

XLDB Use Cases

Collecting data on the use of extreme data

Daniel L. Wang
SLAC/XLDB/LSST
danielw@slac.stanford.edu



XLDB: who we are

Practitioners of extremely large data sets.

Computational science, data management,
database research, etc.

Database vendors, professors, data people at
places with extreme data (e.g., SLAC/LSST,
Google, FB, eBay, NIH, Chevron, Sears, CERN)



XLDB Use Cases Wiki

Mission: Collect and curate examples of problems encountered and solutions in use for extremely large datasets.

- We're used to doing X. How do I build a solution that lets me do X at 10-100x my current scale?
- We want to tell the world how we manage our huge data.
- We are struggling with our data scale. Can we collaborate with others with similar problems?
- What scalability problems can our research solve?
- What are some good research topics in big data?
- Our field seems to be under-served in big data. Big data solutions assume all these things that don't hold for us. Why can't they make something we can use?
- If we knew where people were struggling with data, we could make better solutions for them.
- Why aren't scientists using the same tools as Google, Facebook, etc.?





Page [Discussion](#)

Read [Edit](#) [View history](#)

Main Page

XLDB Use Cases [\[edit\]](#)

Welcome! This wiki is a repository for descriptions of how people/organizations/institutions/projects are managing and analyzing (or want to) large scale data. It exists to ease the sharing of real-world problems, solutions and lessons learned with large-scale data. It is moderated by the XLDB core team (you can reach us through email xldb-admin@slac.stanford.edu).

To read more, including what are the benefits of studying these use cases, what use cases qualify, and how to submit a new use case, see [Hints & Guidelines](#).

By category - slicing and dicing in different ways... [\[edit\]](#)

- Alphabetically: [projects](#) | [use cases](#) | [data sets](#)
- By data model: [array](#) | [graph](#) | [gridded](#) | [relational](#)
- Industry/science: [industry](#) | [science](#)
- By domain: [astronomy](#) | [biology](#) | [earth science](#) | [finances](#) | [geoscience](#) | [healthcare](#) | [oceanography](#) | [oil & gas](#) | [physics](#) | [web analytics](#) | [seismology](#)
- Others: [correlations](#) | [spatial](#) | [subsetting](#) | [time series](#)

(To see all categories currently defined, see [All categories](#). Interested in additional categories? Just add them to [All categories](#) and start using it.)

Appendix [\[edit\]](#)

[Glossary](#)

http://xldb.org/use

This page was last modified on 8 May 2017 at 00:32.

This page has been accessed 209,646 times.

Content is available under [Creative Commons Attribution-ShareAlike 3.0 Unported](#).

[Privacy policy](#) [About XLDB Use Cases](#) [Disclaimers](#)



Help us help you

- Understanding usually precedes solutions.
- Problem statements guide research and products.
- Share your problems and your (incomplete) solutions.
- Wiki URL: <http://xldb.org/use>
- All welcome to add/modify pages. See me or send email to xldb-admin@slac.stanford.edu for an account.

