

議論

2020.3.30@Zoom

データアーカイブの意義・成果・楽しいこと

u 意義・強み

- u 研究着手の高速化・追観測（補足データ取得）の代替
 - u 他波長・装置，他物理量，他天体
- u 時系列・突発天体のリファレンス・イベント前情報
 - u 事前には場所を特定できない
 - u 長期（数十年）規模変化を追える
- u 既存研究を超えるサンプルによる統計的研究
- u レアな天体探査
- u 新技術による新たな解析（生データが必要？）
- u 天文教育のレベル維持に不可欠（~1/3がアーカイブベース）
- u 波長横断して天体の性質を確認できること
- u 観測の専門家でなくても利用できること
- u クロスチェック、エビデンス

挙げた成果，挙げられる成果

u あげた成果

- u レアな天体（レンズQSO，Hy/ELIRG，）
- u 統計的（サイズ進化，LF，ACF，分子吸収系）
- u イベント同定（超新星の親星・質量放出，爆発日，重力波源，矮新星，QSO）
- u 天文（学校）教育の論文

u あげられる成果

- u 超新星爆発前親星の同定
- u 機械学習・反復学習
- u 画像decomposition, super-resolution
- u 天文（学校）教育研究

データアーカイブの要件 1

u データ

u 生データ (FITS?)

- u 生データに戻ることはまだだが，新解析やイベント前情報など保全は必要
 - u 保存すべきデータの選択？ 戦略 vs 共有資産
- u 再解析の必要性
 - u ソフト，解析情報とペアである必要がある
(今後の衛星ミッションはデフォルト)

u 処理済みデータ

- u 最低 1 次処理済み画像
- u 解析可能なメタデータ
- u 共通のもの．個々人のファインチューニング・
全員にとってscience readyということにはなりづらい．
- u データを作った人たちで (アーカイブのことを考えて) 残していく必要がある

u 均質性

- u 希望) プロダクト・装置を横断した統一・共通化された
データフォーマット + 解析方法での公開 (鳥羽・秋山・吉村+)

データアーカイブの要件 2

u 情報

- u 観測コンディション

- u 品質、誤差情報：利用者が評価できる情報

- u あるレベルのQA

- u 解析パイプライン，解析方法，解析用メタデータ

- u ドキュメント

- u ヘルプデスク

データアーカイブの要件 3

u 機能

- u 候補を絞れること（座標検索，測定条件，観測有無）
- u QLができること，品質の俯瞰（データ選別，意味があるか判断）
- u 長期安定運用
 - u 数～数十年の挙動
 - u 長期的な解析の再現性
- u 希望）ビューワ
- u 希望）多波長カットアウト
- u 希望）複数天体の同時カットアウト
- u 希望）数値情報取得
- u 希望）スクリプトからの検索I/F（e.g., astroquery）

ほか

- u 論文誌での情報公開コントロール（カタログ，ソフトウェア）
- u データ処理リソースの確保 à 共同利用機関に期待
- u データ共有PIの貢献（CANFAR?）
- u 公開準備作業の支援，系統化
- u 継続的なデータアーカイブを活用していく情報交換



提案作成

- u データ提供者とデータアーカイブの役割、ユーザのサポートに対して共通理解と合意を作る
- u 光赤天連での提案合意を目指す
- u コアメンバーでたたき台を執筆
 - u 協力者を募る
- u タイムライン？
 - u 他に議論する各論のWS・WGを開く？
 - u 具体的なアーカイブに根差して改善案を議論
 - u よりデータの作り方, and/or 品質・QAに特化したようなWS
 - u よりUIに特化したようなWS
 - u 9月のGopiraシンポまでにある程度のドラフトに
 - u 6月くらいには第1稿？
 - u 予算取り
 - u データアーカイブ用科研費？：規模は大きくない。データ整理など集約的なプロジェクトはあり得る