

Integrierte Datenmanagement-System für marine Forschung in Kiel

Carsten Schirnick, Hela Mehrrens, Pina Springer, Lisa Paglialonga, Claas Faber

- Entstehung, Entwicklung und Aufbau der Infrastruktur
- Integrierte Komponenten
- Ausblick

seit 2009

- Zusammenschluß mit mehreren Großprojekten:
 - IFM-GEOMAR, SFBs, Exzellenz-Cluster
- Prämisse
 - **einheitliches** Datenmanagement
 - eine Arbeitsplattform
 - keine “projekteigenen” DatenmanagerInnen
- 4 Vollzeitstellen für Datenmanagement
 - später um weitere 1/2 Stelle aufgestockt
- sehr heterogenes Umfeld unterschiedlicher Disziplinen, Arbeitsweisen, Arten von Forschungsdaten und “Beziehung” zu diesen
- GEOMAR im Kontext der Helmholtz-Zentren und der marinen Forschung in Kiel

Einheitliche Arbeitsplattform

...aber keine einzelne, “ganzheitliche” Lösung verfügbar

- zentrales Portal zur “Einbindung” verteilter Dienste
 - **dauerhafte** Verwaltung auch **externer** NutzerInnen in flexiblen Strukturen
 - **einfache** Adressierung
- Komponenten für Forschungsgruppen:
 - **zentraler** Austausch von Forschungsdaten und Informationen
 - **Verknüpfung** der Informationen: Personen - Projekt - Publikationen - Daten
 - **Dokumentation** der Arbeiten und Datenentstehung
 - individuelle **Web-Präsenz** der Projekte
 - Koordination der Arbeitsgruppen

Das Portal (portal.geomar.de)

Liferay - OpenSource, kommerzielle Enterprise-Version

- Integrierte Portlets
 - Webseiten, Wikis, Blogs, Kalender, Aktivitäten, Benutzerverwaltung
- Einbindung externer Komponenten als
 - Portlets
 - iFrames
 - Verlinkung

THE DATA MANAGEMENT PORTAL

Provides access to
documentation of
management t
individual blog
es. Moreo
onal login.

infrastructure
(-Richtlinie),
cation in su

TA MAN
en Acce

JPI
OCEANS

Ecolo

JPI Oceans - Ecological Aspects of Deep-Sea Mining



Activities of deep-sea mining will indisputably affect the status of the environment by disturbing the seafloor and the overlying water column. The habitats of unique benthic and pelagic communities may be affected depending on the technologies used as well as on the size of the area impacted directly by mining and indirectly by the dispersion of sediment plumes and mining debris.

BOKNIS Coastal time series

Welcome

Welcome to the homepage of the Boknis Eck Coastal Time Series. Monthly sampling at Boknis Eck is one of the oldest - still operated - time series. This page provides an overview about the activities at Boknis Eck, the setting and the history of Boknis Eck, and access to the data database through the data link.

Boknis Eck is included in the list of the German Long Term Ecosystem Research Network (LTER-D). Since October 2014 Boknis Eck is part of the Marine Ecological Time Series (METS).

In March 2015 Boknis Eck became a member of the German Long Term Ecosystem Research Network (LTER-D).

Boknis Eck is run by the Chemical Oceanography Research Unit at GEOMAR and its activities are coordinated by [Hermann Bange](#).

Key Facts

Project Budget:
Partners: 25 Partners
Duration: Jan 2015 - Dec 2017
Cruises: SO239, SO242-1, SO242-2

Sonderforschungsbereich 754 Climate - Biogeochemistry Interactions in the Tropical Ocean

SFB 754

Project

- Project Summary
- Subprojects 2016-2019
- Subprojects 2012-2015
- Subprojects 2008-2011

Collaborative Research Center 754 / Sonderforschungsbereich 754

"Climate-Biogeochemistry Interactions in the Tropical Ocean"

The Collaborative Research Center (SFB 754) addresses the relatively newly recognized threat of ocean deoxygenation, its possible impact on tropical oxygen minimum zones and implications for the global climate-biogeochemistry system. The overall goal of the SFB 754 is to improve understanding of the coupling of tropical climate variability and circulation with the ocean's oxygen and nutrient balance, to quantitatively evaluate the nature of oxygen-sensitive tipping points, as well as to assess consequences for the Ocean's future.

The key questions of the SFB 754 are:

- How does subsurface dissolved oxygen in the tropical ocean respond to variability in ocean circulation and ventilation?
- What are the sensitivities and feedbacks linking low or variable oxygen levels and key nutrient source and sink mechanisms? In the benthos? In the water column?



Ocean Science Information System - OSIS

Eigenentwicklung

- Grails-basiertes System
- Metadaten-orientierte Struktur gegliedert nach Hauptthemen:
 - Expeditionen mit georeferenzierten “Stationen”
 - numerische Modellierungen mit georeferenziertem Output
 - Experimente
- Austausch von Forschungsdaten mit Zugriffskontrolle durch
 - Datei-Upload im Metadaten-Kontext
- Verlinkung zu spezifischen Diensten und externen Datenzentren
- Metadaten immer sicht- und abrufbar - Forschungsdaten kontrolliert

Ocean Science Information System - OSIS

- Nutzung z.B. bei Expeditionen:
 - Georeferenzierung von Lokalitäten, Expeditionsverlauf, Gerätebewegungen
=> ermöglicht Generierung von KML-Datensätzen
=> on-the-fly Datenausgabe: immer aktuellster Stand
=> Ansicht mit Google-Earth oder GIS-Werkzeugen (inkl. Bezug auf OSIS)
- ... für alle:
 - Personen- und Projekt-bezogene Relationen
zu Forschungsdaten und Fachpublikationen
=> Zugang im Kontext von Forschenden und Projekten

Ocean Science Information System - OSIS

- Nutzung z.B. bei Expeditionen:
 - Georeferenzierung von Lokalitäten, Expeditionsverlauf, Gerätebewegungen
 - => ermöglicht Generierung von KML-Datensätzen
 - => on-the-fly Datenausgabe: immer aktuellster Stand
 - => Ansicht mit Google-Earth oder GIS-Werkzeugen (inkl. Bezug auf OSIS)
- ... für alle:
 - Personen- und Projekt-bezogene Relationen zu Forschungsdaten und Fachpublikationen
 - => Zugang im Kontext von Forschenden und Projekten

OSIS Tutorial

[Overview](#)
[Expeditions](#)
[Numerical Models](#)
[Experiments](#)
[More...](#)
[Login...](#)

 Context:



32 Leg(s)/Phase(s) in this context

Identifier	Project(s)	Chief Scientist	Departure,...,Return	Events/Localities	Files	Links	WWW	
M131	SFB754	Brandt, Peter	2016/10/03 - Recife (Brazil), 2016/11/08 - Walvis Bay (Namibia)	0	0	0		Action...
M130	SFB754	Dengler, Marcus	2016/08/25 - Mindelo (Cape Verde), 2016/09/30 - Recife (Brazil)	0	0	0		Action...
SFB754-HU2016	SFB754	N.N.	2016/04/25 - Callao (Peru), 2016/05/01 - Callao (Peru)	0	0	0		Action...
M119	SFB754	Brandt, Peter	2015/09/08 - Mindelo (Cape Verde), 2015/10/13 - Recife (Brazil)	195	5	7	✓	Action...
M116/1	SFB754	Visbeck, Martin	2015/05/01 - Pointe-à-Pitre (Guadeloupe), 2015/06/03 - Mindelo (Cape Verde)	131	12	9	✓	Action...
M107	SFB754	Sommer, Stefan	2014/05/29 - Fortaleza (Brazil), 2014/07/03 - Las Palmas (Spain)	310	24	7	✓	Action...
M106	CARBOCHANGE,	Brandt, Peter	2014/04/19 - Mindelo (Cape Verde),	206	42	14	✓	Action...

Context: SFB754

Filter legs...

32 Leg(s)/Phase(s) in this context

Identifier	Project(s)	Chief Scientist	Departure,...,Return	Events/Localities	Files	Links	WWW	Action...
M131	SFB754	Brandt, Peter	2016/10/03 - Recife (Brazil), 2016/11/08 - Walvis Bay (Namibia)	0	0	0		Action...
M130	SFB754	Dengler, Marcus	2016/08/25 - Mindelo (Cape Verde), 2016/09/30 - Recife (Brazil)	0	0	0		Action...
SFB754-HU2016	SFB754	N.N.	2016/04/25 - Callao (Peru), 2016/05/01 - Callao (Peru)	0	0	0		Action...
M119	SFB754	Brandt, Peter	2015/09/08 - Mindelo (Cape Verde), 2015/10/13 - Recife (Brazil)	195	5	7	✓	Action...
M116/1	SFB754	Visbeck, Martin	2015/05/01 - Pointe-à-Pitre (Guadeloupe), 2015/06/03 - Mindelo (Cape Verde)	131	12	9	✓	Action...
M107	SFB754	Sommer, Stefan	2014/05/29 - Fortaleza (Brazil), 2014/07/03 - Las Palmas (Spain)	310	24	7	✓	Action...
M106	CARBOCHANGE, SOPRAN, SFB754	Brandt, Peter	2014/04/19 - Mindelo (Cape Verde), 2014/05/26 - Fortaleza (Brazil)	206	42	14	✓	Action...
M105	CARBOCHANGE, SFB754, SOPRAN	Visbeck, Martin	2014/03/17 - Mindelo (Cape Verde), 2014/04/16 - Mindelo (Cape Verde)	221	25	24	✓	Action...
Kristineberg_2013/2	SFB754	Hahn, Johannes	2013/08/20 - Kristineberg (Sweden), 2013/08/21 - Kristineberg (Sweden)	10	5	0		Action...
M131	Kristineberg_2013/1	Hahn, Johannes	2013/08/20 - Kristineberg (Sweden), 2013/08/21 - Kristineberg (Sweden)	13	7	0		Action...
M130	M97	Tanhua, Toste	2013/05/25 - Mindelo (Cape Verde), 2013/06/28 - Fortaleza (Brazil)	275	21	14	✓	Action...
SFB754	M93	Lavik, Gaute	2013/02/06 - Callao (Peru), 2013/03/10 - Cristobal (Panama)	310	37	24	✓	Action...
M119	M92	Sommer, Stefan	2013/01/05 - Callao (Peru), 2013/02/03 - Callao (Peru)	290	47	27	✓	Action...
M116/1	M91	Bange, Hermann W.	2012/12/01 - Callao (Peru), 2012/12/26 - Callao (Peru)	199	44	43	✓	Action...
	MSM23	Visbeck, Martin	2012/11/26 - Mindelo (Cape Verde), 2012/12/20 - Walvis Bay (Namibia)	126	16	8	✓	Action...
M107	M90	Stramma, Lothar	2012/10/28 - Cristobal (Panama), 2012/11/28 - Callao (Peru)	184	20	31	✓	Action...
M106	ortal.geomar.de/metadata/leg/show/322815	Brandt, Peter	2012/10/24 - Mindelo (Cape Verde),	247	52	17	✓	Action...

Edit Show Web Info

Leg/Phase: M106
Cruise/Expedition: M106
Platform: Meteor (M)
Departure/Return: Departure: 2014/04/19 - Mindelo (Cape Verde)
Return: 2014/05/26 - Fortaleza (Brazil)
Add...

Research Area: Tropical Atlantic

Expocode:

Notes:

Project(s): SOPRAN; CARBOCHANGE; SFB754

Participating GEOMAR Research Unit(s):

Organizing Institution: Helmholtz Centre for Ocean Research Kiel

Chiefscientist: Brandt, Peter

Chiefscientist's Affiliation: Helmholtz Centre for Ocean Research Kiel

Vicechiefscientist: N.N.

Created: 2013-07-01 11:53

Last Updated: 2014-09-25 13:44

Community Context: [SFB754, CVOO]

Upload New File for Leg M106...

Filename	Version	Size [kB]	Q-Flag	Description	Uploaded	File Creator (Urheber)	
leg_M106_NetworkLink.kml	Dynamically generated file to view Leg(s)/Phase(s) with Events/Localities e.g. in Google Earth						
ifm13_depl01_glider_gridded.nc	1	55873	final	Gridded glider data from deployment ifm13_depl01. T,S,O calibrated against CTD from M105 and M106.	2016-05-09 11:51	Krahmann, Gerd	Action...
Readme_sfb754_UVP5_bin_data.txt	2	7	calibrated	Readme for uvp5_bin_data. Corrected wrong unit for particle abundance data.	2015-10-28 20:07	Kiko, Rainer	Action...
N2O.docx	1	17	preliminary	Readme for N2O data	2015-09-10 12:15	Grundle, Damian	Action...
N2O_data_Rebecca.xlsx	1	29	preliminary		2015-09-10 12:14	Grundle, Damian	Action...
UVP5_bin_data_M106_20150905.txt	1	54242	calibrated	Contains particle data obtained with the Underwater Vision Profiler 5. Please check the Readme_sfb754_UVP5_bin_data.txt for further details.	2015-09-05 19:58	Kiko, Rainer	Action...
Readme_sfb754_Multinet_item_data.txt	1	3	preliminary	Readme for multinet mesozooplankton item data	2015-09-03 10:56	Kiko, Rainer	Action...
M106_mesozooplankton_net_item_data_20150903.txt	1	22401	preliminary	Single item mesozooplankton data. For further details please see Readme_sfb754_Multinet_item_data.txt	2015-09-03 10:56	Kiko, Rainer	Action...
Readme_sfb754_Multinet_bin_data.txt	1	3	preliminary	Readme for multinet mesozooplankton bin data	2015-09-03 10:56	Kiko, Rainer	Action...
M106_mesozooplankton_net_bin...	1	236	preliminary	Binned mesozooplankton data. F...	2015-09-03	Kiko, Rainer	Action...

Ocean Science Information System - OSIS

- einfaches, kontextbezogenes Ticketing-System
 - Anlegen von “Deliverables”
 - Darstellung zu liefernder bzw. zu erwartender Datensätze
 - Status während langer Projektlaufzeiten nachvollziehbar
- Webservices für Expeditionen und Forschungsprojekte
 - versorgen Webseiten kontext-bezogen mit
 - Expeditionsinformationen (institutsübergreifend)
 - Projektinformationen (nur GEOMAR, aus Verwaltung)

Ocean Science Information System - OSIS

- zentraler Informationsknoten
- stets aktuelle und verlässliche Informationen (single-point-of-truth)

- aber: Forschungsdaten in
 - **Dateien** unterschiedlicher Struktur oder
 - in spezifischen **Datenbanken**
 - an verschiedenen **Orten**
- "...one solution doesn't fit all..."

OPeNDAP - Schnittstelle zu Forschungsdaten

THREDDS Server - OpenSource

- Auffindbarkeit über Kataloge
- Selektion und Abruf strukturierter Daten (z.B. netCDF) per OPeNDAP
 - alternativ Download
- web-basierter, zentraler Zugriff auf georeferenzierte
 - Modelldaten
 - Messdaten

Medien-Server für Filme und Fotos

ProxSys - kommerzielle Software

- Archivierung von Foto und Film nach Großgeräteeinsätzen
- Auffindbarkeit von Medieninhalten anhand
 - ausgelesener EXIF-Metadaten
 - manuell ergänzter Metadaten
- Sichtung, Auswahl und Bereitstellung anhand Vorschauen
- manuelle und automatisierte wissenschaftliche Annotation:
 - für Fotos und Filme implementiert
 - Archivierung der Annotationen in “universeller” Lesbarkeit beim Inhalt
 - georeferenzierbare Integration in OSIS anhand der Metadaten

OceanRep - Institutionelles Volltext-Repository

ePrints Software - OpenSource mit kommerzieller Wartung

- Verknüpfung der Artikel und Forschungsdaten via OSIS mit Personen und Projekten
- kontextbezogene Publikationslisten als Service
- zentrale Pflege durch Bibliotheks-Team
- Synchronisation mit Universität Kiel in Arbeit

Versionierungswerkzeuge

Subversion und Git(Lab) - OpenSource

- für Quellcode numerischer Modellierungen
 - Dokumentation im gekoppelten Wiki
- für Forschungsdaten
 - Subversion
 - Handhabung grosser Dateien
 - Anwendungsfall Fotos: automatisiert in Vignetten zerlegt, manuell sortiert
 - Git
 - dezentrale Versionierung
 - Anwendungsfall Workflow-Dokumentation: Jupyter Notebooks
 - ToDo... oder ToExplore...

Versionierungswerkzeuge

Subversion und Git(Lab) - OpenSource

- für Forschungsdaten
 - einfache Versionierung derzeit im OSIS
 - Client-Werkzeuge für Subversion und Git ermöglichen
 - unschlagbare Synchronisation zentralen und/oder dezentral
 - aber wie komplexe Nutzungsszenarien “verstecken”
 - Umgang mit grossen/binären Dateien in Git

Workflow - Ausführung+Dokumentation

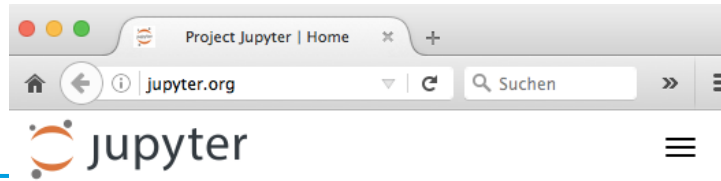
Jupyter Notebook - OpenSource

- ...ermöglicht gleichzeitig
 - **Ausführung**
von Programmen, Skripte, Code in zahlreichen Sprachen und
 - **Dokumentation** der Ausführung in einem
- **lokal** und **remote** ausführbar
- **wissenschaftliche** Verarbeitung der Daten
- Abfrage von **Metadaten** per Notebook Templates
- einfache Versionierung und Verteilung

Workflow - Ausführung+Dokumentation

Jupyter Notebook - OpenSource

- ...ermöglicht gleichzeitig
 - **Ausführung**
von Programmen, Skripte, Code in zahlreichen Sprachen und
 - **Dokumentation** der Ausführung in einem
- **lokal** und **remote** ausführbar
- **wissenschaftliche** Verarbeitung der Daten
- Abfrage von **Metadaten** per Notebook Templates
- einfache Versionierung und Verteilung



Open source, interactive data science and scientific computing across over 40 programming languages.



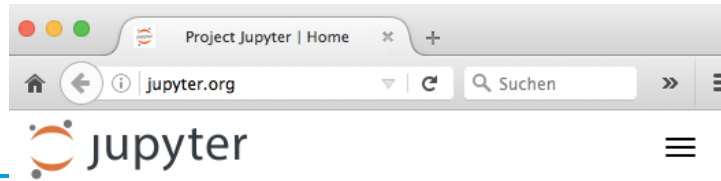
Jupyter Notebook

Komponenten

Workflow - Ausführung+Dokumentation

Jupyter Notebook - OpenSource

- ...ermöglicht gleichzeitig
 - **Ausführung**
von Programmen, Skripte, Code in zahlreichen Sprachen und
 - **Dokumentation** der Ausführung in einem
- **lokal** und **remote** ausführbar
- **wissenschaftliche** Verarbeitung der Daten
- Abfrage von **Metadaten** per Notebook Templates



Open source, interactive data science and scientific computing across over 40 programming languages.



Jupyter Notebook

Skript um Liste der Dateipfade für ein Projekt auf hsm:/proxsys/toolbox/ zu kopieren

```
In [1]: import os, sys, inspect, datetime
# realpath() will make your script run, even if you symlink it :)
cmd_folder = os.path.realpath(os.path.abspath(os.path.split(inspect.
if cmd_folder not in sys.path:
    sys.path.insert(0, cmd_folder)

# use this if you want to include modules from a subfolder
cmd_subfolder = os.path.realpath(os.path.abspath(os.path.join(os.pat
if cmd_subfolder not in sys.path:
    sys.path.insert(0, cmd_subfolder)

import pypxtools.pypxtools as pxt
```

- UUID des Projektes eintragen und dann mit Login und Passwort anmelden:

```
In [2]: # get project data
query = pxt.PxQuery()
prj = query.get_project('067558e03fbec922013fc32e4db80023')
```

```
user [cschirnick]:
password:.....
```

```
In [3]: # setup output options
```



Open source, interactive data science and scientific computing across over 40 programming languages.



Jupyter Notebook

Skript hsm:

```
In [1]: import  
# rea  
cmd_fc  
if cmd  
sy  
  
# use  
cmd_su  
if cmd  
sy  
  
import
```

• UU

```
In [2]: # get  
query  
prj =
```

user [
passwo

```
In [3]: # setu
```

```
1 {  
2 "cells": [  
3 {  
4 "cell_type": "markdown",  
5 "metadata": {},  
6 "source": [  
7 "## Skript um Liste der Dateipfade für ein Projekt auf hsm:/proxsys/toolbox/ zu kopie  
8 ]  
9 },  
10 {  
11 "cell_type": "code",  
12 "execution_count": 1,  
13 "metadata": {  
14 "collapsed": false  
15 },  
16 "outputs": [],  
17 "source": [  
18 "import os, sys, inspect, datetime\n",  
19 " # realpath() will make your script run, even if you symlink it :)\n",  
20 "cmd_folder = os.path.realpath(os.path.abspath(os.path.split(inspect.getfile( inspect  
21 "if cmd_folder not in sys.path:\n",  
22 " sys.path.insert(0, cmd_folder)\n",  
23 "\n",  
24 "# use this if you want to include modules from a subfolder\n",  
25 "cmd_subfolder = os.path.realpath(os.path.abspath(os.path.join(os.path.split(inspect.  
26 "if cmd_subfolder not in sys.path:\n",  
27 " sys.path.insert(0, cmd_subfolder)\n",  
28 " \n",  
29 "import pypxtools.pypxtools as pxt"  
30 ]  
31 },  
32 {  
33 "cell_type": "markdown",  
34 "metadata": {},
```



data science and scientific
10 programming languages.



Notebook

Vielen Dank!

...auch für Fragen und Anregungen