

Data Harvesting - Machine to machine

Lotta Fyrberg, SMHI (Sweden), Lotta.Fyrberg@smhi.se
Nils Nexelius, SMHI (Sweden), Nils.Nexelius@smhi.se
Arnold Andreasson, SMHI (Sweden), Arnold.Andreasson@smhi.se
Lisa Sundqvist, SMHI (Sweden), Lisa.Sundqvist@smhi.se

Users of data will always need access to highest quality, largest collection, and most recent version of data. Advances in technology can meet these demands via technical solutions. One way forward is to grant open access using machine to machine interfaces. SMHI has developed the SHARKdata.se system to handle these data flows. The data are currently being harvested in the DwC-A format, by portals such as EMODnet-Biology (and hence to EurOBIS/OBIS), and by the Swedish LifeWatch system.

There is now an ongoing study in co-operation with ICES Data centre. The aim of this study is to automate the national reporting of biological data (phytoplankton, zooplankton, phytobenthos and zoobenthos) from OSPAR and HELCOM regions, to the Marine Environmental data portal (DOME).

This will result in the following:

- I. Reduce the amount of manual work (and human errors)
- II. Ensure the most recent version of data in DOME
- III. Automated quality controls through DATSU REST API on data.
- IV. Higher quality of data

All technology used is open source (MIT license) and hence open for everyone to download and build their own system. For data users it is possible to use R and Python to set up any type of analysis on data from the Swedish National Oceanographic Data Centre. Examples are published on the server SHARKdata.se