Use of XML technology in the Baltic Sea Fishery Database

P. Sandbeck

DIFRES, Danish Institute for Fisheries Research
Jægersborgvej 64, 2800 Lyngby, Denmark

E-mail: pes@dfu.min.dk

INTRODUCTION

The assessment of the sizes and composition of the fish stocks in the Baltic Sea is coordinated by ICES. In order to make fish catch data from the different countries available for all partners, a common Internet based database was designed and implemented by DIFRES. With this application the users can upload data to the database via the Internet, and data can also be requested from the database and received in this way. The structure of the tables in the database is the same as the column separated ICES formats normally used to exchange fishery data. However for several reasons it was decided to use XML formats for the exchange of data. The benefits of this are described in this presentation.

XML SCHEMAS

The structure of the XML file is defined in a XML schema file, which is a XML file itself. When data are uploaded to the database, the XML file is validated against the XML schema. In the XML schema it is possible to check for the order and occurrence of record types, the data types (integer, real, date), minimum and maximum values and allowed enumeration data. From the database website the user can download a program, which converts the data from ICES formats to XML formats according to the defined XML schema. By using a XML aware editor the user can instantly check the validity of the data before uploading. The validity of the file can also be checked by a downloaded program. Hence the XML schema is both a specification of the exchange format and a data validation program, which the data from all partners must apply to.

XML AND DATA HANDLING

Since XML is a standard, the Baltic Sea fishery database system and its web program can use XML directly. The data records in the XML file are stored in the database without much programming. When the users download requested data, the data can easily be converted to e.g. spreadsheets for analysis. By using the XPath and XSLT (XML Style sheet Language Transform) facilities, which are part of the XML standard, it is easy to pass a XML file and use the data elements in a custom made program. In the Baltic Sea fishery database XSLT is used to pass the uploaded XML files in order to perform further data validation checks and to log some information about the uploaded data. When the users want to browse the database, XML is used to build a tree structure of the data in the left frame on the HTML page. When the user clicks on a data node in the tree, e.g. a year or a cruise, the tree expands and the selected data record is shown in the right frame. By using XSLT it is also possible to make a dynamic HTML form based on the XML schema.

THE PRESENTATION

The presentation will focus on the use of XML and XML schemas in the Baltic Sea Fishery database. The structure of the XML schema and the built in configuration of data validation will be shown. The uploading of data to the database and the user interface with the XML tree structure will also be demonstrated.