

LIGO Hanford Observatory



LIGO Livingston Observatory



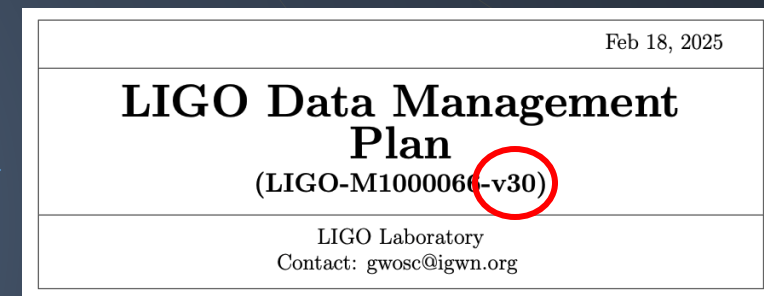
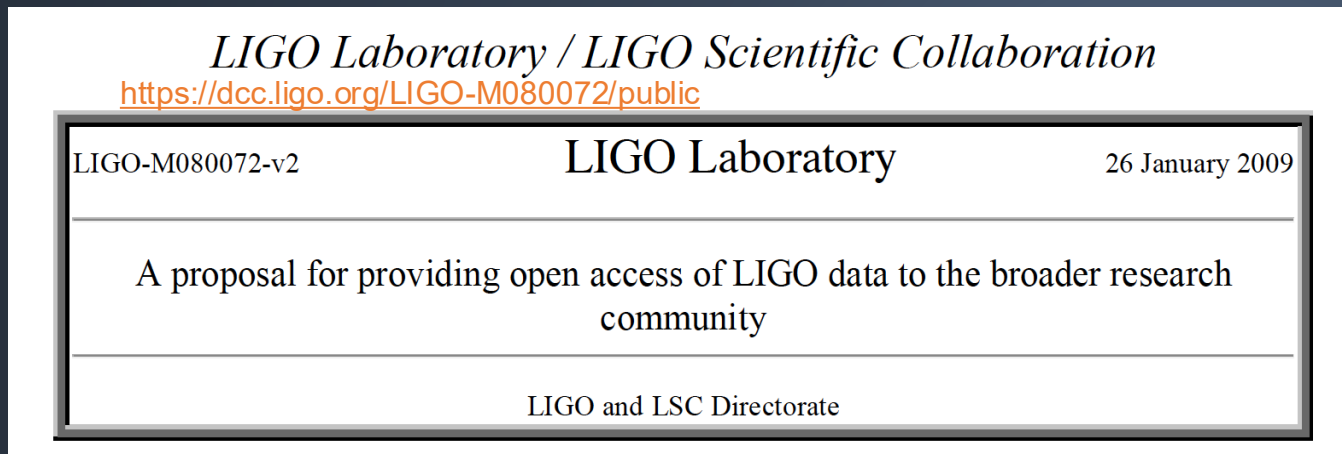
Cosmic Explorer Data Policy Discussion

Dave Reitze, LIGO Laboratory, Caltech

What Do I Know About This?



- I was the LSC Spokesperson in 2008 when the National Science Foundation required LIGO to develop and implement an Open Data Policy for LIGO data



<https://dcc.ligo.org/LIGO-M1000066/public>

- ▶ This eventually evolved into what is now the LIGO-Virgo-KAGRA Open Data Policy
 - Gravitational-Wave Open Science Center ([GWOSC](https://gwosc.ligo.org))

- 2014 AAAS Meeting Session on Open Data: <https://dcc.ligo.org/LIGO-G1400060/public>



What Should a Data Policy Should Hope To Achieve?



- Maximize the scientific productivity of the group/collaboration/facility/... that produces and analyzes the data
- At the same time, incentivize the researchers who produce and analyze the data
 - Or, at the very least, not disincentivize them!
- Conform to agency requirements as well as national/international norms (and in some cases laws) on the availability/curation/preservation of data
 - E.g. Findable, Accessible, Interoperable, Reusable (FAIR), "Data in Perpetuity"

Major Implementation Issues Facing Public Data Release Programs



- *What is the nature of the data, the size of the data stream, and what analysis tools are needed to manipulate and analyze the data?*
- *What data products should be released? When should the data products be released?*
- *Who are the broader research community of potential users?*
- *What is the added cost of releasing the data and who pays?*
- *What are the implications regarding intellectual property?*
- *What are the implications of making data publicly available which is 'owned' by international collaborations?*

Data Policy Issues Specific to CE



- Negotiating data release policies with international partnerships: Einstein Telescope (and LIGO-India)
 - For some science goals, CE cannot 'go it alone'
- Low latency alerts are, by definition, a zero-delay data release policy
 - What types of information to release? The LVK has only recently evolved to a point where we release coarse mass bins for compact object mergers
- Is a 'Data Verification Period' needed -- how long does it realistically take to clean, calibrate, and package the data?
 - This sets the minimum release time for bulk data releases

Lessons from LIGO-Virgo-KAGRA



- Once collaborations have privileged access to data, it is difficult to switch to a fully open policy
 - Collaborations get comfortable with having privileged access!
- Even with a proprietary data period, some collaboration members may fear getting scooped once the data is released.
 - This may be true for certain analyses, however the collaboration is the power user of the data. It's likely that collaboration papers will be the more cited papers.
- No matter what data policy you eventually settle on, not everyone will be happy with it ;-)
 - Overarching guideline: Maximize the scientific productivity of the group/collaboration/facility/... that produces and analyzes the data!