

Tomo-e Gozen データプラットフォームの開発 3

○瀧田 怜¹ 酒向 重行¹ 森 由貴¹ 大澤 亮² Tomo-e Gozen コラボレーション³ (1.東京大学 2.国立天文台)

東京大学木曽観測所では 105 cm シュミット望遠鏡に広視野動画カメラ Tomo-e Gozen を搭載し、2 fps の動画サーベイをもとに超新星等の突発天体や小惑星等の高速移動天体の探査を行っている。この動画データの生成レートは一晩あたり最大で約 30 TB に達するため、長期アーカイブとしては時間方向に圧縮した二次元画像データ (スタック画像) のみを保存している。しかしスタック画像であっても年間で 200 TB 弱のペースでデータが蓄積されるため、木曽観測所内でアーカイブし続けるのは困難な状況である。そこで我々は、東京大学 柏キャンパスに設置された「データ活用社会創成プラットフォーム mdx」に Tomo-e Gozen のデータプラットフォームを構築する計画を進めている。

2025 年度から mdx の大容量ストレージ (HDD) の利用料が無償化されて十分なストレージ容量を確保できたため、Tomo-e Gozen が本格稼働を開始した 2019 年 10 月以降の全スタック画像の転送を行った。転送したデータのメタデータも順次データベース化することで、2025 年 12 月にデータアーカイブシステムとしての公開を開始した。転送済みのデータ総量は約 900 TB、ファイル数にして 1 億以上の大規模データアーカイブとなっている。データベースは列志向型の Apache Arrow 形式とし、ファイルフォーマットに Arrow IPC を用いることでストレージベースのデータベースでありながら座標検索にかかる時間を 10 秒程度に短縮している。また 2025 年より、毎晩一時間の動画サーベイデータについても mdx での長期アーカイブを開始している。この動画データについても共同研究者向けには上記のアーカイブシステムから取得できるようにしている。本講演では新アーカイブシステムと今後の開発方針について紹介する。