Cloud-Based Visual Discovery in Astronomy: Novel EOSC Services for Exploring Large Scale Catalogues

Modern radioastronomy is currently undergoing a revolution e.g. with a new generation of interferometers such as the game changing Square-Kilometre Array (SKA)[1], which is an international collaboration involving over 100 scientific and technological partners from 20 countries; upon its completion this will be the largest radio interferometer on earth. Such instrumentation infrastructures spearhead the next generation all-sky surveys towards a brand-new paradigm shift as multiwavelength all-sky surveys are expected to replace traditionally executed single-object observations, enabling to perform novel statistical studies and thus revealing previously unknown patterns and trends[2]. The scientific community is facing an emerging flurry of enormously large, incredibly rich and highly complex data volumes, e.g. when fully operational SKA is expected to generate data rates largely exceeding the global Internet traffic. The associated science-ready data products will pose extremely challenging technological demands on traditional approaches, e.g. analysis and imaging software tools will have to be adapted at various levels, if not completely re-designed from scratch, so as to run efficiently at scale with satisfactory performances. An earlier blog article[3] discussed recent technological advances in the game engines and virtual reality fields relating to NEANIAS[4] which is exploiting new tools and approaches to prototype services for agile, fit-for-purpose and sustainable offerings accessible through the European Open Science Cloud (EOSC[5]) to handle emerging, next generation radioastronomy data. The NEANIAS S1 services[6] are delivering optimised cloud-based visual discovery for multiwavelength astrophysics datasets, ensuring scalability and exploiting innovative interfaces through virtual reality. We are building upon the Astra Data Navigator (ADN[7]), which is a 3D visualisation environment supporting interactive data exploration and navigation for large stellar catalogues developed around Unity3D[8]. This article presents a brief survey of technological solutions for accessing and exploring large scale astronomical catalogues. We then focus on ADN describing its core functionalities, current release and roadmap of future developments. Enabling ADN features for suitable VR interfaces is also discussed prior to summarising our work results and pointing towards potential future developments.

There are several visualisation solutions for real-time space simulation and catalogue exploration, e.g. Celestia[9] which is an open source, multi-platform environment

---

[1] https://www.skatelescope.org/
[2] The Challenges of Radioastronomy in the Era of SKA, C. Bordiu, F.Bufano and E.Sciacca, 2020.
[3] Cloud-Based Visual Discovery in Astronomy: Big Data Exploration Using Game Engines and VR on EOSC P. Gallo, S. Terzuolo, E. Topa, B. Amuneni, L. Marlow and M. Krokos, 2020.
[4] https://www.neanias.eu/
[5] https://eosc-portal.eu/
[6] https://docs.neanias.eu/projects/s1-service/en/latest/services/adn.html
[7] Development of an Immersive Virtual Reality Application for the Exploration of the Universe, PhD Thesis, N. Lanza, 2020.
[8] https://unity.com/
[9] https://celestia.space/

supporting scripting and numerous add-ons[10]. Other sky visualisers offer extragalactic views of our universe based on astronomical imagery, e.g. the WordWide Telescope[11] and Google Sky[12] among others. The Universe Sandbox[13] is a physics based space simulator, providing an interactive educational application simulating gravitational effects in various scales, including also interactions among pairs of galaxies. SpaceEngine[14] is an open source 3D space simulation environment and game engine that can be exploited for building VR enabled 3D planetariums based on a combination of real astronomical datasets and procedural generation algorithms. Their latest version supports various VR headsets such as HTC Vive[15], Oculus Rift[16] and mixed realities via SteamVR[17]. The CosmoScoutVR[18] is a virtual universe which allows exploration, analysis and visualisation of large planetary / simulation datasets in real-time developed at the German Aerospace Centre and making use of data from Hipparcos[19] and Tycho2[20] catalogues. Their core VR software layer allows scaling experiences from single laptop applications to large CAVE immersive multi-projection systems, supporting VRPN[21], a comprehensive set of classes within a library and a set of servers designed to implement a network-transparent interface between application programs and physical devices in a VR system. Open source visualisation suites originating from the creative industries such as Blender[22] have also been deployed for scientific visualisation in astronomy, e.g. through the open-source AstroBlend[23] which is built upon the analysis tools suite provided by yt[24], for visual discovery in computational and observational astrophysics. Some recent astrophysical visualisation developments include iDaVIE-v which is a new VR software environment[25] to be released imminently (at time of writing) and is based on the Unity3D and SteamVR focusing on data cubes exploration. Similarly, Astera[26] which is a cosmological visualiser for real-time rendering of large cosmological simulations using Unreal[27] and is focused on customised renderings of large-scale universe structures.

The Astra Data Navigator, unlike its competitors described in the previous section, is being developed in very close integration with Unity3D, yet decoupled sufficiently thus allowing easily incorporating long-term updates of all its core functions relating to the underlying graphical and physics engines. ADN does not require complex build or installation procedures to be executed; a simple configuration mechanism is provided

[10] http://celestiamotherlode.net/
[11] https://worldwidetelescope.org/
[12] https://www.google.com/sky/
[13] http://universesandbox.com/
[14] http://spaceengine.org/
[15] https://www.vive.com/
[16] https://www.oculus.com/
[17] https://www.steamvr.com/
[18] https://github.com/cosmoscout/
[19] https://www.cosmos.esa.int/web/hipparcos/catalogues/
[20] https://www.cosmos.esa.int/web/hipparcos/tycho-2/
[21] https://github.com/vrpn/vrpn/
[22] https://www.blender.org/
[23] https://arxiv.org/abs/1602.03178
[24] https://yt-project.org/
[25] https://arxiv.org/abs/2012.11553
[26] https://astera.soton.ac.uk/
[27] https://www.unrealengine.com/

via XML, which can be fully customised prior to execution, e.g. to define parameter presets concerning some aspects of data loading or various simulation characteristics. A group of parameters supports settings for defining the system set up from which to retrieve motions of revolution information. For this there are the following core possibilities:

- **Position Management System** (PMS[28]) which is a web service developed within the core services offerings of NEANIAS, and requires pointing to a url for execution;
- **SpiceKernels[29]** which allows ADN to read the information requested by the kernels locally, and requires setting a path for the desired metakernel required;
- **Kepler** which calculates the positions for the revolution with mathematical formulas internal to AND and does not require any further additional information.

ADN supports both dynamic and static loading. The former allows partial data loading, and recovers the necessary information from the catalogues during the simulation as necessary, thus creating minimum latencies for loading with a lower waste of resources. The latter provides a holistic loading mechanism for all available data at the beginning, and although it may result in slowing down when starting the simulation it poses no further data loading requests. The possibility of loading personal datasets from different sources (e.g. databases or individual files) is also offered for both loading mechanisms. At a minimum ADN currently offers the Tycho2 star catalogue containing 2.5M of objects, while it is currently being updated through cross-matching with the GAIA EDR3 catalog[30]. The overall configuration system in ADN provides the possibility to specify the date from which to start a simulation. The system is initiated with a loading system displaying to the prospective user the step of simulation initialisation (Figure 1).

Fig. 1: Loading system

An ADN simulation always starts with the observation point on the Sun, which represents the origin of the system. The main functionalities are offered in groups

[28] https://docs.neanias.eu/projects/s1-service/en/latest/services/adn.html
[29] https://naif.jpl.nasa.gov/naif/data.html
[30] https://www.cosmos.esa.int/web/gaia/earlydr3

located at the bottom left/right of the window for temporal aspects / graphics and on the right for other information. At a temporal level, ADN allows changing the simulation date as specified by the user (including setting this to current dates), repositioning all objects in the correct position, speeding up/down, or even stopping the overall time evolution. The functions that concern multiple graphic features allow the display of the orbits of celestial bodies, where it is possible to calculate them, and the possibility of enabling selectors on the closest objects - these selectors indicate the objects that can be selected from a certain point by clicking the mouse. Both the orbits and the selectors have different colours based on the type of celestial body - green for planets, blue for satellites, while the white selectors indicate the stars and the red one indicates the Sun.

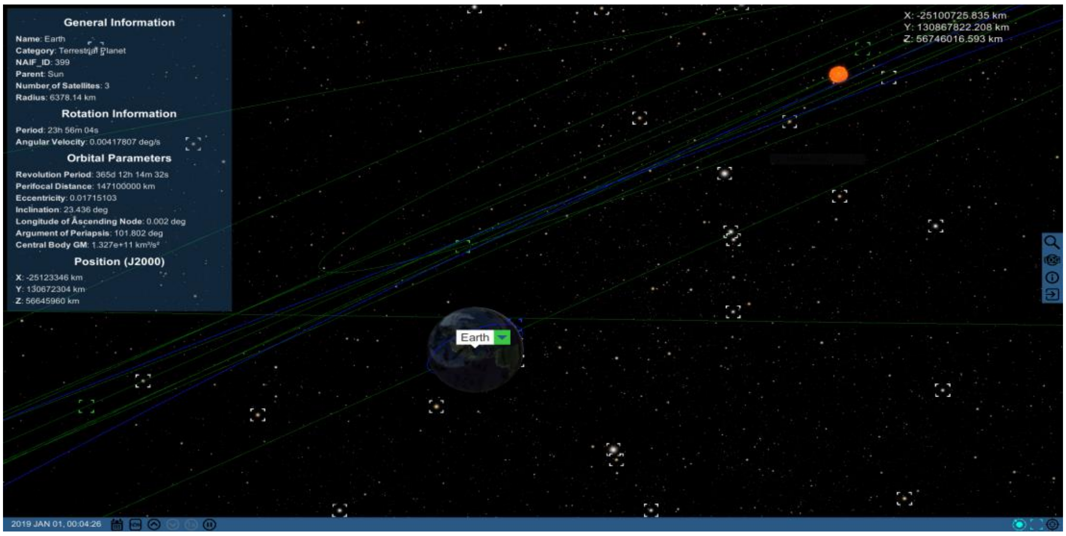

Fig. 2: Selectors and Orbits

From the menu on the right it is possible to open different panels or close the application. The first panel is the Info Panel, which is shown at the top left in Figure 2. This panel can contain various information based on the type of object selected. The other two panels available are the search panel and the log panel, as shown in Figure 3. The former allows searching for celestial objects on users can move, while the latter shows if there are various warnings or error logs. ADN also contains a series of basic features, such as free movement via WASD or increased movement speed with mouse scrolls. The modular structure of its software architecture allows an easy extension of its functionalities, making it a tool capable of evolving with a minimum effort.

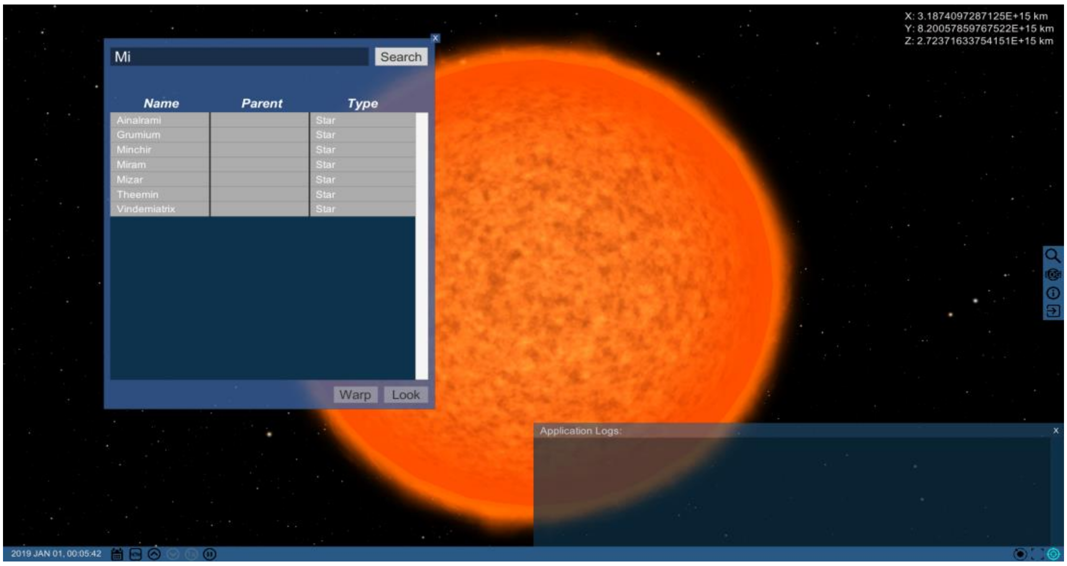

Fig. 3: Search and Logs panel

ADN exploits the advantages offered by SteamVR to provide an interface for OpenVR[31] to support ways to interact with VR displays without relying on specific hardware vendor's SDKs. SteamVR can be updated independently of the underlying gaming environment to add support for new hardware and/or software updates as required. The abstraction and modularity enabled through SteamVR means that ADN can be experienced through nearly any VR enabled Head Mounted Display (HMD) and controllers and SteamVRs interface will allow remapping of controls as necessary. SteamVR has long term support for legacy and state of the art HMDs, thus making ADN functionalities as inclusive as possible, requiring minimal maintenance (figure 4). To utilise SteamVRs modularity, the VR user interface needs to be easily interpreted and remaped. This is achieved by projecting ADN’s desktop GUI in 3D and using a laser pointer so that the resulting GUI is as simple to experience as using an ordinary desktop PC approach but embedded in a VR environment. This allows user familiarity between the desktop PC and VR versions, but also implies that extensions on the desktop PC version can be easily replicated in VR. To our experiences so far the benefit of VR enabling for ADN is that users are provided with the possibility to interpret and appreciate in detail real scale celestial objects with truly immersive high sensory impacts.

The main activities of current ADN developments are focused on improving functionalities across the board. Firstly, from the data point of view, users will be provided with wider astrophysical catalogues, our final objective is to be able to offer the GAIA EDR3 catalogue consisting of 1.8 billion objects. Secondly, several improvements of data management mechanisms are planned, together with display of more information on the objects themselves, with the possibility to export such information externally. Thirdly, for data analysis, various functionalities are being introduced to realise advanced analysis workflows ranging from sophisticated data filtering to execution of meaningful visual comparisons. Finally, for the graphic aspects, new models are planned to be inserted while continuing improving further existing

[31] https://github.com/ValveSoftware/openvr

ones by introducing features allowing meaningful visualisations of additional celestial object information, e.g. relating to the galactic plane or displacement vectors of the stars.

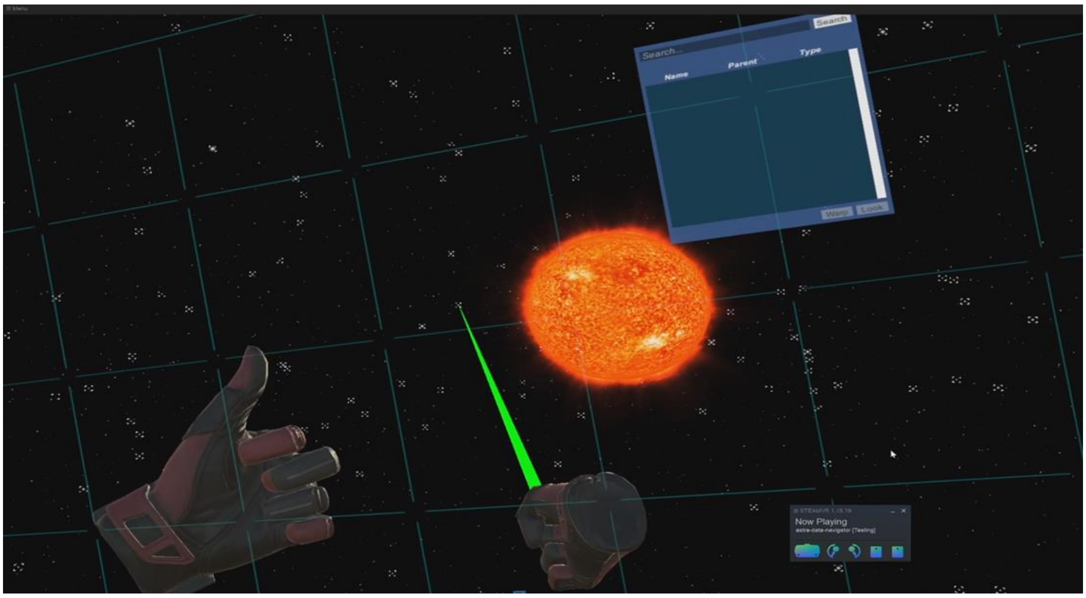

Fig. 4: SteamVR view of ADN

The present version of ADN is 1.6 and following a preliminary testing it is currently undergoing validation by external communities of users comprising students, scientific experts, VR experts and non-specialists, to drive the course of our developments for future enhancements. One direction for future work is to realise a mechanism to comprehensively deal with the vast range of scales inherent within a virtual universe representation. To maintain realism the deployed units of measurement need to be able to represent objects with dimensions far below single units while numerically be able to also contain the entire universe dimensions. The present ADN version employs fixed 1Km units which may restrict visualisation of relatively close objects from a single point of view. The idea is to implement a runtime re-scaling system to also allow maintenance of very small objects, through real-time modifications of the viewing camera representation matrices. A further direction for future developments is to allow the possibility of ADN operating collaboratively with groups of multiple users simultaneously for data analysis ideally by accessing different ADN instances from different locations. This functionality will require the implementation of an add on system for streaming and visualising large data amounts through a centralized server mechanism.

Paolo Gallo, E. Topa (ALTEC Team)
Basiru Amuneni, Lewis Marlow and M. Krokos (University of Portsmouth Team)