Astro2020 Science White Paper

Increasing the Discovery Space in Astrophysics The Exploration Question for Stars and Stellar Evolution

Thematic Area: Stars and Stellar Evolution

Principal Author:
Name: G. Fabbiano
Institution: Center for Astrophysics | Harvard & Smithsonian (CfA)
Email: gfabbiano@cfa.harvard.edu
Phone: 617 495 7203 (but email is much better)

Co-authors:
M. Elvis (CfA), A. Accomazzi (CfA), G. B. Berriman (Caltech/IPAC-NExScI), N. Brickhouse (CfA), S. Bose (CfA), D. Carrera (Penn State Univ.), I. Chilingarian (CfA), F. Civano (CfA), B. Czerny (Center for Theoretical Physics, Warsaw; CfA collaborator), R. D'Abrusco (CfA), J. Drake (CfA), R. Emami-Meibody (CfA), J. R. Farah (CfA), G. G. Fazio (CfA), E. Feigelson (Penn State Univ.), F. Fornasini (CfA), Jay Gallagher (Univ. Wisconsin), J. Grindlay (CfA), L. Hernquist (CfA), D. J. James (CfA), M. Karovska (CfA), V. Kashyap (CfA), D.-W. Kim (CfA), G. M. Lacy (NRAO), J. Lazio (JPL, Caltech), W. P. Maksym (CfA), R. Martinez Galarza (CfA), J. Mazzarella (Caltech/IPAC), D. Sanders (Institute for Astronomy, Hawaii), N. Scoville (Caltech), I. Shapiro (CfA), A. Siemiginowska (CfA), A. Smith (STScI), H. Smith (CfA), A. Szentgyorgyi (CfA), S. Tacchella (CfA), A. Thakar (Johns Hopkins Univ.), V. Tolls (CfA), B. Wilkes (CfA), D. Wilner (CfA), S. P. Willner (CfA), S. J. Wolk (CfA)

Abstract:

We write in response to the call from the 2020 Decadal Survey to submit white papers illustrating the most pressing scientific questions in astrophysics for the coming decade. We propose exploration as the central question for the Decadal Committee's discussions. The history of astronomy shows that paradigm-changing discoveries were not driven by well-formulated scientific questions, based on the knowledge of the time. They were instead the result of the increase in discovery space fostered by new telescopes and instruments. An additional tool for increasing the discovery space is provided by the analysis and mining of the increasingly larger amount of archival data available to astronomers. Revolutionary observing facilities, and the state-of-the-art astronomy archives needed to support these facilities, will open up the universe to new discovery. Here we focus on exploration for stars and stellar evolution.

1. The exploration question

There has been a long-standing tension in our discipline between the 'exploration' approach and the more physics-based 'question-driven' approach.

1.1 - *The **question-driven** approach*

This approach seeks to formulate the most important open questions in our discipline. It is based on our *present knowledge* of the field (both theoretical and observational) and is formulated usually as a way to constrain and/or advance a currently proposed cosmological or astrophysical scenario. 'Question/hypothesis-driven' has been the preferred approach in the last few decades and is used to justify both observing proposals and proposals for new instruments and telescopes. Most talks at conferences and papers are framed based on this question and answer approach. This is how we teach our students to approach research. This is the approach formulated in the Decadal Survey call for white papers. This approach addresses the '*known unknowns'*: for example, the way to best constrain the cosmological parameters of our universe, and lately the search for dark matter and dark energy, and the definitive discovery of gravitational waves.

The question-driven approach continues to be fruitful, and it gives us a certain sense of control in our progress, but -by its own nature- is also a limited and limiting epistemology. For example, it can bias our knowledge. As expounded in a recent article with reference to extra solar planets "the key is to make sure that science policy permits discovery for the sake of discovery and not for finding Earth-like planets, which we have prejudiced to be of greatest interest (D. J. Stevenson, CalTech, Physics Today, Nov. 2018)". The same opinion can be easily shaped to apply to other fields of astrophysics.

The question-driven approach does not address the *'unknown unknowns'* that by their nature cannot be addressed as well-defined 'important questions'.

1.2 - *The **exploration** approach*

This approach, i.e. gaining the capability to find new questions, rather than solving known ones, is the only way we can address the *unknown unknowns*. Harwit (1984) calls this '*discovery space*'. The notion that most of science is undiscovered and that 'out of the book' thinking may be needed for real progress is making fast inroads (e.g., see the book 'Ignorance: How it Drives Science' by S. Firestein, 2012). How to best foster the discovery of *unknown unknowns* is particularly poignant for astronomy, which throughout its history has been first and foremost exploratory.

The real big paradigm shifts in astronomy and astrophysics have occurred when new approaches have significantly opened up the discovery space, revealing unforeseen views of the universe. These approaches may have been framed as a way to address important questions of the time, but the real advances were from serendipitous discoveries. The discovery space may have been increased by means of new telescopes and instruments (both hardware and software), and also by *unanticipated data repurposing*.

Famous examples of discoveries stemming from exploration include:

- The Galilean Moons of Jupiter, the metal composition of the Sun and stars, the HR diagram, the expansion of the Universe, large scale structure, hot Jupiters (driven by improvements in optical telescopes and spectrographs);

- Quasars, radio galaxies, the microwave background, pulsars, superluminal motion, fast radio bursts (following the invention of radio telescopes, VLBI, and search in the archives in the case of bursts);
- Black holes and their mass range, dark matter, dark energy, super-starburst galaxies (from the availability of new space-based observing windows, X-ray, IR, and high resolution optical imaging with HST, and availability of multi-wavelength archives).

These foundational discoveries for the present understanding of the Universe and its evolution were not in any way anticipated. Most of them were fostered by the use of increasingly larger telescopes and more sensitive instruments, able to explore different parts of the electromagnetic spectrum. Others were surprising results of the data analysis.

Given the increasing availability of large and survey data sets in our open archives, a new hybrid approach, **question-driven exploration**, has emerged, where astronomers have mined these data and researched the literature guided by relatively vague questions, finding answers, new questions, and surprises (a similar approach is making inroads in biology; Elliott et al 2016).

In this white paper we discuss the '*exploration question'*, providing examples relevant for the field of resolved stellar populations. We include both serendipitous discoveries and question-driven explorations, resulting from unanticipated analyses of multi-wavelength survey data (Section 2). In Section 3, we address our recommendations for increasing the discovery space.

2. Exploration in Stars and Stellar Evolution

Given the nature of exploration, it is not possible to give definite questions that need to be addressed in the near future. Rather, we provide a few examples of (1) serendipitous unexpected discoveries (*unknown unknowns*) and their potential for changing established paradigms; and (2) new research avenues posed by asking very general questions (*known unknowns*). We do not mean to provide an exhaustive survey of such discoveries, but only to illustrate our case with a few representative studies.

2.1 *Unanticipated discoveries from Hubble and other space observatories*

Past the first discovery period that defined stellar physical and chemical properties, and spurred the first theoretical work on the stellar engine and stellar evolution, stellar astronomy / astrophysics is an area where discovery guided by theoretical models has been predominant and works well. Despite this, the advent of observations with *Hubble*, with its high angular resolution, has provided unexpected discoveries, for example (1) the complex stellar populations of globular clusters (Gratton et al. 2012; Piotto et al. 2015), and (2) the major role of binary star evolution in mass ejection, e.g., in stellar systems ranging from the production of planetary nebulae to binary star mergers (Jones & Boffin 2017; Smith et al. 2016, Tylenda & Soaker 2006).

New generations of *multi-wavelength data* have led to new insights into stellar physics. The enhanced capabilities provided by space observatories fostered multi-wavelength stellar research leading to unanticipated results, such as:

- The detection of black hole transients, and routinely finding dusty disks accreting matter onto protostars (with an accretion shock detected with *Chandra* in TW Hya), and producing jets (Burrows et al. 1996, McCaughrean & O'Dell 1996, Kastner et al 2002; Prisinzano et al. 2008).
- The detection of stellar oscillations, especially in evolved red giant stars by *CoRoT* and *Kepler*, which opened a powerful avenue to test stellar models (Chaplin & Miglio 2013),
- The detection of γ-rays from novae in outburst, which revealed their ability to act as shock-powered high-energy particle accelerators (Ackermann et al. 2014).

- The discovery of coronal activity in a nearby brown dwarf with *Chandra*, showing that these ultra-cool objects undergo energetic reconnection flares similar to fusion-powered stars like the Sun (Rutledge et al. 2000).

The ability to connect a variety of multi-wavelength observations to sophisticated models is a key factor behind these advances. Progress thus depends on access to data by researchers with a wide variety of backgrounds.

2.2 *New understanding from archival studies*

New research avenues are opened by the presence of the growing amount of stellar data available in astronomy archives, and by the availability of powerful software tools to exploit these data. In these types of projects exploration comes from the capture of high quality data for large samples of objects. Hundreds of thousands of high quality stellar spectra have been collected by different observatories over the last three decades and put into data archives. Tools to determine stellar atmospheric parameters (effective temperature, surface gravity, elemental abundances, rotation) evolved over time and became mature only a few years ago thanks in particular to the development in the field of studies of exoplanets. New methods of data exploration are also coming to the fore, especially based on classification and clustering based on machine learning. Examples of this approach include:

- AMBRE, a project that re-processed some 52,000 high-resolution stellar spectra collected with 4 different instruments at the European Southern Observatory (Worley et al. 2012; de Laverny et al. 2013). By combining precise measurements of stellar parameters and radial velocities obtained from spectra with *Gaia* information on distances and proper motions, archival studies in the area of Galactic astronomy have an enormous potential in the next decade. Analyzing the archival stellar spectra in a uniform way and determining stellar properties will help to pin down some of the unsolved questions of stellar evolution and synthesis of heavy elements in stellar interiors.
- The discovery of a diagonal rotation period gradient across the main sequence in stars from *Gaia* Data Release 2, matched with rotation periods from *Kepler*, which may be due to metallicity effects (Davenport & Covey 2018; Davenport 2019).

A variety of stellar spectroscopic surveys are in progress or planned for the next decade, with a variety of goals. These surveys will provide huge accessible archival databases. They include:

- APOGEE, a program in the Sloan Digital Survey IV designed to study stellar abundances patterns in the Milky Way.
- ULLYSES, the recently announced project from STScI that seeks to produce a full library of ultraviolet spectra of massive stars in the SMC, to provide a knowledge base for modeling stars at low metallicity. This project will gain support from a number of PI-programs of optical spectroscopy of massive stars in the SMC, such as those that were enabled by the capabilities at ESO for moderate-to-high spectral resolution multi-object observations over substantial fields of view.

These studies will further benefit from connections to future multi-messenger surveys such as the photometric studies by WFIRST and SPHEREx, as well as new generations of gravity wave and neutrino observatories. The results will impact extragalactic astronomy because new generations of stellar population models can be created utilizing improved stellar astrophysics and used to quantitatively explore the evolution of galaxies.

The time domain also is an area where enhanced capabilities led to new discoveries and where more progress will come, e.g., with the advent of LSST. Examples include the OGLE gravitational lensing surveys that led to new understanding of the structure of the Magellanic Clouds. Many optical transient surveys have enriched greatly our understanding of supernovae and stellar variability. Key features of these programs include systematic approaches to the capture, analysis, and delivery of results to the community.

3. Increasing the Discovery Space

3.1 *Observing facilities that expand boundaries*

Any new observing facilities/missions for the next decade should significantly improve performance in some key metric (e.g., energy range, sensitivity, exposure time, angular resolution, higher dimensional data, rapid response), and be well characterized and calibrated, so to provide flexibility for new observing avenues. *Hubble*, *Spitzer* and *Chandra* provide examples in the discovery of Dark Energy, the detection of z=11 galaxies, and the nature of Dark Matter (Bullet Cluster), respectively. Beyond hardware capabilities these discoveries require: mission longevity, community driven science, high-quality data products in readily accessible, interoperable, archives and a well-supported user/observer community.

3.2 *Multi-wavelength and multi-messenger capabilities*

Many historical examples also demonstrate a strong synergy between different wavebands and messengers. Having contemporaneous access to the entire electromagnetic spectrum was vital to finding the first counterpart to a gravitational wave source, for example. This multi-wavelength coverage of the sky that we are currently enjoying needs to be preserved.

3.3 *Curated Data Archives and Powerful Data Analysis tools*

These new facilities will generate increasingly larger and complex multi-wavelength and multi-messenger data sets and catalogs. These data will need to be properly reduced and curated to fully enable their discovery potential. *Archives must provide both easy access to these data and (with the community) the means to exploit them*. These goals translate into:

(1) Ensure that any operational (old and new) facility/mission explicitly include in their scope the proper processing of software so to produce well documented and calibrated data products, as well as the capability for data recalibration and reprocessing.

(2) Organize these data products in well-maintained archives, following the International Virtual Observatory Alliance (IVOA)[1] standards, so to allow a basic level of access and *interoperability*, as well as *repurposing.* Much of this is already in place in the NASA archives, and they are collaborating in extending and evolving these capabilities to meet the demands of new data types and research methods through the 2020s. Data products should be replicable and reproducible, ranging from basic observation data to high-level aggregated data and catalogs.

(3) Ensure that data centers engage in the development and refinement of interoperability standards, via the well-established processes of the IVOA, and work with groups such as *Astropy*[2] to ensure support for these standards in present in community developed, open source software.

(4) *New facilities* (Sections 3.1, 3.2) *will demand a transformation in the way data are analyzed.* The early phases of this transformation are already underway (e.g., the use of *Python* as an environment, cloud computing). But, resources must be made available for full development, which will demand remote Science Platforms[3] and Server-side analytics[4], implementation of

---

[1] The forum for the development of the interoperability standards used by major astronomy datacenters (http://www.ivoa.net)

[2] http://www.astropy.org/acknowledging.html

[3] See LSST Science Platform Design document https://ldm-542.lsst.io

complex fault-tolerant workflows, data mining and machine learning, and advanced visualization.

(5) Foster the development of *next generation* interoperable, user-friendly visual interfaces, data mining tools, the ability to construct and implement analysis workflows easily, both via visualization and scripting, and the ability to work with data both locally and remotely (current-generation well-know examples include TOPCAT, DS9 and CSCView).

(6) Support interdisciplinary research in astrostatistics and astroinformatics and the transfer of methods from the statistics, computer science, and machine learning communities, for development and application of innovative data analysis methods and algorithms.

(7) Ensure that facilities and archives participate in curation efforts and initiatives to link together datasets, related ancillary data (e.g., atomic and molecular databases), objects, and the literature.

Data are an important legacy of major astronomical facilities, and proper data maintenance will insure that new science will be produced for the future, even after the first crop of scientific papers and discoveries have been published. Statistics of data usage from the NASA archives demonstrate that archival data is used for new published scientific work several times (Fig. 1).

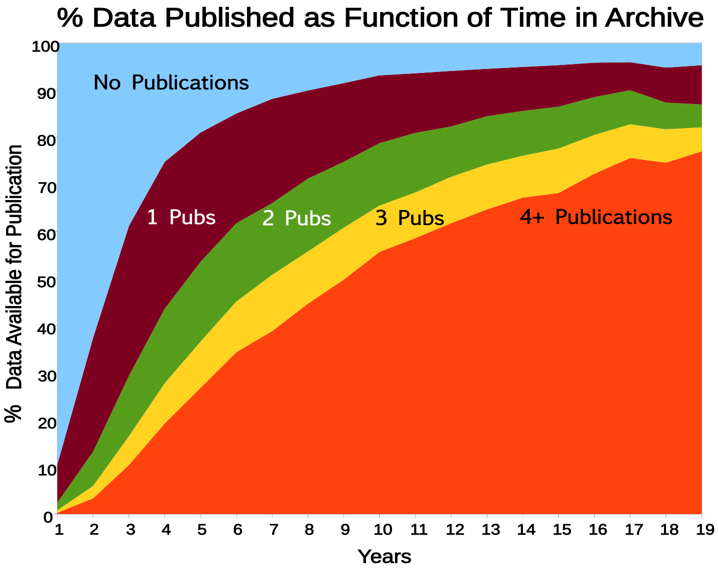

Figure 1. – Percentage of *Chandra* exposure time published versus years in the archive. The scientific use of archival *Chandra* data is increasing with time in the public archive. For example, 19 years from launch, ~75% of the observation have been published in more than 4 papers. A similar trend is observed for the HST data.

4. Conclusions

We propose *exploration* as the central question for the Decadal Committee's discussions. The history of astronomy shows that paradigm-changing discoveries were not driven by well-formulated scientific questions, based on the knowledge of the time. They were instead the result of the increase in discovery space fostered by new telescopes and instruments. An additional tool for increasing the discovery space is provided by the analysis and mining of the increasingly larger amount of archival data available to astronomers. We urge the Decadal Committee to **(1)** *keep multi-wavelength and multi-messenger exploration center stage* in their deliberations of new facilities, including consideration for flexible and well-calibrated modes of operation that could foster adaptation for use with new discovery space; and **(2)** *recognize the importance of data and*

---

[4] NASA Big Data Task Force (https://science.nasa.gov/science-committee/subcommittees/big-data-task-force)

their stewardship, and computational services, as major elements of any new scientific development for the next decade. ***Revolutionary observing facilities, and the state-of-the-art astronomy archives needed to support these facilities, will open up the universe to new discovery.***

References

Ackemann, M. et al. the Fermi-LAT collaboration 2014, Science 345, 554.

Burrows, C. J. et al. 1996, ApJ, 473, 437.

Chaplin, W. J. & Miglio, A. 2013, ARA&a, 51, 353.

de Laverny, P., Recio-Blanco, A., Worley, C. C., et al. 2013, The Messenger, 153, 18.

Davenport, J. R. A. & Covey, K. R. 2018, ApJ 868, 151.

Davenport, J. R. A. 2019, AAS 126.2.

Elliot , K. C. et al. 2016, Bioscience, 66(10), 880. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5862324/

Firestein, S. 'Ignorance: How it Drives Science' 2012, Oxford Univ. Press .

Gratton, R. G.; Carretta, E.; Bragaglia, A. 2012, ARA&A, 20, 50.

Harwit, M. 1984, 'Cosmic Discovery: The Search, Scope and Heritage of Astronomy', The MIT Press

Jones, D., Boffin, H. M. J.,2017, Nature Astronomy, 1, 117.

Kastner, J. H. et al. 2002, ApJ, 567, 434.

McCaughrean, M. J., O'Dell, C. R. 1996, AJ, 111, 1977.

Piotto, G. et al. 2015, AJ, 149, 91.

Prisinzano, L. et al. 2008, ApJ, 677, 401.

Rutledge, R. E. et al. 2000, ApJ, 538L, 141.

Tylenda, R., Soaker, N. 2006, A&A, 451, 223.

Remillard, R. A., McClintock, J. E. 2006, ARA&A, 44, 49.

Smith, N. et al. 2016, MNRAS, 458, 950.

Worley, C. C., de Laverny, P., Recio-Blanco, A., et al. 2012, A&A, 542, A48

.