Member Node Description: DataNet Federation Consortium (DFC)

Version 1.0 3/9/16 Lisa Stillwell, Chris Lenhardt

General

Name of resource:	DataNet Federation Consortium (DFC)
URL(s):	http://datafed.org/
Institutional affiliation(s):	University of North Carolina at Chapel Hill, University of Virginia, University of California at San Diego
Primary geographic location:	UNC, Chapel Hill, NC
Project Director & contact info:	Reagan Moore, rwmoore@email.unc.edu
Technical Contact & contact info:	Lisa Stillwell, lisa@renci.org
Age of resource:	Since 2013
Funding support:	NSF
Proposed Unique Identifier:	urn:node:DFC

Content

Content description/collection policy (1 paragraph, domain and spatial/temporal coverage, uniqueness of content, exclusions, as applicable):
The DataNet Federation Consortium (DFC) Hydro-NEXRAD rainfall accumulation data for the Carolinas, Maryland, and Virginia uses Hydro-NEXRAD algorithms (http://www.crwr.utexas.edu/gis/gishydro08/SpaceAndTime/hydroNEXRAD.htm) and NEXRAD data from the NOAA National Centers for Environmental Information (NCEI) to provide spatial datasets that estimate rainfall accumulation at 15-minute time intervals for parts of the Carolinas, Maryland, and Virginia. The temporal domain for these data span August 2013 to the present. The data and web services are hosted at the computational facilities of the Renaissance Computing Institute (RENCI), in Chapel Hill, NC.

Types of data (complex objects, text, image, video, audio, other):
3D NetCDF formatted data – rainfall accumulation, based on latitude, longitude & time

Data and metadata availability (rights, licensing, restrictions):
These data are made available under a Creative Commons license: Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). See http://creativecommons.org/licenses/by-nc/4.0/ for more information.

Option for embargo (yes/no, duration):
No

Size of holdings (number and size of datasets, mean and median granules (files) per dataset):
Collection holdings are increasing daily. Every day, 24 hours worth of NEXRAD data (as available) from 6 radar stations, is processed with Hydro-NEXRAD model to create hourly NetCDF files, for each station. Each NetCDF file averages 20Mb. Currently data is available since 2/2016. More historical (since 2014) data is coming soon.

Please describe recent usage statistics, if known, including information on annual data product downloads, annual number of users, annual number of data products used in publications:
No usage statistics are available at this time as this is a newly published data collection.

User interactions

How does a user contribute data? (what can be deposited, how are data prepared, are specific software required, documentation/support available)
N/A

How does a user acquire / access data?
Download links as part of the metadata returned using the DataOne Search interface

What user support services are available (both for depositing and accessing/using data)?
Questions can be directed to Lisa Stillwell, lisa@renci.org

How does the resource curate data at the time of deposit?
Data is organized in hourly (24) increments in a day. It is also grouped by the NEXRAD station that produced the radar data. Currently the radar data from 6 stations are used:

KRAX - Raleigh/Durham, NC
KCAE - Columbia, SC
KGSP - Greer, SC
KLTX - Wilmington, NC
KLWX - Sterling, VA
KMRX - Knoxville, TN

Technical characteristics and policies

Software platform description, incl. data search and access API(s):
iRODS, Jargon, JBOSS-RestEasy, Java, ElasticSearch

Service reliability (including recent uptime statistics, frequency of hardware refresh, if known):
Not known

Preservation reliability (including replication/backup, integrity checks, format migration, disaster planning):
Offsite replication, using institutional resources, for back up, at the University of North Carolina at Chapel Hill.

User authentication technology (incl. level of create/modify/delete access by users):
Read Only

Data identifier system and data citation policy, if available:
Handle Server (https://www.handle.net)

Metadata standards (including provenance):
NetCDF CF-1.4: This format is based on the attribute data tags defined by the Cooperative Ocean/Atmosphere Research Data Service (COARDS) and Climate and Forecast (CF) metadata conventions

Capacity/services to DataONE

At what functional tier will you initially be operating? (see http://bit.ly/MNFactSheet for definitions)

☒ **Tier 1: Read only, public content**

- ☐ **Tier 2: Read only with access control**
- ☐ **Tier 3: Read/write using client tools**
- ☐ **Tier 4: Able to operate as a replication target**

If you can host data from other member nodes, what storage capacity is available?
None

Can you provide computing capacity to the broader network? If so, please describe.
No

Other Services

What other services or resources (such as expertise, software development capacity, educational/training resources, or software tools) can be provided of benefit to the broader network?
We are developing a docker version of the DFC-DataONE integration as well as developing other types of services such as HIVE ontology comparison capabilities. DFC would be interested in discussing potential collaborations of benefit to the broader network.