

DDWG Quarterly Report - 202205

Author: Kevin Krieger

With information from the period February 1 2022 to April 30 2022

The report periods are:

1. February 1 to April 30 with an update due by May 15
2. May 1 to August 31 with an update due September 15
3. September 1 to October 31 with an update due November 15
4. November 1 to January 31 with an update due by February 15

Executive Summary

A current file listing for differences between USASK and BAS, USASK and NSSC was generated, please see files attached with this report. 210 files have been retrieved since November 2021 due to gaps file checking.

General Announcements

The following are general updates:

1. The consortium that hosts data using Globus for the University of Saskatchewan, Compute Canada, is shut down effective March 31 2022. In its place is the Digital Research Alliance of Canada.
2. Work has started on uploading 2019 and 2020 datasets to **FRDR** - this would be a DOI'd dataset much like the rest of the datasets already there.
3. Corrupted JME files from 20220113-20220128 were found, and replaced by NSSC on their server in mid-March. BAS has temporarily stopped transfers of JME.
4. Globus was unavailable for a scheduled downtime to update the Operating system, system software and drivers on March 21 to March 24 2022.
5. An issue with Canadian files from CLY, INV and RKN with lag0 data from far ranges was found in late 2021. The issue was fixed and affected files replaced by March 29 2022. Please ensure you have the updated versions.
6. 48 files from 2020 and 2021 DCE and DCN were found from the gaps check and subsequently uploaded to the Globus and NSSC mirrors. This makes a total of 210 files since November 2021 when gaps emails were sent to the data managers.
7. An analysis of file differences between BAS, USASK and NSSC was performed April 28th, relative to USASK. File lists will be sent along with this report. Here is a summary:
 1. 1303 blocked files are on BAS (mostly SAS commissioning files removed and blocked in 2021)

2. 12,170 files are different between USASK and BAS from 202001 to present (there are previous differences not included here)
3. 1580 files are different between USASK and NSSC from 202104 and 202111
4. 1581 files are missing from BAS that USASK has (201603 onwards)
5. 839 files are missing from NSSC that USASK has - most of these failed a DMAP check (200807 onwards)

Outstanding Tasks & Future projects

1. Ensure all mirrors have the same standards for file inclusion
2. Get all 2020 files to include in **FRDR**
3. Get all 2021 files ready to include on FRDR by the end of December 2022.