The Keck Observatory Archive (KOA) Hurtles Into the 2030s

G. Bruce Berriman (Caltech/IPAC- NExScI) on behalf of the KOA Team

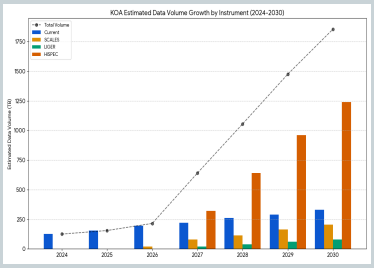
KOA Today: A Private Observatory as a Public Resource

- **11 Active Instruments, data ingested on acquisition**
- **31,300 Nights**
- **7 Million** Level 0 (Raw) Files
- **6 Million** Pipeline processed files

KOA Tomorrow: Challenges for the 2030s

- **10x Surge in Data Volume with 3 new instruments**
- **Higher Cadence Observing and Bigger Files (>1 GB)**
- **Same people!**

Meeting the Challenges: Automation, adaptation, new S/W architecture (including AI)

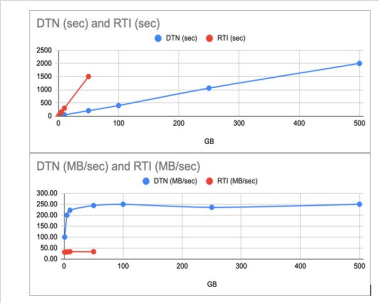


Anticipated data growth at KOA

- **New S/W Infrastructure In place**
- **Plotly-DASH Query Infrastructure and R-trees for fast access**
- **Real Time Ingestion triggered on file generation**
- **Observers Data Access Portal: authenticated access to protected data**

High data volumes, high cadence observing: Dedicated Data Transfer Node

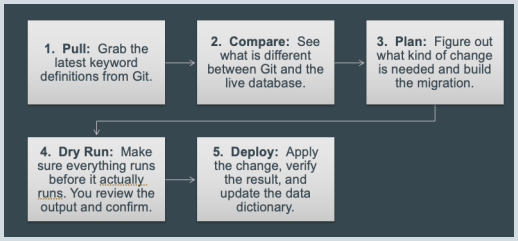
- Adaptation of data transfer for transfer of Gemini Speckle Data to IPAC
- Internet2 required
- HP machine running RHEL 9.6; 24 cores 2.6 GHz; 263 GB RAM
- Solid state disks
- 40Gbit/s bandwidth (under upgrade)
- Fast Data Transfer (FDT; fast-data-transfer.github.io).
- Logs transfer errors, verifies completeness before DB commit
- 30 GB tar file ingested in 2 min. x7 performance improvement



Performance of DTN vs Real Time Ingestion

Automated Synchronization of Keyword Repository With KOA Oracle Database

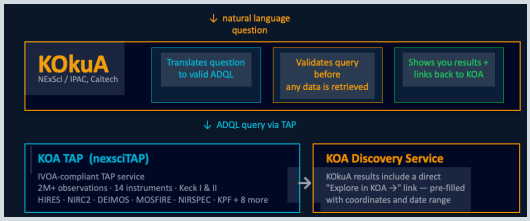
- Replaces error prone manual SQL Scripts
- Synchronizes Keyword Git Repository with the KOA Oracle Database
- Gemini AI used to to build
- Reconstruct history of keyword changes



Flow Diagram of Keyword Automation in Action

Natural Language Queries: Exploring Agentic AI

- KOkuA builds complex, multi-component queries without rebuilding ADQL.
- TAP server underpins architecture.
- 3-Layer Harness Engineering Harness > Context > Prompt) to eliminate hallucinations.
- Evaluates schema constraints and validates astronomical coordinate targets prior to execution.



Architecture of KOKuA

KOA is a collaboration funded by NASA between the W.M. Keck Observatory and the NASA Exoplanet Science Institute