

Activities and Status of the Space Physics Data Facility (SPDF) Archive for the Past Year

Robert Candey, Lan Jian, and the rest of the SPDF team

Fourth IHDEA Meeting

2020 Oct. 19-22

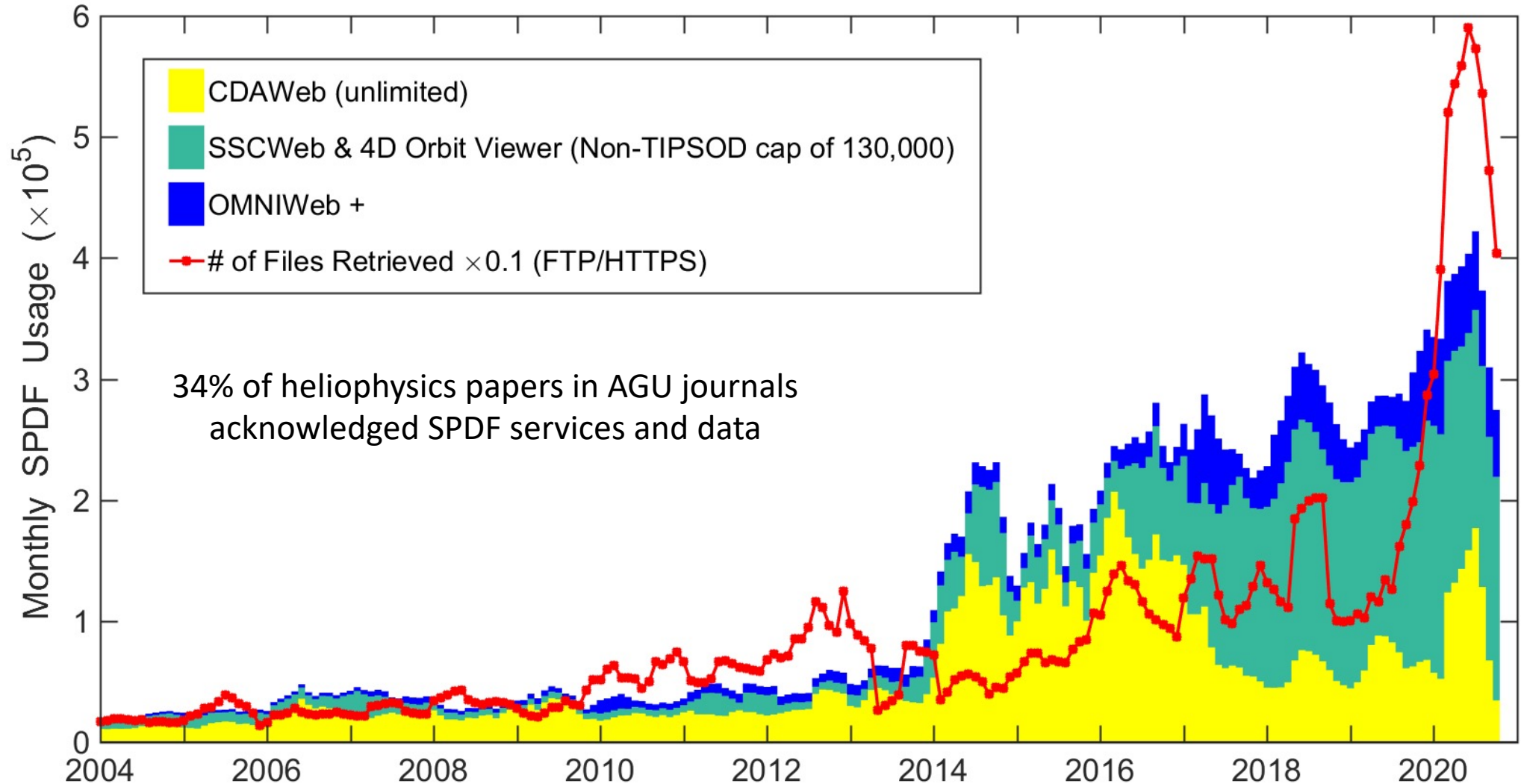
SPDF <<https://spdf.gsfc.nasa.gov>>

SPDF activities in past year and near future

- Added AIM, GOLD, Parker Solar Probe (85), IBEX data (40), and Explorer-35
- Working on Solar Orbiter (74) and ICON data, final data from Van Allen Probes, TWINS.
- New mission requirements <https://spdf.gsfc.nasa.gov/guidelines/archive_newdata_req.html>
- Exploring changes to become part of NASA's Heliophysics Digital Resource Library, such as additional outreach materials, user support, and support for data in cloud services
- Will add webservice for event lists for burst mode data and science events (CMEs, bow shock crossings, etc.) and support to SSCweb and CDAWeb
- Adding SPASE Resource ID and DOI to CDAWeb metadata and display
- Redirected alternative server names, such as `cdaweb.sci.gsfc.nasa.gov`, and switched FTP to FTPS, so please update your codes and scripts
- CDAWeb added netCDF support, `variable_purpose` attribute, IBEX mapped images
- CDF: planning new features and tools, systematize ISTP metadata guidelines, explore cloud support
- Adding support for end dates in filenames
 <dataset>_<beginTime>-<endTime>_<varyingInfo>_v<version>.cdf [hyphen or underscore between times]
- Developing alternatives to the Java-based 4D Orbit Viewer and SKTeditor tools

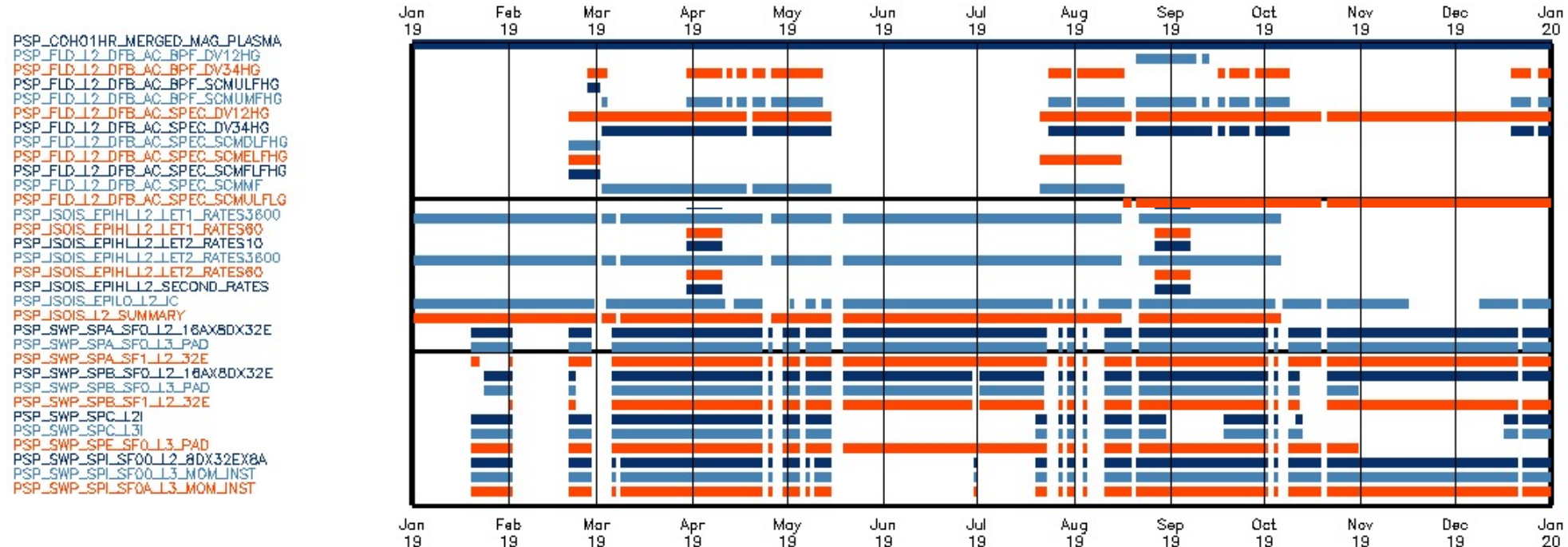
SPDF Statistics

(see reports at <https://cdaweb.gsfc.nasa.gov/publiclogs/>)



Parker Solar Probe

Part of 2019 Inventory Plot



- Since the launch in Aug 2018, PSP has finished five orbits around the Sun, reaching a perihelion of 25 solar radii
- Three in situ instrument suites: SWEAP, FIELDS, ISOIS. About 75 data sets in CDAWeb.
- To achieve the flexibility in data products and data rates necessary for an encounter-based mission of exploration, the FIELDS instrument incorporates a broad range of selectable configurations for many of its core functions, resulting in many data sets or many variables in some data sets
- Due to different telemetries in encounter phase and cruise phase, there are many data gaps. The inventory plot provided by SPDF helps the users to search for data availability
- SPASE group supported by NASA has made SPASE description for these data sets

Solar Orbiter

- Solar Orbiter is a joint mission between ESA and NASA, but primarily an ESA mission. Solar Orbiter Archive (SOA) is the primary archive
- NASA mirrors the data at SOA and then SPDF ingests and archives the in-situ data from the mirror
- There are 16 low latency data sets served at CDAWeb
- There are more than 80 science data sets. We are still ingesting and testing them at CDAWeb
- The SPASE descriptions for all these data sets are not available yet

SPDF action items from the last IHDEA meeting

- 1. Action on SPDF: to circulate on a regular basis, at least yearly, to IHDEA members the roadmap of CDF development
 - See next page for CDF status
- 2. Action on SPDF: ISTP rules to allow filenaming with start and end dates, in such a way that these filenames will not cause errors or warning by SKTeditor
 - See following page for proposed changes
- 3. Action on NASA and ESA/ESDC: to coordinate/facilitate the transfer of all science datasets of the Cluster mission from ESDC to SPDF
 - No action yet due to so many new missions this year, but we still plan to copy Cluster and SOHO data

CDF Status and Recent Development

- Released CDF 3.8.0.1
- Continued CDF support and general development, plus added features
- Use FILLVAL to fill the missing data if it exists, instead of the pad value
- Added options to cdfconvert (remove dimension without DEPEND_*, use FILLVAL if exists in place of pad value)
- Use FORMAT attribute for data listing if it exists
- Update zlib package to V2.1.11, for code improvements and bug fixes
- More generalized CDF epoch data encoding and parsing functions, default encoded epoch data now ISO-8601 format
- CDF epoch data conversion to/from Unix time
- String typed variables padded with a single space and followed by NULs
- cdf2skt tool option to choose how to display variable's metadata and data

CDF Plans

- High-level functions to read variables or whole CDF into a map structure for IDL, Java, Perl and C#
- CDF-JSON converter
- Improve Windows installer, autoconf/make build/install, Mave/Ant/Gradle installs
- Improve documentation, beginner's guides, add to Wikipedia CDF entry
- Standardize ISTP/IACG Metadata Guidelines with version control, etc.
- New SKTeditor in Javascript or Python, perhaps also SPASE metadata creation
- Looking into supporting CDFs in cloud object storage, perhaps Zarr like netCDF is exploring
- Define CDF MIME type and international standard
- Apache 2 license in place of current custom license
- Update CDFML and its corresponding JSON representation with cdf.xsd use more specific datatype (e.g., xs:dateTime, xs:integer, xs:float, etc.) instead of just xs:string
- Add support for CDF to command line netCDF tools, such as NCO, NCAR, ANTS, NCtools
- CDF gap checker to write filename, variable name, begin and end time, number of records, and any gaps greater than a certain amount (G-good, M-missing, F-fill, R-outside range, B-backward time)
- Add CDF support to Octave, Gnu Data Language (GDL), Excel, Ruby, C++, WebWinds, LinkWinds, Opendap, SWIG.org

File naming changes

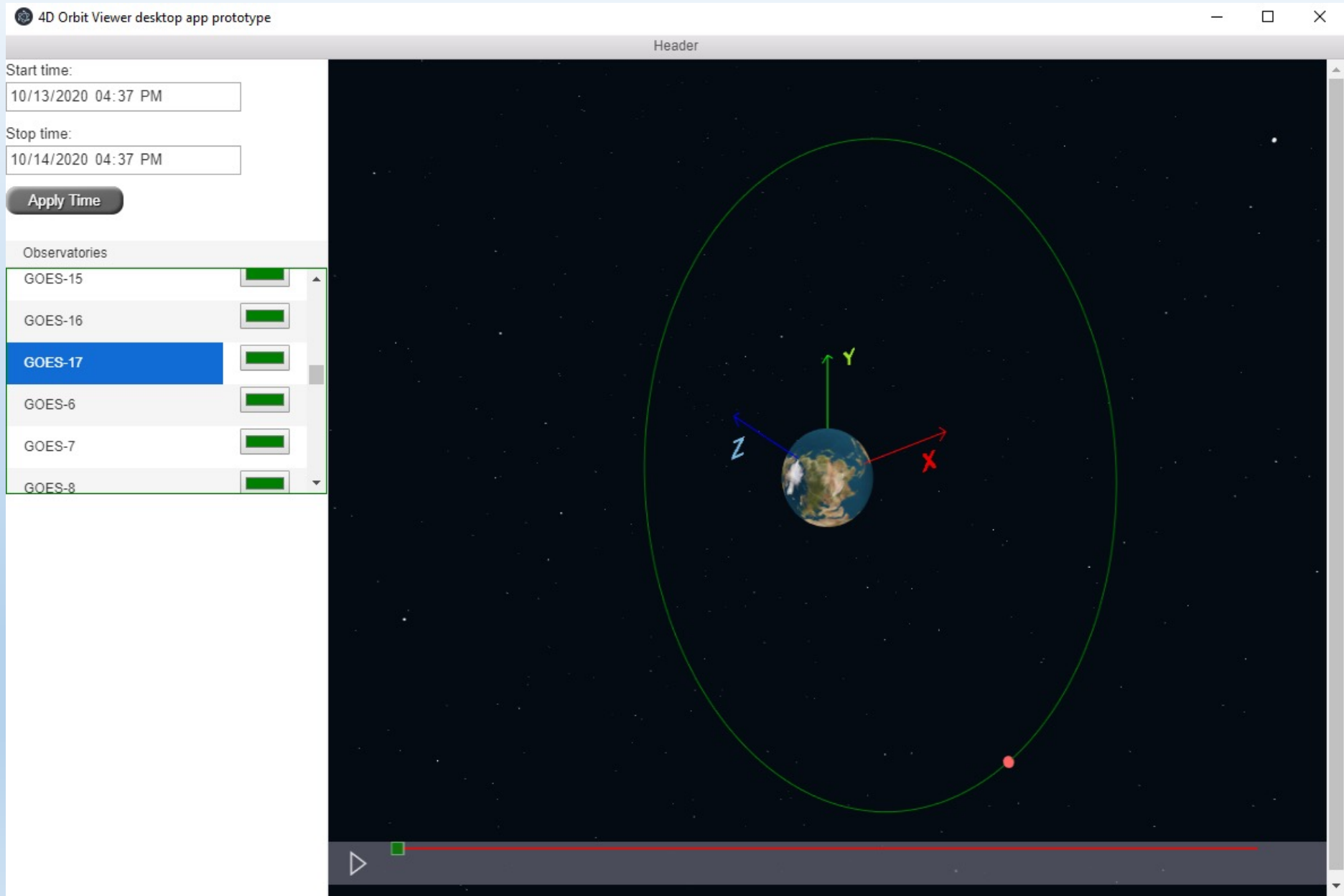
- Need to add following to the current recommended file naming guidelines <https://spdf.gsfc.nasa.gov/guidelines/file_naming_recommendations.html>:
- Add support for end times to the overall pattern:
 - <dataset>_<beginTime>_<varyingInfo>_v<version>.cdf [underscore between sections, hyphens within sections]
 - <dataset>_<beginTime>-<endTime>_<varyingInfo>_v<version>.cdf [hyphen between times]
 - <dataset>_<beginTime>_<endTime>_<varyingInfo>_v<version>.cdf [underscore between times]
- Formats for the begin/endTime :
 - YYYYMMDD
 - YYYYDOY
 - YYYYMMDDHHMMSS
 - YYYYMMDDTHHMMSS
 - Probably others.
- Version/revision formats (keep only greatest one)
 - Normal/standard: V01
 - Complex RBSP version: v1.1.0-00, e.g. rbsp-a-rbspice_lev-3_isrheld_20161208_v1.1.9-00.cdf

Replacements to our Java tools

- Developing alternatives to the Java-based 4D Orbit Viewer and SKTeditor tools
- New SKTeditor may also enable the user to define the SPASE metadata at the same time as defining the internal metadata and structure of the CDF/netCDF to be created.
- Screen shots of prototypes of HTML5/Javascript-based interfaces for the 4D orbit viewer are below. We plan to expand this project to add interactive data plotting and sonification tied to the orbit display, similar to CDAWeb and perhaps data glyphs along the orbits as well.

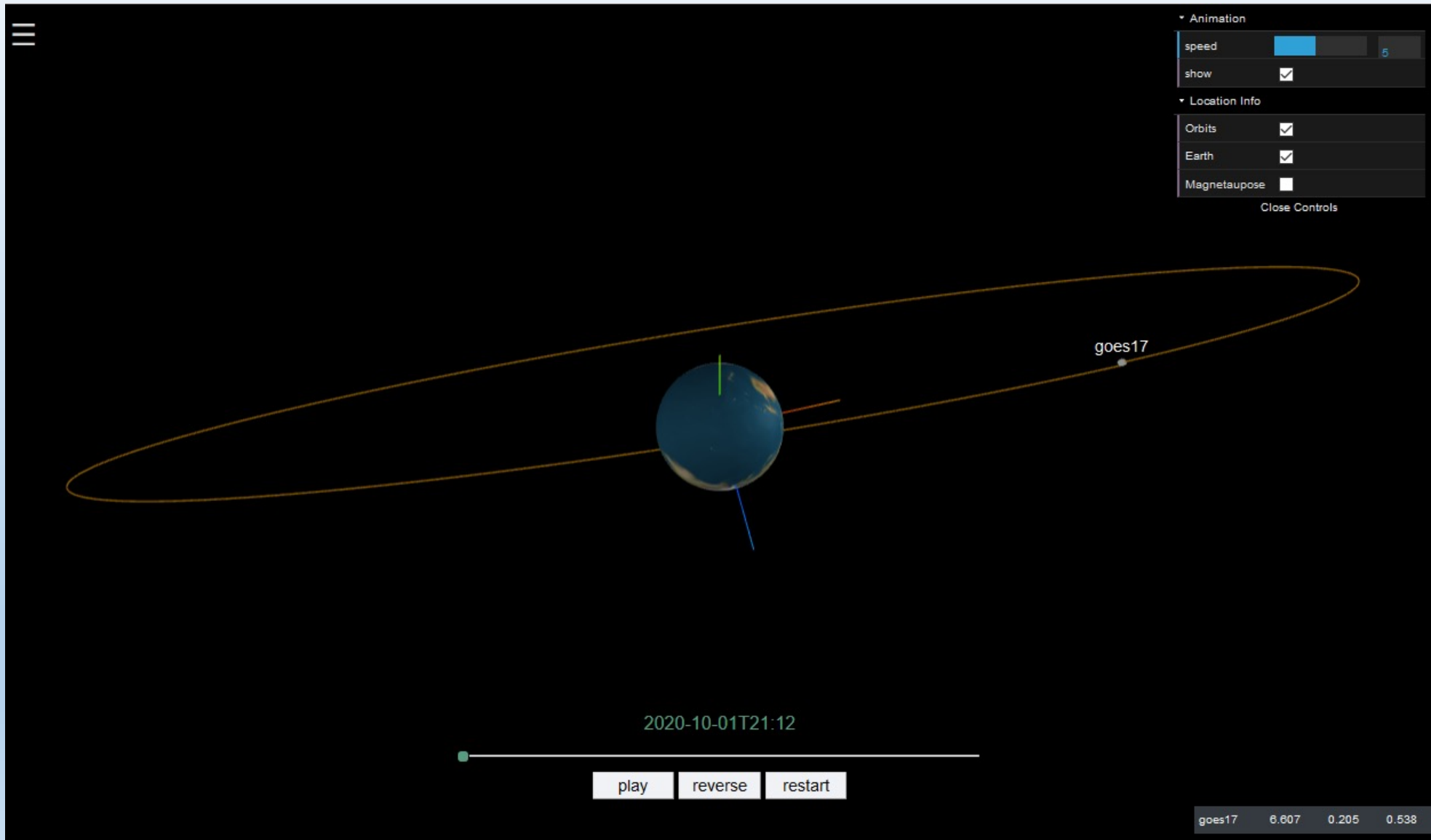
4D Orbit Viewer Prototypes

3D application with Electron and BabylonJS using Typescript



4D Orbit Viewer Prototypes

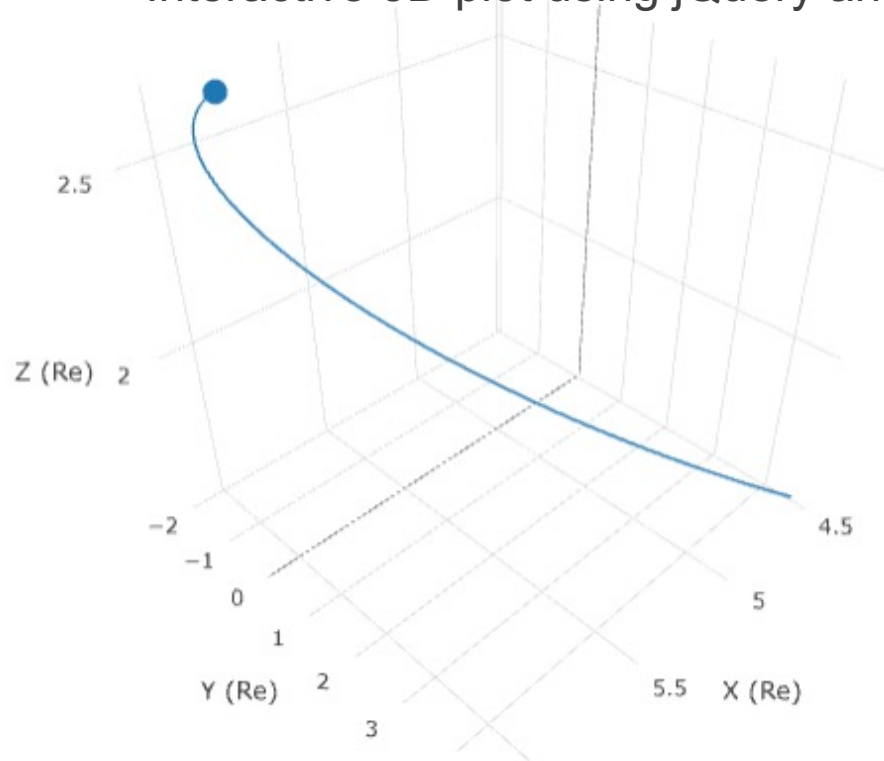
3D application with jQuery and Three.js using TypeScript



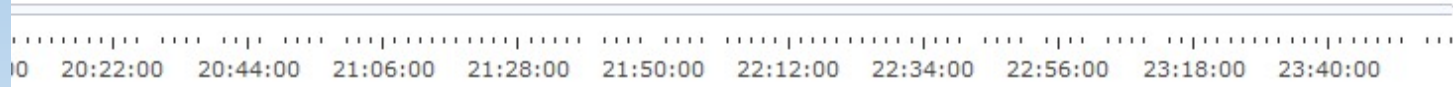
GSE Orbit Plot

GOES-17

Interactive 3D plot using jQuery and 3D Plotly (D3)



Time:20:00:00



4D Orbit Viewer Prototypes

CDAWeb and SSCWeb replacement using OpenMCT Visualization Framework

The screenshot displays the OpenMCT interface for a 4D Orbit Viewer. The interface is divided into several sections:

- Top Bar:** Includes a '+ CREATE' button, status indicators for 'NO TIMER BEING FOLLOWED', 'OFF-LINE STORAGE', and the current time '2020/10/13 17:48:08 UTC'. The 'openMCT' logo is on the right.
- BROWSE Panel (Left):** A tree view showing a hierarchy of data sources. Under 'GOES17_EPHEMERIS_SSC', 'GEO_LON' is selected and highlighted in blue.
- Main View (Center):** Displays a line graph titled 'GEO_LON' with a value of '222.85391'. The y-axis is labeled 'Deg' and ranges from 222.845 to 222.855. The x-axis is labeled 'UTC' and shows dates from 21-01-0202 to 31-01-0202. A cyan line represents the data, with a cyan square marker at the current time. Below the graph is a 3D visualization of Earth with a satellite orbiting it, labeled 'goes17'.
- INSPECT Panel (Right):** Shows the properties of the selected 'GEO_LON' object. It includes a 'Summary View' dropdown, 'PROPERTIES' (Title: GEO_LON, Type: cdaweb), and 'ORIGINAL LOCATION' (Observatories).
- Bottom Panel:** Shows the current time range: 'Start 2020-10-12 17:16:35.000Z' to 'End 2020-10-13 17:46:35.203Z'. It also includes 'Fixed Timespan Mode' and 'UTC' options.

SPDF access protocols

- Files available through HTTPS and FTPS
<<https://spdf.gsfc.nasa.gov/pub/>>
- HAPI <<https://cdaweb.gsfc.nasa.gov/hapi>>
- REST and SOAP web services for
 - CDAWeb <<https://cdaweb.gsfc.nasa.gov/WebServices/>>
 - SSCweb orbits <<https://sscweb.gsfc.nasa.gov/WebServices/>>
- Autoplot <<http://autoplot.org/help#CDAWeb>>
- Other methods such as IDL
<https://cdaweb.gsfc.nasa.gov/alternative_access_methods.html>