LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 26, 2022

LIGO Data Management Plan

(LIGO-M1000066-v26)

LIGO Laboratory
Contact: gwosc@igwn.org

California Institute of Technology
LIGO Project, MS 100-36
Pasadena, CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project, NW22-295
Cambridge, MA 02139
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

LIGO Hanford Observatory
PO Box 159
Richland, WA 99352
Phone (509) 372-8106
Fax (509) 372-8137
E-mail: info@ligo.caltech.edu

LIGO Livingston Observatory
19100 LIGO Lane
Livingston, LA 70754
Phone (225) 686-3100
Fax (225) 686-7189
E-mail: info@ligo.caltech.edu

`http://www.ligo.caltech.edu/`

1 Introduction

1.1 Overview

LIGO Laboratory [1] is the world's premiere facility for studying the Cosmos through the direct observation of gravitational radiation. The LIGO observatories, in operation since 2001, are located in Hanford, Washington [2] and Livingston, Louisiana [3], and are managed and operated by the LIGO Laboratory, a joint enterprise of the California Institute of Technology (Caltech) and the Massachusetts Institute of Technology (MIT) under a cooperative agreement with the NSF. LIGO's second generation detector, Advanced LIGO, began operating and observing gravitational waves alongside Advanced Virgo in 2015 [4, 5, 6]. In the years since, it has recorded gravitational-wave signals from over 90 different compact object mergers [7]. Data from the LIGO observatories have made a tremendous impact on the scientific community. This includes discoveries of never-before-seen astrophysical objects, detailed studies of fundamental physics, publications with thousands of citations, and a long list of awards and recognition.

As part of its responsibilities in operating an NSF facility, LIGO Laboratory has sole fiduciary responsibility for curating its data as long as LIGO is in operation. Should LIGO cease operation, all preserved data will be delivered to a repository to be defined by the NSF. This LIGO Data Management Plan (DMP) follows NSF policies for data sharing and data management [8], and describes how data from the LIGO observations are ingested, stored, shared, and preserved for future use. LIGO is also working with our international partners, Virgo [9] and KAGRA [10], to coordinate broader access to all gravitational-wave data. However, this plan addresses only the LIGO component. The plan is intended to be a living document, updated as needed. The latest version is publicly available at `https://dcc.ligo.org/LIGO-M1000066/public/`.

1.2 Collaborations

To promote the best possible scientific use of LIGO data, the LIGO Scientific Collaboration (LSC) [11] has worked for more than a decade to develop and refine data characterization techniques, data quality vetoes, and extensive analysis pipelines that take the instrument data and produce astrophysical results. The LSC is an open collaboration, allowing scientists from the global community to join based on their willingness to contribute to the LSC Program [12]. In addition to the two LIGO observatories [2, 3], the LSC includes the GEO600 detector in Germany [13].

Many LSC activities are coordinated with the Virgo Collaboration and KAGRA Collaboration, as guided by a Memorandum of Agreement (MoA) [14]. In practice, this forms a larger LIGO/Virgo/KAGRA collaboration (LVK), which includes thousands of scientists and five observatory locations around the world. Data from instruments within this global network are initially made available only to LVK members. The LVK takes responsibility for a number of steps to prepare and annotate the data, including calibration, data characterization, segment production, and noise subtraction, while pursuing a scientific agenda [12]. In addition, all open data releases are carried out in collaboration with the LVK, with responsibilities as described in the MoA.

2 Data Life Cycle

This section describes a data life cycle for LIGO data. The life cycle follows this structure, and is visualized in Figure 1:

- Collection: Data are collected in well-defined observing runs.
- Description: Data are described using an international standard.
- Storage and Preservation: Data are stored at LIGO Data Grid computing centers, with redundant storage at multiple locations.
- Assurance and Analysis: LIGO Laboratory works closely with the LVK to validate data calibration, document data quality, and perform astrophysical analysis.
- Public Data Release: Key data products are made available to the public, to be findable, accessible, interoperable, and reusable (FAIR) [15].
- Long-term Data Archive: We expect the public data release will remain available indefinitely, even after the existence of LIGO Laboratory.

Below, we describe each of these stages.

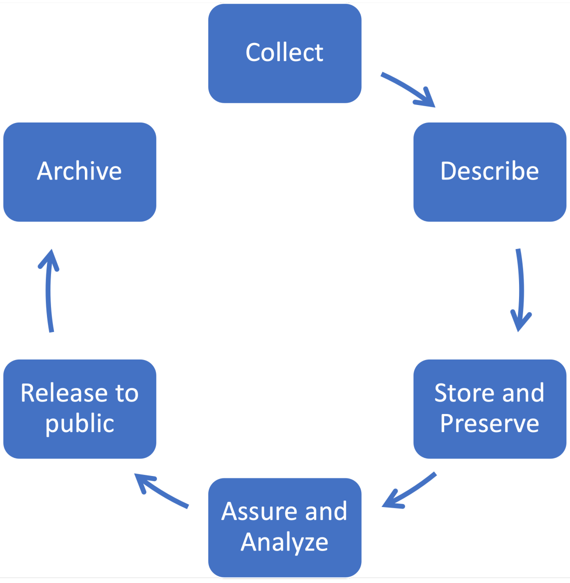

Figure 1: *The LIGO Data Life Cycle*

2.1 Collection

LIGO includes two gravitational-wave detectors in the United States: one near Hanford, Washington, and the other near Livingston, Louisiana. These detectors operate together in a series of observation runs separated by periods of detector improvements and commissioning. The expected timelines and target sensitivity ranges for each run are described in [16, 17] and illustrated in Figure 2, which provides the context for the rest of this document.

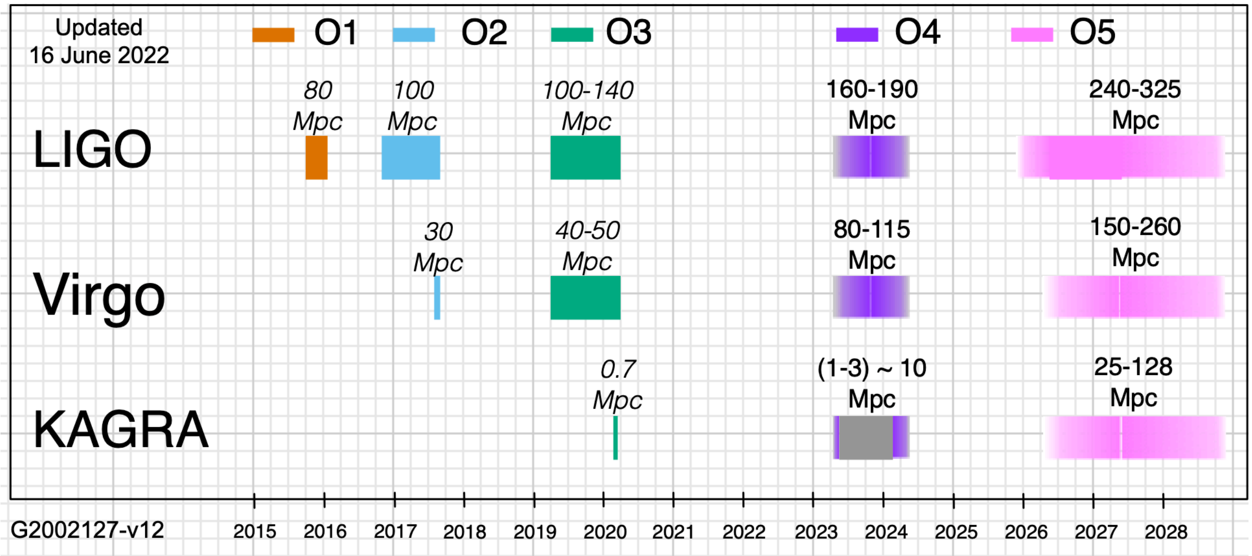

Figure 2: *Currently envisioned timeline of past and future observing runs. The expected schedule and sensitivity may evolve with time. Figure adopted from `https://dcc.ligo.org/G2002127-v12/public/`.*

Figure 2 shows an approximation of the expected "BNS Range" for each observation run. This is an estimate of the distance to which the LIGO detectors could observe a binary neutron star (BNS) merger with a signal-to-noise ratio (SNR) of 8, after averaging over source location. Detector upgrades and/or commissioning efforts between observation runs are designed to increase sensitivity for each run. The LIGO detectors operate within a global network of ground-based gravitational wave detectors including Virgo in Italy and KAGRA in Japan. Data sharing agreements require that data from all detectors in the network be jointly analyzed.

During an observing run, timeseries data from both LIGO instruments are collected and recorded in Gravitational Wave Frame (GWF) files along with appropriate metadata (See Section 2.2). Data are recorded from a large number of sources, including sensors that monitor the state of the instrument (e.g., photo-diodes), sensors that monitor the local environment (e.g., seismometers), and the status of a large collection of digital processing filters and controls. All told, around 250,000 channels of information are associated with each instrument, for a total data rate of ~ 1.5 petabyte per year for both LIGO instruments stored in *raw frames* as described in Section 2.3. While any of these channels may be used for diagnostic purposes, the critical output for astrophysical observations are the so-called "STRAIN" channels, which are calibrated to produce channels representing a sum of instrumental noise and astrophysical signals in units of gravitational wave strain. These "calibrated strain" channels are the main input to all astrophysical observations.

2.2 Description

LIGO data are stored in GWF files that comply with the international standard, "Frame Format for Interferometric Gravitational Wave Detectors" [18]. This format has been used by gravitational-wave detectors around the world since 1997, and nearly all specialized gravitational-wave software accepts this format. The recorded frame files include self-describing metadata. Each LIGO timeseries, or channel, includes fields for key bits of metadata to make the timeseries useful, including a start time, duration, sample rate, and channel name. The channel name uses a standard scheme that puts the instrument name first, followed by abbreviations indicating the interferometer subsystem that generated the data [19]. For example, `H1:DCS-CALIB_STRAIN_C01` refers to data from the Hanford interferometer (H1), generated with the Data and Computing Systems (DCS). In this case, the channel name further indicates this is calibrated strain data using the `C01` version of calibration.

2.3 Storage and Preservation

LIGO manages the following data products with one copy preserved at a LIGO observatory and a second off-site copy at a central data archive at Caltech:

- **Raw Frames** - written by data acquisition systems at each observatory as a 64-s GWF file per interferometer containing all collected channels that are to be preserved. These data are preserved indefinitely for observing runs, but inter-run data are only kept in their entirety until the next observing run.

- **Reduced Data Set** - a subset of the *raw frames* channels at potentially reduced sample rates. These data are primarily used to preserve select data collected between observing (typically $\sim 10\%$) for long-term preservation.

- **Trend Frames** - all of the channels in *raw frames* GWF files are decimated to second-trend (1 sec sample rate) and minute-trend (1 min sample rate). These are preserved indefinitely for both observing and non-observing time periods.

- **Low-latency Calibrated Strain** - generated at each observatory and distributed in low latency for multi-messenger astrophysics trigger generation, these data initially take the form of a one-second duration GWF files per observatory. After initial processing, these data are aggregated for offline processing and long-term storage and preservation.

- **Aggregated Strain** - for offline post-processing and long-term storage and preservation, the one-second *low-latency calibrated strain* data are aggregated to 4096s-duration GWF files. These are preserved indefinitely for both observing and non-observing mode time periods.

- **Recalibrated Strain** - if needed, *aggregated strain* data are regenerated with updated calibration models based on improved knowledge of interferometer performance. This typically only happens for observing-mode time periods, and the data are preserved indefinitely.

- **Analysis Results** - data analysis results from processing one or more versions of strain data. These data are backed up and, starting with O4, will include off-site disaster recovery copies. While in practice all historical analysis results have been preserved to date, this policy only commits to long-term preservation of published results.

- **Published Results** - subset of *analysis results* that are published. These data are preserved for the long-term: currently 20 year commitment from `zenodo`.

In addition to copies stored at the LIGO observatories and central data archive, the *aggregated* and *recalibrated strain* data are distributed to LIGO Data Grid sites in various locations around the world [20] and made globally available via CernVM-FS [21] ensuring that these key data cannot easily be lost or destroyed.

The LIGO data-retention policy is that all data collected during observing runs are preserved indefinitely, and that data collected between runs are kept for 6 months into the following observing run, after which time they are pruned back for longer-term storage and preservation. In particular, *raw frames* are preserved for 1% of the time (1,000 contiguous seconds every 100,000 seconds) at full resolution along with a *reduced data set* specified by the LSC Detector Characterization working group for data collected between observing runs. In very rare cases, *raw frames* may also be preserved for segments around an outstanding astrophysical event that occurs between observing runs, at the request of the LSC spokesperson. In addition, all *trend frames* and *strain* data are preserved regardless of observing run status. For long-term preservation, observing run data are physically write protected at both geographically separated locations to prevent accidental or malicious loss of data.

2.4 Assurance and Analysis

Data validation and analysis are performed by LIGO Laboratory staff working in close collaboration with LVK working groups. Data assurance and analysis steps include calibration [22, 23], detector characterization [24], and astrophysical analysis (e.g., [7]). The close coupling between these activities ensures that final data products are high quality and well documented. This work is done on computing resources [25], including the LIGO Data Grid and other computing centers, which have access to all LIGO data described in Section 2.1, with data access described in [26].

2.5 Public Data Release

Following a proprietary period by the LVK, key data products are made available to the public, as described in Section 3, and on the Gravitational Wave Open Science Center (GWOSC) website at `https://gwosc.org`. The public data release follows FAIR data principles [15]:

- FINDABLE: Timeseries data are discoverable through the GWOSC web server, which includes both human-readable and machine-readable interfaces. The GWOSC website also includes documentation and metadata for all data products. Data analysis products are uploaded to `zenodo` [27] and tagged to the LVK community, making them easy to find.

- ACCESSIBLE: In most cases, both timeseries and analysis results may be directly downloaded through the discovery interface, or accessed via CernVM-FS [21] and a network data server (NDS2) [28].

- INTEROPERABLE: LIGO timeseries data are released in two international standards: HDF5 [29] and Gravitational Wave Frames [18]. Data from KAGRA and Virgo are released in identical formats. The GWOSC website provides a list of open source tools and libraries that are designed to work with these formats, many of which are used in LVK analyses.
- REUSABLE: The GWOSC website provides rich documentation and metadata for all released data, and the LVK hosts free, annual workshops to train interested persons on using public data. All data are released with a license, ensuring they may be freely reused with minimal restrictions. To facilitate the re-use of data the LIGO Laboratory has an open source software license policy for the tools used to generate and analyze those data [30]. More broadly the LSC is considering adopting a similar policy.

2.6 Long-term Data Archive

The public data release, as described in Section 3, contains all of the key data and metadata required to make use of LIGO observations for a wide range of scientific activities. For this reason, the public data release is exactly the data that should be preserved in a long-term archive on time-scales of ten years or more. Moreover, the calibrated strain data grow at a rate of only ∼10 TB per year, and so may be readily copied and stored with current computing standards.

Currently, redundant copies of the public strain data are maintained at LIGO Data Grid sites, and we will continue to maintain this archive as long as LIGO Laboratory exists. In the event that LIGO Laboratory dissolves, the public strain-data archive should be copied to a suitable alternative repository. Similarly, published analysis results are archived on `zenodo`, whose user agreement promises to maintain this release for at least 20 years [27].

3 Data Release Policy

3.1 Timing of Data Releases

This section describes the planned timing of data releases from LIGO, Virgo, and KAGRA, and is summarized in Table 1 and Figure 3. Subsequent sections describe the various types of data releases in more detail.

- Limited information from low-latency alerts will be available starting minutes after each transient detection, as described in Section 3.2 and the Public Alerts User Guide [31].
- Strain data and data-quality segments of up to 4096 seconds around detected transients are made available at the time of publication in a refereed, scientific journal (see Section 3.4.3). In most cases, these data releases also include some analysis results, such as Bayesian parameter estimation posterior samples.

- Full strain data releases – including final, calibrated strain and data quality segments for an entire observing run, as described in Section 3.4.4 – are made available in blocks of up to six months. These data are released to the public after a period of 18 months from the end of each observing epoch. For example, the first 6 months of the O3 observing run (O3a) includes data recorded from April through September, 2019. These data were made publicly available approximately 18 months after the end of the observing period, in April of 2021. The same 18-month latency period will apply to O4 strain data. This is shown in Figure 3.

We expect these target dates to apply to data from O4. While we strive to make these target dates for data releases, it is possible that unexpected delays can occur. We will keep the scientific community informed of any such delays.

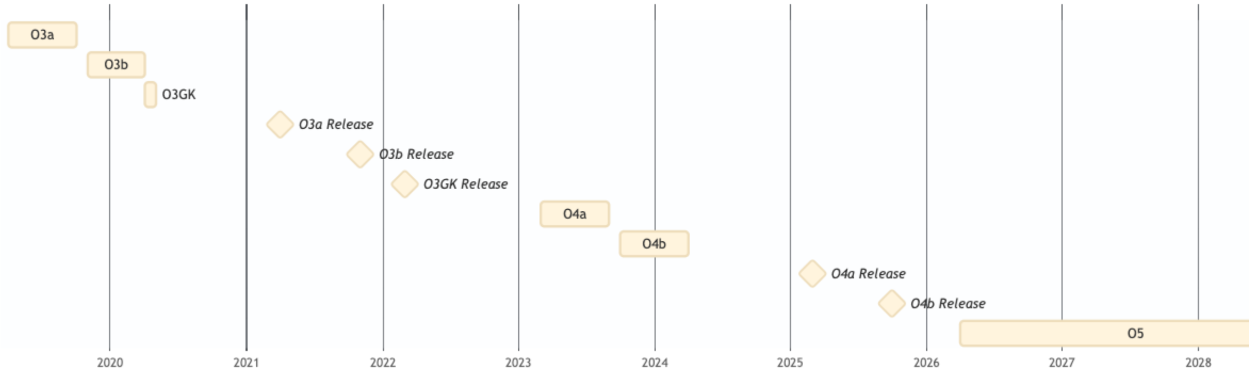

Figure 3: *Possible timeline for full-strain data releases. Observing runs are divided into 6 month blocks (rectangles). For example, O3 was divided into two 6 month periods, called O3a and O3b. The O3GK label refers to a 3 week period where GEO and KAGRA observed together. For O3 and O4, strain data releases (diamonds) are planned for 18 months after the end of each 6 month observing block.*

3.2 Low-latency Public Alerts

Starting with O3, LIGO and Virgo began producing public low-latency alerts for detected transient GW events. A primary purpose of the public alerts program is to support a rapid follow-up program that enables astronomers to seek potential counterparts to GW transients. The low-latency public alerts are designed to provide enough information to enable prompt electromagnetic follow-up observations. While the alerts do not include any release of strain data, they do include a number of preliminary analysis results for each event, including an event time and estimated source position. A list of alerts may be accessed through the online service GraceDB [32], and a technical description is available in the Public Alerts User Guide [31]. In general, trigger information in GraceDB is not curated or updated beyond a few days after the initial alert. Final analysis results are curated in the Event Portal on the GWOSC website as described in Section 3.5.

Release type	Section	Notes	Timing
Low-latency alert	3.2	Enables astronomical follow-up of GW transients. Includes sky position, but no strain data.	Starting minutes after each detection
Published events	3.4.3	Strain data at times of GW transients	Released at time of publication in a scientific journal
Analysis products	3.5	Includes parameter estimation samples and other details of observed GW transients	Typically released at time of publication in a scientific journal
Full strain data	3.4.4	Strain data for an observing period	Released in 6-month blocks, after 18 months from the end of each block

Table 1: *Summary of different types of data releases and their associated timing.*

3.3 Gravitational Wave Open Science Center

The Gravitational Wave Open Science Center (GWOSC) [33], formerly known as the LIGO Open Science Center, was created to provide public access to gravitational-wave data products [34]. Starting in 2011, the LIGO Scientific Collaboration sought input from the astronomy and astrophysics community in the form of surveys and a workshop to guide development [35]. Since that time, LVK scientists have been working to build the tools and infrastructure for public release of these data.

Today, the collaborations running LIGO, Virgo, GEO600, and KAGRA have all agreed to use GWOSC services as the primary access points for public data products. This collaborative approach benefits users by creating a uniform interface to access data from multiple observatories, and provides cost savings to the various observatories by sharing the tools, services, and human resources. At this time, public data products from LIGO, Virgo, KAGRA, and GEO600 are accessible from the GWOSC website. The remainder of this section provides a short overview of these data products and services.

3.4 Strain Data and Segments

3.4.1 Strain Data

Astrophysical observations with LIGO, Virgo, and KAGRA are derived from timeseries channels representing calibrated strain, or $h(t)$. These calibrated strain channels represent the main scientific data product of the experiment. The versions used for analysis are recorded at a sample of 16384 Hz, leading to a data rate of around 450 MB per hour of livetime for each detector.

For ease of access in public releases, the final versions of the strain vectors are packaged into both Gravitational Wave Frame files (GWF), a standard within the gravitational wave community, and the more broadly used HDF5 format. Currently, only times when the instrument is collecting valid strain data (i.e., times passing CAT 1 data quality) are included in public releases. For the first three observing runs of Advanced LIGO, we have released both the full-sampled, 16384 Hz strain vector used for LVK publications, and for convenience to the user, a smaller strain vector down-sampled to 4096 Hz.

3.4.2 Data Quality Segments

Important pieces of metadata accompanying strain data are segment lists representing times when instruments are operating in a known good state. In a typical observing run, duty cycles might range from 50 - 90%, so knowing when an instrument was operating (or not) is a very basic first step in any investigation.

In addition, an important component of any search for astrophysical signals involves labeling times when instrument performance is limited by any of a wide range of potential issues. These may include times plagued by excess environmental noise, interferometer angular misalignment, power saturation in photo-diodes, or other factors making the data significantly harder to interpret. LVK working groups classify and identify these times. For transient searches, data quality issues are categorized as CAT 1, CAT 2, or CAT 3, in order of decreasing severity [36]. Times of good quality in each of these categories are represented as segment lists, and are released along with any public data release. Segment lists are also used to represent times when simulated signals, known as hardware injections, are intentionally added to instrument data for testing purposes.

Data quality and hardware-injection segment lists are shared with the public through the “Timeline” feature of the GWOSC website. Segment lists may be downloaded in ASCII or JSON format, or displayed in an interactive plot in a web browser. The same segment lists are also represented in a bitmask that is packaged inside files containing the strain vector so that data quality and hardware injection segments may be accessed in the same place as the corresponding strain data.

3.4.3 Data Around Published Events

When a discovered gravitational wave signal is published in a scientific journal, the LVK releases a number of data products associated with each detection claim. These include strain data around the time of transient events, as well as data quality segments, as described in sections 3.4.1 and 3.4.2. These data products are made available through the Event Portal on the GWOSC website. Each release includes 4096 seconds of strain data around the event time, if available, so that open data users can assess noise levels in nearby times.

3.4.4 Full Strain Data Releases

The “full” strain data release represents the final data release for an observing period. These releases include the final version of calibrated strain data for an extended period of time – typically 6 months (see Figure 3). For example, there is one full-strain data release for each of the first two LVK observing runs (O1 and O2), as well as three more data releases for

O3, corresponding to the first six month block (O3a), the second six month block (O3b), and a three week GEO/KAGRA run (O3GK). Along with the calibrated strain, these data releases include data-quality segment information and documentation.

Beginning with O3, we expanded the full data release by adding several versions of the calibrated strain data, in addition to the final version. These may include alternative versions of calibration, and/or data with or without some noise subtraction methods applied. The O3 data release includes an "alternate calibration" release with several of these variants. These additional channels are accessible through CernVM-FS [21] for cluster access, and a network data server (NDS2) for individual computers, both described in Section 3.6.

3.5 Analysis Results

In addition to strain data and associated segments, a number of analysis products produced by LVK working groups are publicly available. These include Bayesian posterior samples from parameter estimation pipelines, data represented in figures in publications, and skymaps of detected events. These data products are uploaded into repositories where experts may maintain and curate them, including the LIGO Document Control Center (DCC; [37]) and Zenodo [27]. The LVK has recently adopted Zenodo as the preferred repository for published results, and now posts results to the Zenodo LVK Community [38]. Where possible, these analysis products are linked from the GWOSC website in documentation supporting the associated data release to make these data products easy to find.

One important data analysis result is a catalog of detected gravitational wave signals. Currently, GWOSC hosts the "Gravitational Wave Transient Catalog" (GWTC), a curated list of gravitational-wave events identified by the LVK, including estimates of the parameters for each source. This catalog is available in a number of different formats, and is periodically updated by the LVK working groups. For archival purposes, snapshots of the catalog are uploaded to Zenodo after each update [39].

3.6 Data Discovery and Access

The GWOSC website provides easy point-and-click access to discovering all available public data products. In addition, the website supports queries through a JSON API, which can allow scripts to search for, discover, and download data products associated with a given time or event. API access is supported by a lightweight Python client, called `gwosc` [40], available through `pypi` [41].

To discover strain data and analysis products for a given event, GWOSC provides the "Event Portal" user interface. The Event Portal provides a number of tools to discover sets of events, including:

- Catalogs listing sets of discoveries and source parameters, including the "Gravitational Wave Transient Catalog" (GWTC) maintained by the LVK working groups
- Queries for events based on their parameters
- Sortable lists of all events in the database

In addition, strain data are now made available through two additional interfaces: CernVM File System (CernVM-FS) [21] and a Network Data Server (NDS2) [28]. CernVM-FS has been adopted internally by the LVK to provide access to data on a range of computing systems. Once installed and configured, CernVM-FS gives a user access to remote data products as if they are available through a local disk drive. Data files are automatically cached locally so that repeated access is offered without multiple downloads. This interface has allowed users to access public data products on general computing clusters, and thereby is enabling large scale scientific analysis. NDS2 is used internally by the LVK for streaming data access; it allows a user to easily access small segments of data through a client [28, 42, 43].

3.7 Software and Documentation

GWOSC development work emphasizes ease of use of LIGO, Virgo, and KAGRA data [34]. A number of services are provided on the GWOSC website to lower the bar for entry into data analysis as much as possible. All data sets are accompanied by documentation on the GWOSC website. This documentation includes data set descriptions, overviews of available time and sensitivity, and links to relevant publications.

The website also includes a number of software tutorials, which emphasize finding, loading, and plotting strain data, as well as basic analysis tasks such as identifying signals through matched filtering. The tutorials include descriptive text and software examples to provide a “hands-on” introduction to working with gravitational-wave data, including many of the tasks described in [44]. In addition to tutorials available on the website, GWOSC provides a collection of links to a number of open-source libraries that are available for working with gravitational wave data, many of them in active use by the LVK. These include well-documented Python packages such as `gwpy` [43], `pycbc` [45], and `bilby` [46], which all include their own descriptive web pages and usage examples. Currently, the LIGO Laboratory has an open-source software license policy for the tools used to generate and analyze LIGO data [30], and the LSC is considering formalizing the currently common practice of using an Open Source Initiative (OSI)-approved license for all collaboration software [47].

Starting in 2018, GWOSC began hosting “Open Data Workshops” to encourage LVK members to meet face-to-face (or screen-to-screen) with researchers interested in working with gravitational-wave data. For these workshops, data analysis experts prepare lectures and software examples, and spend time working with participants. All materials from these workshops, including lecture videos and software examples, are available as part of the tutorials package offered on the GWOSC website.

4 Summary of Resources

This document describes how LIGO data are managed and made available to a global scientific and educational community as well as the general public. To learn more about working with LIGO data, a variety of resources are available:

- Gravitational Wave Open Science Center [33]

- LIGO-Virgo-KAGRA Public Alerts User Guide [31]
- LIGO-Virgo-KAGRA Publications List [48]

This document will be updated occasionally as the LIGO detectors and policies evolve. Over the next few years, we expect the sensitivity and impact of the LIGO detectors to continue to grow, as described in the observation schedule [17].

References

[1] `https://www.ligo.caltech.edu`. LIGO Laboratory.

[2] `https://www.ligo.caltech.edu/WA`. LIGO Hanford Observatory Home Page.

[3] `https://www.ligo.caltech.edu/LA`. LIGO Livingston Observatory Home Page.

[4] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[5] J. Aasi et al. Advanced LIGO. *Class. Quant. Grav.*, 32:074001, 2015.

[6] F Acernese et al. Advanced virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, dec 2014.

[7] R. Abbott et al. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run. 11 2021.

[8] `https://www.nsf.gov/bfa/dias/policy/dmp.jsp`. Dissemination and Sharing of Research Results - NSF Data Management Plan Requirements.

[9] `http://public.virgo-gw.eu/the-virgo-collaboration/`. Virgo Collaboration Home Page.

[10] `https://gwcenter.icrr.u-tokyo.ac.jp/en/organization`. KAGRA Collaboration Home Page.

[11] `https://www.ligo.org`. LSC Home Page.

[12] `https://dcc.ligo.org/LIGO-M2100100/public`. 2021 program. *LIGO DCC*, M2000130-v4, 2021.

[13] J. Abadie et al. A Gravitational wave observatory operating beyond the quantum shot-noise limit: Squeezed light in application. *Nature Phys.*, 7:962–965, 2011.

[14] `https://dcc.ligo.org/LIGO-M1900145/public/`. Memorandum of Agreement between VIRGO, KAGRA, and LIGO. 2019.

[15] Mark D. Wilkinson et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:160018, 2016.

[16] B. P. Abbott et al. Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA. *Living Rev. Rel.*, 21(1):3, 2018.

[17] `https://observing.docs.ligo.org/plan/`. LVK Observing Plan.

[18] `https://dcc.ligo.org/LIGO-T970130/public`. Specification of a Common Data Frame Format for Interferometric Gravitational Wave Detectors (IGWD). 2019.

[19] `https://dcc.ligo.org/LIGO-M080375-v1/public`. LIGO Acronym List. 2020.

[20] `https://computing.docs.ligo.org/guide/computing-centres/ldg/`. LIGO Data Grid.

[21] `https://cernvm.cern.ch/fs/`. CernVM File System.

[22] Ling Sun et al. Characterization of systematic error in Advanced LIGO calibration in the second half of O3. arXiv:2107.00129, 2021.

[23] Ling Sun et al. Characterization of systematic error in advanced LIGO calibration. *Classical and Quantum Gravity*, 37(22):225008, oct 2020.

[24] Derek Davis et al. LIGO detector characterization in the second and third observing runs. *Class. Quant. Grav.*, 38(13):135014, 2021.

[25] `https://computing.docs.ligo.org/guide/grid/`. IGWN Computing Grid.

[26] `https://computing.docs.ligo.org/guide/data/`. Accessing IGWN data.

[27] `https://zenodo.org`. Zenodo. 2022.

[28] `https://computing.docs.ligo.org/guide/data/#nds`. IGWN Computing Docs on NDS2.

[29] `https://www.hdfgroup.org/solutions/hdf5/`. HDF5: High-performance Data Management and Storage Suite.

[30] `https://dcc.ligo.org/LIGO-M1500244/public`. LIGO Laboratory Open Source Software License Policy.

[31] `https://emfollow.docs.ligo.org/userguide`. LVK Public Alert User Guide.

[32] `https://gracedb.ligo.org`. GraceDB.

[33] `https://gwosc.org`. GWOSC.

[34] Rich Abbott et al. Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo. *SoftwareX*, 13:100658, 2021.

[35] `https://dcc.ligo.org/LIGO-P1100182/public/`. Community Input for LIGO Open Data. 2011.

[36] Rich Abbott et al. Open data from the first and second observing runs of advanced ligo and advanced virgo. *SoftwareX*, 13:100658, 2021.

[37] `https://dcc.ligo.org/dcc`. LIGO Document Control Center.

[38] `https://zenodo.org/communities/ligo-virgo-kagra/`. Zenodo LVK Community.

[39] `https://doi.org/10.5281/zenodo.6344403`. Evnet Portal Snapshots.

[40] `https://pypi.org/project/gwosc/`. GWOSC Python Client. 2021.

[41] `https://pypi.org/`. Python Package Index (PyPI).

[42] `https://www.gw-openscience.org/O3/O3alt/`. O3b data release documentation.

[43] `https://gwpy.github.io`. GWpy.

[44] Benjamin P Abbott et al. A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals. *Class. Quant. Grav.*, 37(5):055002, 2020.

[45] `https://pycbc.org`. PyCBC.

[46] `https://lscsoft.docs.ligo.org/bilby/index.html`. Bilby.

[47] `https://opensource.org/licenses`. Open Source Initiative Licenses.

[48] `https://pnp.ligo.org/ppcomm/Papers.html`. LVK Publication List.