INFRAESTRUTURA DE DADOS OCEANOGRÁFICOS - INDO / FASE 1

Matheus Artur Zanella[1], Rafael Medeiros Sperb[2], Luiz Cláudio Dalmolin[3], Nilson Ribeiro Modro[4]

[1] Acadêmico do Curso de Bacharelado em Sistemas de Informação CEPLAN- PROIP/UDESC
[2] Professor do Instituto de Oceanografia, Universidade Federal do Rio Grande - FURG - rsperb@furg.br
[3] Professor, Departamento de Sistemas de Informação - CEPLAN – luiz.dalmolin@udesc.br
[4] Orientador, Departamento de Sistemas de Informação - CEPLAN – nilson.modro@udesc.br

Palavras-chave: Dados Oceanográficos, netCDF, Procedimento Operacional Padrão.

Há uma crescente demanda para definição de políticas de gestão de dados oceanográficos, no Brasil. Seja ela oriunda dos acordos internacionais dos quais o Governo Brasileiro é signatário; seja em virtude da inexistência ou precariedade com que os dados vêm sendo armazenados. Experiências internacionais, realizadas em infraestruturas de dados oceanográficos, indicam que o emprego do *Network Common Data Form* (*netCDF*) como meio de sistematização, armazenamento e consulta à dados oceanográficos se mostra como uma forte tendência. Principalmente quando se considera interoperabilidade e compartilhamento de dados. Assim, a problemática abordada neste projeto consiste em conhecer os padrões, as tecnologias e as boas práticas para sua adoção, ao passo que se determina as dificuldades operacionais para sua adoção.

Neste contexto, este projeto que teve vigência entre agosto de 2017 a julho de 2019, objetivou levantar as dificuldades operacionais a serem enfrentadas na adoção do *netCDF* como formato padrão para sistematização, armazenamento e consulta à dados oceanográficos junto à Infraestrutura de Dados Oceanográficos do Instituto de Oceanografia da Universidade Federal do Rio Grande (IO/FURG). O tema, fruto da colaboração entre o Laboratório de Sistemas de Informações Geográficas - LABSIG/UDESC e o Laboratório de Gerenciamento Costeiros - LABGERCO/FURG, abrange o levantamento dos protocolos e padrões internacionais para manipulação e armazenamento de dados oceanográficos, o estudo do conjunto de bibliotecas e formatos de dados, e dos pacotes de manipulação distribuídos para as linguagens Java, C/C++ e Fortran, desenvolvidos e mantidos como padrão pela UNIDATA. Finalmente, conjuntos de dados oceanográficos foram encapsulados como *netCDF*.

O projeto teve as seguintes macro atividades:

1. Estudo de protocolos e padrões internacionais para manipulação e armazenamento de dados oceanográficos;
2. Estudo de conjunto de bibliotecas e formatos de dados, dos pacotes de manipulação distribuídos para as linguagens Phyton, Java, C/C++ e Fortran desenvolvidos e mantidos como padrão pela UNIDATA;
3. Manipulação de dados: espelhou-se um servidor de dados do LABGERGO/UFRG no LABSIG/UDESC. Foi realizada uma carga amostral de 150 GB. A amostra contém os metadados de dados físicos da Lagoa dos Patos (RS), seguindo o perfil de Metadados DwC (Darwin Core) e possui os seguintes campos: a) medida de temperatura e salinidade; b) data e profundidade da amostragem; c) metadados em .xml. Também foram utilizados os dados das boias de La Parguera – Porto Rico e Efin Cave – Canadá, disponibilizados pela NOAA - *National Oceanic & Atmospheric Administration*.

4. Testes de encapsulamento de conjuntos de dados oceanográficos utilizando as seguintes tecnologias e plataformas: PyCharm, *Numpy*, *NetCDF4 e Pandas*;

5. Mapeamento dos passos necessários para boas práticas de armazenamento de dados de dados Oceanográficos em formato *NetCDF* e criação do Procedimento Operacional Padrão – POP (Figura 1).

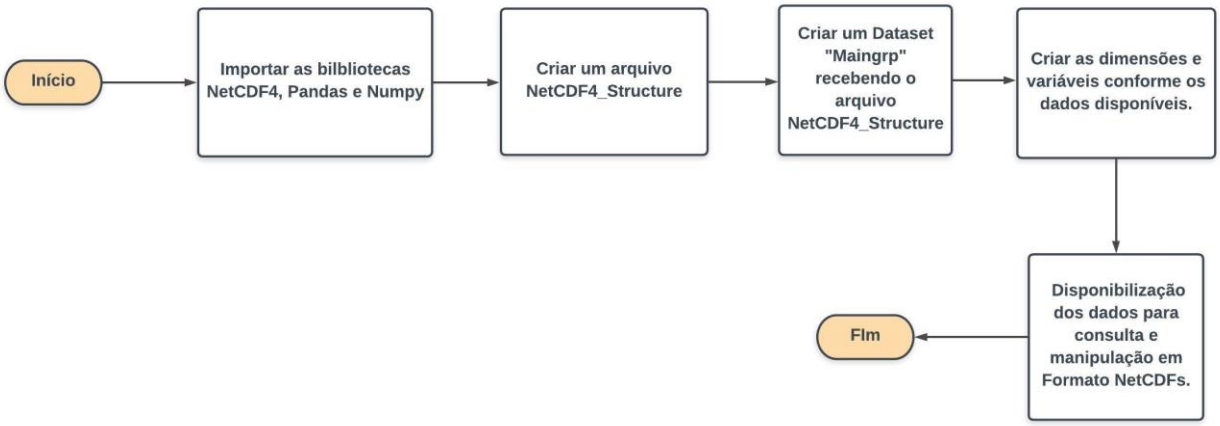

Figura 1 – Boas práticas de armazenamento de dados oceanográficos em formato NetCDF