

Анализ коллекций климатических данных для тематической виртуальной исследовательской среды

(Analysis of climatic data collections for thematic virtual research environment)



Институт мониторинга климатических и экологических систем СО РАН, Томск, Россия; Томский государственный университет, Томск, Россия; Институт оптики атмосферы СО РАН, Томск, Россия



Presented the first results of preprocessing and integration of climate data collections, obtained from observations and modeling, into developed thematic virtual research environment for analysis, assessment and predicting of global climate change impact on the environment and climate in the selected region. Description of climate data collections, a selected data storage model, a metadata database containing key characteristics of data sets and procedures for their processing and analysis, as well as an administrative web console for it are given.

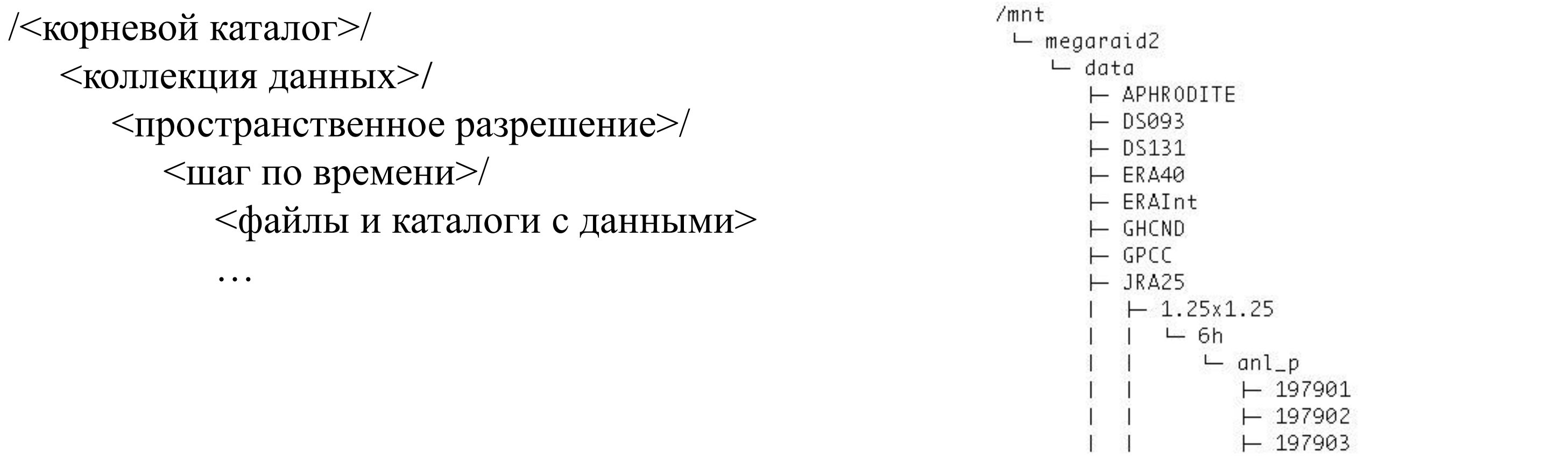
Введение

Актуальность создания виртуальной исследовательской среды обусловлена произошедшей во многих отраслях науки сменой характера научного труда, вызванной прогрессом компьютерных технологий. Благодаря этому прогрессу ученые хотят получать доступ к различным данным, ресурсам и сервисам обработки через веб-браузер, вне зависимости от места их расположения и принадлежности. В области климатической науки и ее приложений такой подход к решению научных задач уже стал необходимостью. Состояние климатической системы характеризуется большим набором величин, в число которых входят метеорологические элементы, отражающие свойства атмосферы и атмосферных процессов. Они регистрируются при выполнении стандартных метеорологических наблюдений, однако разреженность сети наблюдательных пунктов не позволяет простыми методами интерполяции получить поля необходимого для анализа глобальных и региональных климатических процессов пространственного разрешения. Для решения этой проблемы используются глобальные метеорологические модели, обеспеченные системами усвоения многолетних рядов инструментальных наблюдений. Полученные поля метеорологических элементов (так называемые реанализы) обеспечивают глобальное покрытие по всему Земному шару на равномерной сетке и, в определенной степени, восстанавливают метеорологическую ситуацию для выбранного временного интервала. Зачастую, эти поля хранятся в различных форматах и заданы на несовпадающих пространственных сетках, поэтому, прежде чем выполнять совместный сравнительный анализ различных характеристик, их необходимо привести к единому формату файлов и единой пространственной сетке. В рамках данной работы был выполнен сбор, анализ и подготовка к интеграции в создаваемую виртуальную исследовательскую среду (ВИС) наборов геопривязанных данных, детально описывающих динамику современных и возможных в будущем климатических и экологических изменений. Для обеспечения возможности выполнения сравнительного анализа различных характеристик были разработаны специализированные программные процедуры, приводящие поля геопривязанных данных к единому формату и пространственной сетке. В подготовленный архив геопривязанных данных вошли следующие наборы:

Набор данных	Организация	Охват по времени	Пространственное разрешение
APHRODITE	RHN-MRI/JMA	1951 - 2007	0.25°×0.25°
ERA-40	ECMWF	1957 – 2004	2.5°×2.5°
ERA Interim	ECMWF	1979 – 2012	0.75°×0.75°
GPCC	GPCC	1901 - 2009	0.5°×0.5°
JRA-25	JMA/CRIEPI	1979 – 2009	2.5°×2.5°
JRA-55	JMA/CRIEPI	1958 – 2013	1.25°×1.25°
MERRA	ECMWF	1979 - 2014	0.67°×0.5°
NCEP/DOE AMIP II	NCEP/DOE	1979 – 2003	2.5°×2.5°
20th Century Global Reanalysis Version II	NOAA/OAR/ESR L PSD	1869 – 2011	2.0°×2.0°
Модель PlaSim	ИМКЭС СО РАН	2000 - 2100	2.5°×2.5°
Метеостанции РФ	ВНИИГМИ-МЦД	1910 – 2016	600 станций

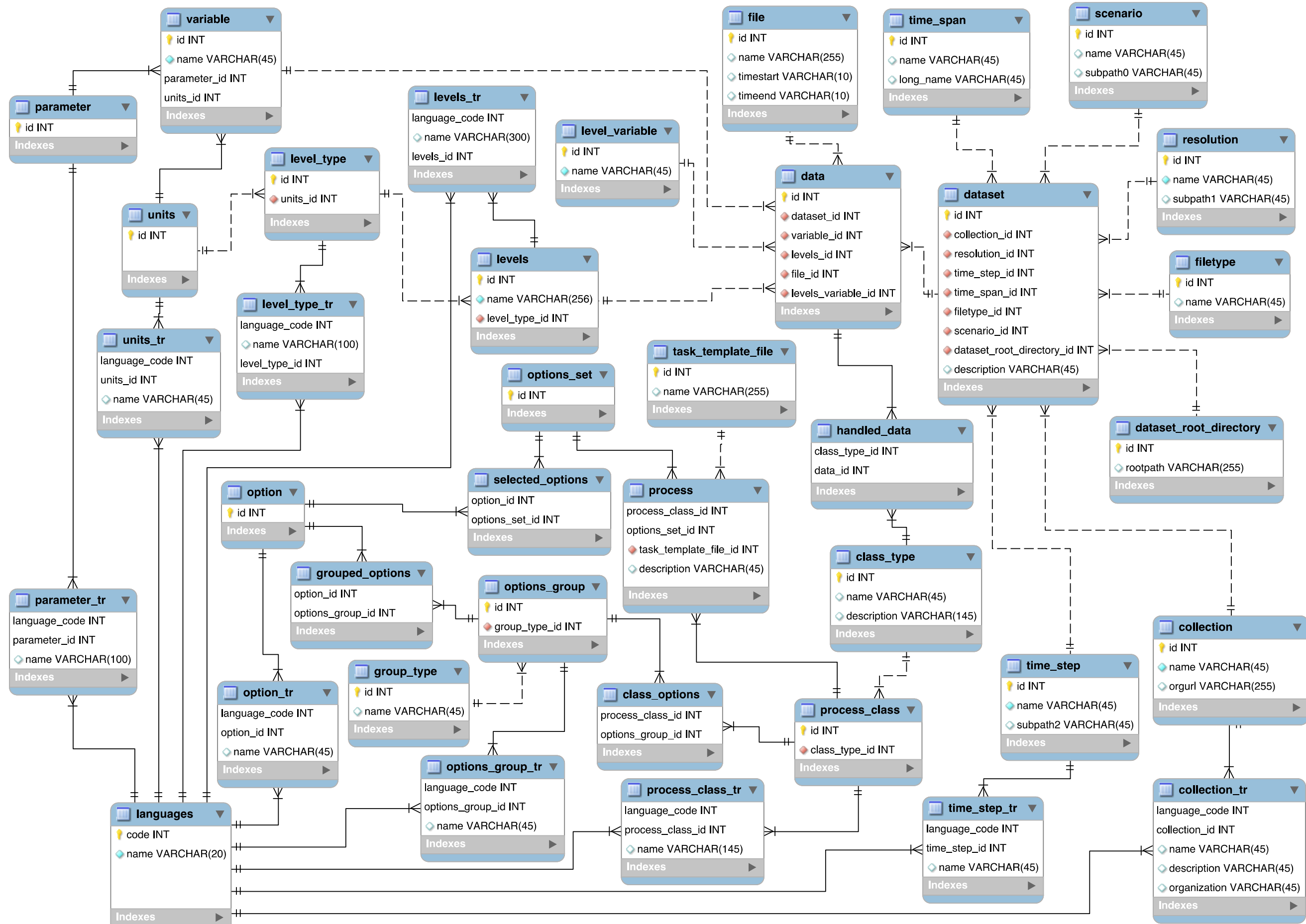
Модель хранения данных

В качестве единого формата файлов для хранения полей геопривязанных данных, описывающих метеорологические характеристики и характеристики поверхности был выбран формат netCDF. Он формально признан научно-исследовательскими и стандартизирующими органами во всем мире, в том числе членами OGC, как кандидат в стандарты OGC для стимулирования международного сотрудничества и увеличения интероперабельности между клиентами и серверами, обменивающимися данными в двоичной форме. Он отлично приспособлен для хранения геопространственных данных, поддерживается широким спектром различного коммерческого и открытого программного обеспечения. С использованием разработанных процедур исходные файлы, содержащие поля геопривязанных данных, были преобразованы в формат netCDF с приведением к нескольким кратным пространственным и временным сеткам для обеспечения быстрой обработки на различных масштабах, и размещены в следующей иерархии каталогов на системах хранения данных:



База метаданных

Для быстрого поиска и выборки необходимых файлов данных по заданным параметрам была разработана база метаданных, содержащая описания пространственно-временных характеристик доступных для обработки наборов данных, расположение в иерархии файловой системы на системах хранения данных файлов с данными, описание и параметры запуска программных компонент для анализа данных, а также взаимосвязи между ними:

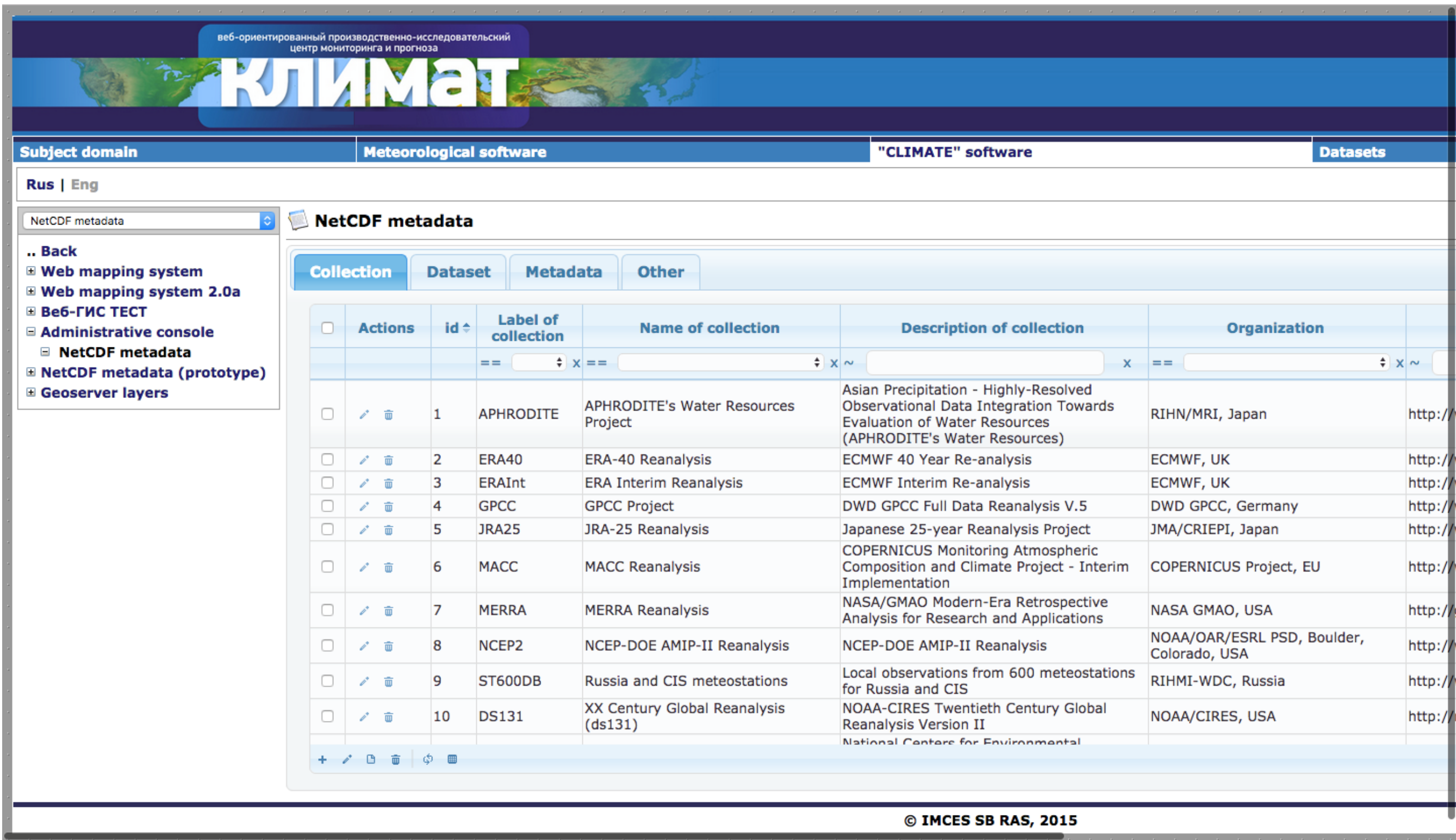


Административная веб-консоль базы метаданных

Для обеспечения интерактивного взаимодействия конечного пользователя с базой метаданных была разработана административная веб-консоль базы метаданных, являющаяся частью веб-портала ВИС. Она представляет собой функционально-законченный модуль программного обеспечения ВИС, предоставляющий возможности создания, отображения, редактирования и удаления информации из базы метаданных с поддержкой целостности базы. Разработанная административная веб-консоль базы метаданных выполняет следующие функции:

- отображение значений таблицы базы метаданных;
- сортировка по отдельной колонке или множественная сортировка;
- фильтрация результатов по ограничению на значение в колонке, также возможны различные ограничения и их сброс;
- динамическая загрузка и пейджинг страниц результатов по скроллингу;
- добавление записи с автоподстановкой результатов при печати в формы ввода;
- редактирование записи с автоподстановкой результатов при печати в формы ввода;
- просмотр полной записи по клику на значок или двойному клику по строке, при этом возможен просмотр записей влево-вправо в рамках подгруженной страницы отображения;
- динамическое обновление данных;
- фильтрация колонок отображения;
- удаление одной или нескольких записей в таблицах (с учётом взаимосвязей) одновременно.

Административная веб-консоль состоит из набора сценариев, написанных на языках PHP, JavaScript, HTML, с использованием JavaScript-библиотек jQuery, jqGrid, и обеспечивающих пользователя ВИС интуитивно-понятным онлайн-инструментом. Её веб-интерфейс также предоставляет интерфейс программирования приложений для реализации соответствующих PHP-контроллеров на серверной стороне, обрабатывающих запросы веб-клиента.



Заключение

В результате работы была получена первая версия архива климатических данных, интегрированного в создаваемую виртуальную исследовательскую среду и содержащего надежную, подготовленную для интерактивного анализа информацию о происходящих и возможных в будущем климатических и экологических процессах. Работа выполнена при поддержке гранта РНФ № 16-19-10257.