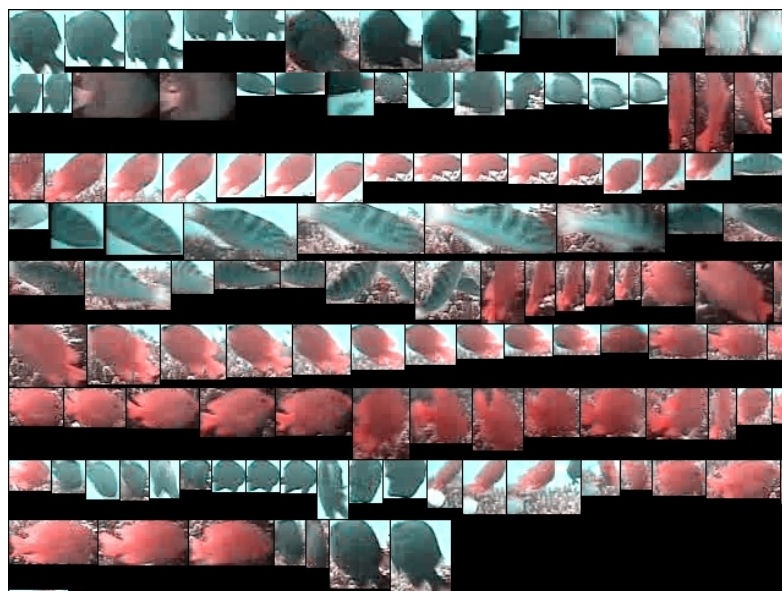
User Involvement, Promotion and Awareness

We have:

- Introduced the project to the Scientific Advisory Board, and engaged them on a few issues

- Developed a draft Second Life public portal to the project:
  http://maps.secondlife.com/secondlife/Edinburgh%20University/70/198/26

- Made arrangements for the first of 4 technical seminars hosted by the project, to take place this summer at: The Intelligent Workflow, Cloud Computing and Systems Invited Session at the 5th International KES Symposium on Agents and Multi-Agent Systems – Technologies and Applications Manchester, UK, 29 June - 1 July 2011.

Future Work

As the project is just starting, our main activity is developing shared ideas on the types of marine biology query that can be supported, and what the underlying technology is that will enable the answering of the queries. This includes visits to marine biologists and, in April 2011, a visit to the data capture sites in Taiwan.