



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

شیوه نامه استاندارد سازی داده ها، فراداده ها و گزارش گشت های دریایی

پژوهشگاه ملی اقیانوس شناسی و علوم جوی
مرکز ملی داده های اقیانوسی و دریایی

تاریخ آخرین ویرایش: آبان ۱۳۹۸

گشت های دریایی

مقدمه:

مرکز ملی داده‌های اقیانوسی و دریایی^۱، داده‌های مختلفی را از منابع متفاوت جمع‌آوری می‌کند. منابع این داده‌ها از مراکز داده بین‌المللی، داده‌های حاصل از طرح‌های پژوهشی و گشت های دریایی پژوهشگاه می‌باشند. یکی از قدم‌های مهم در فرایند مدیریت داده‌های اقیانوسی، جستجوی داده، استخراج و انتقال داده‌ها به متقاضیان داده می‌باشد؛ که برای تحقق این امر، نیازمند به مجموعه داده‌هایی یکپارچه و با ساختاری مشابه می‌باشیم. از آنجایی که داده‌ها از منابعی متفاوت دریافت می‌شود، مرکز داده‌ها با ساختاری متفاوت از داده‌ها مواجه است. گوناگونی در ساختار داده، به‌طور بالقوه‌ای زمینه‌ساز بسیاری از مشکلات در زمینه کشف و استخراج داده می‌شود. جهت مدیریت این امر، ایجاد رویه‌ای جهت استانداردسازی داده‌های اقیانوسی و فراداده‌های حاصل از آن‌ها ضروری می‌باشد. ارزش استانداردسازی کاملاً شفاف و واضح است. این رویه به ایجاد مجموعه داده‌های یکپارچه، استاندارد و باکیفیت کمک می‌کند. کیفیت، سازگاری و انسجام داده‌ای که از منابع مختلف صادر می‌شود، با پذیرش متدولوژی‌های استاندارد تضمین می‌شود. به فرایند دستیابی به توافق بر روی تعاریف مشترک داده‌ها، چارچوب‌ها، اشکال نمایش و ساختارهای همه لایه‌های داده‌ای و عناصر داده‌ای، "استانداردسازی داده" اطلاق می‌شود. استانداردسازی داده‌ها نخستین گام در تحقق بخشیدن به مقوله به اشتراک‌گذاری داده‌ها می‌باشد. یک استاندارد داده، قادر است داده‌های مورد نظر را به شکلی منسجم و واحد، تحت چارچوب‌ها و تعاریفی که مورد توافق همگان قرار دارد، به نمایش گذاشته و از این طریق سازگاری را افزایش، تکرار را کاهش و کارایی تبادل و گردآوری داده‌ها را بهبود بخشد.

در این راستا برنامه بین‌المللی تبادل داده‌ها و اطلاعات اقیانوسی (IODE^۲) به کمک کمیسیون فنی مشترک اقیانوس‌شناسی و هواشناسی دریایی (JCOMM^۳) اقدام به تهیه برخی از استانداردهای داده‌های اقیانوسی نموده است که استفاده از آن‌ها در نشست‌های مختلف برنامه IODE به مراکز داده توصیه شده است. در همین راستا، مرکز ملی داده‌ها، با استفاده از این استانداردها که خود طیف وسیعی از مواردی چون نشانه‌های کنترل کیفیت، تاریخ و زمان، کد کشورها، مدل فراداده‌ها و واژگان مشترک^۴ دریایی را شامل می‌شود، اقدام به ایجاد رویه‌ای جهت استانداردسازی داده‌ها، فراداده‌ها و گزارش گشت های دریایی در مرکز داده‌ها پرداخته است. جهت تسهیل در به‌کارگیری و اعمال موارد مطرح شده که باید توسط مجریان طرح‌های پژوهشی و محققین ارشد گشت های دریایی انجام شود، مرکز ملی داده‌ها اقدام به تهیه این شیوه نامه در پنج ماده نموده است.

ماده ۱: تعاریف

در مبحث مدیریت داده‌های اقیانوسی، تعاریف زیر برای واژه‌های داده، اطلاعات، فراداده و گزارش گشت دریایی در نظر گرفته می‌شود.

¹ Iranian National Center for Ocean Data

² International Oceanographic Data and Information Exchange

³ Joint Technical Commission for Oceanography and Marine Meteorology

⁴ Common vocabularies

داده^۵: داده‌ها متشکل از مشاهدات اقیانوس شناسی، داده های استنتاج شده و شبکه ای (از طریق اینترپوله کردن بدست می آیند مانند مشاهدات هواشناسی) هستند. این مقادیر می توانند عددی^۶ باشند مانند درجه حرارت و شوری و یا اسمی^۷ باشند مانند لیست گونه‌ها در یک منطقه خاص.

اطلاعات^۸: به داده‌ای که پردازش یا تفسیر شده باشد، اطلاعات اطلاق می‌شود.

فراداده^۹: فراداده‌ها، داده‌ای در مورد داده هستند و قادرند اطلاعات مفیدی را در رابطه با یک دسته داده در اختیار گذارند؛ مثلاً اینکه داده مورد نظر در چه بازه زمانی و در کدام محدوده جغرافیایی برداشت شده است و اصولاً شامل چه پارامترهایی می‌باشد. فراداده ها نوعی خاص از اطلاعات هستند.

گزارش گشت دریایی^{۱۰}: این گزارشات به نوعی فراداده گشت های دریایی هستند و اطلاعات گشت های دریایی را توصیف می کنند.

ماده ۲: فایل های که باید به مرکز داده ها تحویل داده شود

در هنگام تسویه حساب طرح‌های پژوهشی می‌بایست فایل‌های داده به همراه فایل فراداده توسط مجری طرح به مرکز داده‌ها تحویل داده شود. اگر داده تولید شده، حاصل از گشت های برگزار شده توسط پژوهشگاه باشد می بایست پس از خاتمه هر گشت، فایل گزارش گشت دریایی نیز علاوه بر داده و فراداده توسط محقق ارشد گشت به مرکز داده ها تحویل داده شود. کلیه فایل ها می بایست به آدرس ایمیل incod_data@inio.ac.ir ارسال شود.

ماده ۳: برنامه کاربردی تولید فراداده

الزامات زیر می‌بایست در فایل‌های فراداده در نظر گرفته شود:

الف- برنامه کاربردی تولید فراداده، از پورتال مرکز داده‌ها به آدرس زیر دانلود شود. کاربر بایستی فایل دانلود شده را از حالت فشرده خارج کرده و سپس فایل اجرایی `MetadataTool.exe` را اجرا نماید. توصیه می‌شود قبل از دانلود و اجرای برنامه کاربردی، به موارد مطرح شده در همین آدرس توجه شود.

<http://incod.inio.ac.ir/Products/Software/Metadata-Tool>

ب- فیلدهایی که با ستاره قرمز مشخص شده اند، فیلدهای اجباری هستند.

ج- خروجی رکوردهای فراداده تولید شده می بایست به صورت فایلی با فرمت XML در اختیار مرکز داده ها قرار داده شود.

⁵ Data

⁶ Numerical

⁷ Nominal

⁸ Information

⁹ Metadata

¹⁰ Cruise Summary Report (CSR)

د- ترجیحا تکنیک های کالیبراسیون و پردازش داده همچنین جزئیات تجهیزات و سنسورها در قسمت کنترل کیفیت¹¹ فراداده، ذکر شود.

ه- لازم بذکر است، برنامه کاربردی اشاره شده در ماده ۳-الف در تاریخ ۹۵/۸/۴ توسط مرکز داده ها به اعضای محترم هیئت علمی و کارشناسان پژوهشی، آموزش داده شده است.

ماده ۴: برنامه کاربردی تولید گزارش گشت دریایی

الزامات زیر می‌بایست در فایل‌های گزارش گشت دریایی در نظر گرفته شود:

الف- برنامه کاربردی تولید گزارش گشت دریایی، از پورتال مرکز داده‌ها به آدرس زیر دانلود شود. کاربر بایستی فایل دانلود شده را از حالت فشرده خارج کرده و سپس فایل اجرایی CSRTool.exe را اجرا نماید. توصیه می‌شود قبل از دانلود و اجرای برنامه کاربردی، به موارد مطرح شده در همین آدرس توجه شود.

<http://incod.inio.ac.ir/Products/Software/CSR-Tool>

ب- فیلدهایی که با ستاره قرمز مشخص شده اند، فیلدهای اجباری هستند.

ج- خروجی رکوردهای گزارش گشت دریایی تولید شده می‌بایست به صورت فایلی با فرمت XML در اختیار مرکز داده ها قرار داده شود.

ماده ۵: استانداردسازی داده ها

الزامات زیر می‌بایست در فایل‌های داده در نظر گرفته شود:

الف- همه فایل‌های داده باید شامل ستون‌های عرض جغرافیایی، طول جغرافیایی، تاریخ، زمان و پارامتر (های) اندازه‌گیری شده باشند.

ب- برای نام‌گذاری ستون‌های ذکر شده در ماده ۵-الف باید از اسامی انگلیسی مطابق جدول ۱ استفاده شود:

جدول ۱: عناوین ستون‌های اجباری

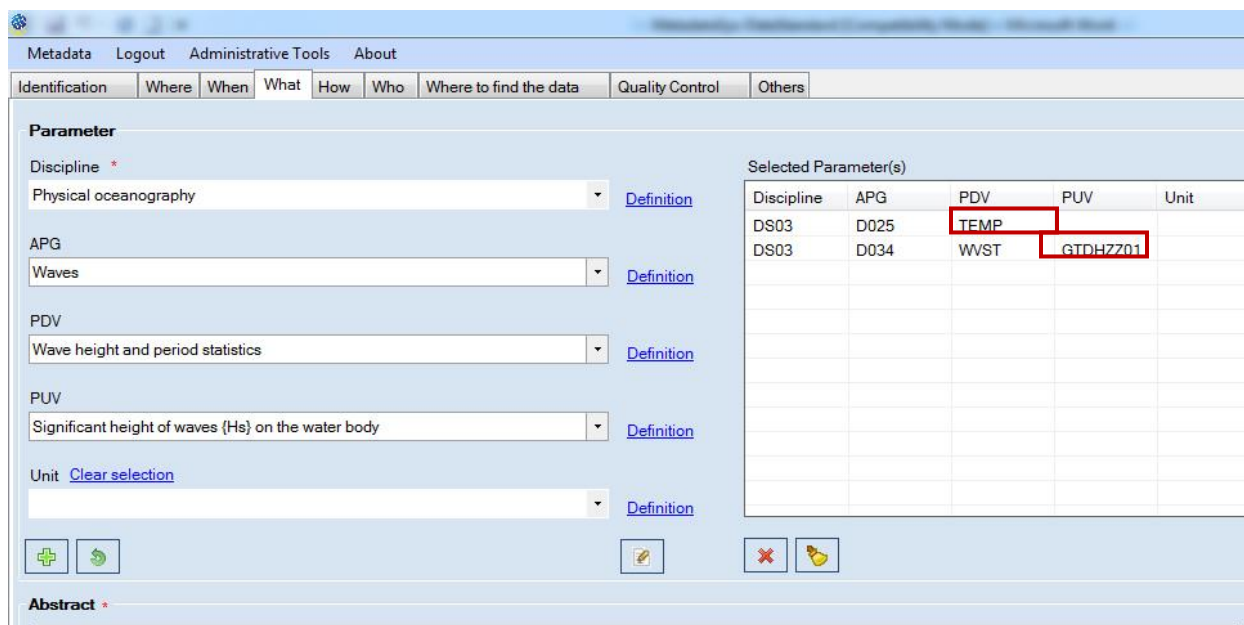
عرض جغرافیایی	Latitude
طول جغرافیایی	Longitude
تاریخ	Date
زمان	Time

پ- تاریخ باید به صورت میلادی و به فرمت YYYY/MM/DD ایجاد شود مانند 2017/02/17.

ت- زمان (در صورت وجود) باید به فرمت hh:mm:ss ایجاد شود مانند 23:05:17 یا 08:15:03.

¹¹ Quality control

- ث- طول و عرض جغرافیایی باید به فرمت درجه اعشاری^{۱۲} باشد برای مثال 63.84. از ارائه مقادیر، به صورت درجه-دقیقه-ثانیه خودداری شود. به عنوان مثال مقادیری مانند "63° 28' 48" یا اشکال مشابه، قابل قبول نیست.
- ج- فایل‌های داده باید ماهیت داده‌ای (ساختار رسمی جداول) داشته باشند.
- چ- ترجیحاً از فرمت ASCII استفاده شود.
- ح- از کلمات فارسی، در فایل‌های داده استفاده نشود.
- خ- تعداد و اسامی پارامترها در فایل داده باید با تعداد و اسامی پارامترها در رکورد فراداده، همخوانی داشته باشد.
- د- اسامی پارامترها می بایست استاندارد شوند. برای نام‌گذاری اسامی پارامترها در فایل‌های داده، از آخرین کد انتخاب شده در طبقه‌بندی پارامترها در برنامه کاربردی فراداده مطابق شکل ۱ استفاده شود. به عنوان مثال برای نام‌گذاری ستون «دما» از کد TEMP و برای نام‌گذاری ستون «ارتفاع شاخص موج» از کد GTDHZZ01 استفاده شود.



Discipline	APG	PDV	PUV	Unit
DS03	D025	TEMP		
DS03	D034	WST	GTDHZZ01	

شکل ۱: انتخاب اسامی پارامترها در برنامه کاربردی فراداده

ذ- اگر پارامتری اندازه‌گیری نشده باشد، با مقدار -۹۹,۹۹۹ جایگزین شود.

ر- با توجه به توضیحات ذکر شده در مواد ۵-الف تا ۵-د، نمونه‌ای از یک داده مطلوب می تواند مشابه جدول زیر باشد:

Name	Depth	TEMP	CNDC	PSAL	Latitude	Longitude	Date	Time
Chabahar	0.21	23.51	-99.999	36.47	25.26	60.63	2007/01/11	11:41:01
Chabahar	0.77	23.49	-99.999	36.47	25.26	60.63	2007/01/11	11:41:05

¹² Decimal Degrees