Lecture 1: Exploring Space in Cyberspace: Astronomy and Data Science

- Setting the stage: an ongoing transformation of science
- Astronomy in the era of an exponential data growth: from Virtual Observatory to Astroinformatics
- Exploration of parameter spaces and other outstanding challenges
- Science on the carbon-silicon interface: the rise of the machines
- Methodology transfer in action
- Concluding comments

Some readings:

- *The Fourth Paradigm: Data-Intensive Scientific Discovery*, eds. T. Hey, S. Tansley, & K. Tolle (2009), Published by Microsoft Research, free download from https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery/
- *Frontiers in Massive Data Analysis*, National Academy Press (2013), free download from https://www.nap.edu/catalog/18374/frontiers-in-massive-data-analysis
- *Astroinformatics 2016*, Proc. IAU Symposium 325, eds. M. Brescia, S. Cavuoti, G. Longo, S. G. Djorgovski, & E. Feigelson, Cambridge, UK: Cambridge Univ. Press.
- *In general, talks from all annual Astroinformatics conferences are available on their websites*, astroinformatics201X.org, where X = 0, 1, 2, ... 8 (so far).
- *Sky Surveys*, Djorgovski, S.G., Mahabal, A., Drake, A., Graham, M., & Donalek, C. 2012, in Astronomical Techniques, Software, and Data (ed. H. Bond), Vol.2 of Planets, Stars, and Stellar Systems (ser. ed. T. Oswalt), p. 223. Berlin: Springer Verlag; arXiv/1203.5111
- *Challenges for Cluster Analysis in a Virtual Observatory*, Djorgovski, S.G., Brunner, R., Mahabal, A., Williams, R., Granat, R., & Stolorz, P. 2002, in: Statistical Challenges in Astronomy III, eds. E. Feigelson & J. Babu,p. 125, New York: Springer; arXiv/0208246
- *Virtual Observatories of the Future*, eds. R. Brunner, S.G. Djorgovski & A. Szalay 2001, Proceedings of a conference held in Pasadena, Ca., in June 2000. A.S.P. Conf. Ser. Vol. 225; available through ADS, http://adsbit.harvard.edu/cgi-bin/nph-journal_query?volume=225&plate_select=NO&page=1&plate=&cover=&journal=ASPC.
- *Exploration of Large Digital Sky Surveys*, Djorgovski, S.G., Brunner, R., Mahabal, A., Odewahn, S., de Carvalho, R., Gal, R., Stolorz, P., Granat, R., Curkendall, D., Jacob, J., & Castro, S. 2001, in: Mining the Sky, eds. A.J. Banday etal, ESO Astrophysics Symposia, p. 305, Berlin: Springer Verlag; arXiv/0012489

Lecture 2: Data Visualization: What Can you See In Your Data?

- The essential roles of data visualization, and the challenge of multidimensionality
- Data types and their representations
- From data space to visualization space
- The uses and limitations of color
- From 1-D to multi-D, and the types of plots
- Virtual and Augmented Reality as a new visualization platform
- Where are we going next?

Some readings:

- Books by Edward Tufte
- Videos of talks and panel discussions from the Keck Institute for Space Studies, *Virtual and Augmented Reality for Space Science and Exploration* (2018) are available at http://kiss.caltech.edu/symposia/space_science/index.html

- *Immersive and Collaborative Data Visualization Using Virtual Reality Platforms*, Donalek, C., Djorgovski, S.G., Cioc, A., Wang, A., Zhang, J., Lawler, E., Yeh, S., Davidoff, S., Norris, J., & Longo, G. 2014, Proc. IEEE Big Data 2014, p. 609, IEEE press; arXiv/1410.7670
- *The MICA Experiment: Astrophysics in Virtual Worlds*, Djorgovski, S.G., Hut, P., Knop, R., Longo, G., McMillan, S., Vesperini, E., Donalek, C., Graham, M., Mahabal, A., Sauer, F., White, C., & Lopes, C. 2013, Ref. proc. of the SLACTIONS 2012 International Research Conference on Virtual Worlds, eds. L. Morgado et al., p. 49. Vila Real, PT; arXiv/1301.6808

Lecture 3: Astroinformatics in the Time Domain: Classification of Light Curves and Transients

- The rise of Time Domain Astronomy (TDA) and synoptic sky surveys
- The challenges of TDA classification: data heterogeneity and data sparsity
- A variety of classification methods
- From light curves to feature vectors
- Dimensionality reduction: Principal Component Analysis (PCA) and its uses
- Feature selection methods
- Including the contextual information and Bayesian Networks
- Symnbolic Regression: machine-assisted discovery of relationships
- Automating the optimal follow-up

Some readings:

- *Discovery, Classification, and Scientific Exploration of Transient Events From the Catalina Real-Time Transient Survey*, Mahabal, A., Djorgovski, S.G., Drake, A., Donalek, C., Graham, M., Moghaddam, B., Turmon, M., Williams, R., Beshore, E., & Larson, S. 2011, Bull. Astr. Soc. India, 39, 387; arXiv/1111.0313
- *Data Challenges of Time Domain Astronomy*, Graham, M., Djorgovski, S.G., Mahabal, A., Donalek, C., Drake, A., & Longo, G. 2012, Distributed and Parallel Databases, 30, 371; arXiv/1208.2480
- *Real-Time Data Mining of Massive Data Streams from Synoptic Sky Surveys*, Djorgovski, S.G., Graham, M., Donalek, C., Mahabal, A., Drake, A., Turmon, M., & Fuchs, T. 2016, Future Gen. Comp. Sys., 59, 95-104 (Elsevier); arXiv/1601.04385
- *Feature Selection Strategies for Classifying High Dimensional Astronomical Data Sets*, Donalek, C., Kumar, A., Djorgovski, S.G., Mahabal, A., Graham, M., Fuchs, T., Turmon, M., Sajeeth Philip, N., Yang, T.-C., Longo, G. 2013 proc. Scalable Machine Learning: Theory and Applications, IEEE BigData 2013, IEEE Press; arXiv/1310.1976
- *An Analysis of Feature Relevance in the Classification of Astronomical Transients With Machine Learning Methods*, D'Isanto, A., Cavuoti, S., Brescia, M., Donalek, C., Longo, G., Riccio, G., & Djorgovski, S.G. 2016, MNRAS, 457, 3119; arXiv/1601.03931
- *Machine-Assisted Discovery of Relationships in Astronomy*, Graham, M., Djorgovski, S.G., Mahabal, A., Donalek, C., & Drake, A. 2013, M.N.R.A.S. 431, 2371; arXiv/1302.5129

Lecture 4: Big Data Meet Big Black Holes: Quasars in the Time Domain

- Statistical descriptors of quasar variability: Damped Random Walk (DRW) and Continuous Autoregressive Process (CAR(1))
- Quasar variability feature space and quasar discovery using variability
- Characteristic time scale and deviations from the DRWOutliers in the variability feature space and their physical interpretations
- Quasar megaflares and their possible physical origins

- Measuring periods, and periodicity superposed on a red noise
- Supermassive Black Hole Binaries and how to find them
- Predictive stochastic behavior

Some readings:

- *Detection of quasars in the time domain,* M. Graham, S. G. Djorgovski, D. Stern, A. Drake, & A. Mahabal 2017, in ref. proc. IAU Symposium 325, Astroinformatics, eds. M. Brescia et al., p. 231, Cambridge, UK: Cambridge Univ. Press; arXiv/1612.07271
- *A Novel Variability-Based Method for Quasar Selection: Evidence for a Rest-Frame ~54 d Characteristic Time-Scale*, Graham, M., Djorgovski, S.G., Drake, A., Mahabal, A., Chang, M., Stern, D., Donalek, C., & Glikman, E., 2014, M.N.R.A.S., 439, 703; arXiv/1401.1785
- *A Comparison of Period Finding Algorithms*, Graham, M., Drake, A., Djorgovski, S.G., Mahabal, A., Donalek, C., Duan, V., & Maher, A. 2013, M.N.R.A.S. 434, 3423; arXiv/1307.2209
- *Using Conditional Entropy to Identify Periodicity*, Graham, M., Drake, A., Djorgovski, S.G., Mahabal, A., & Donalek, C. 2013, M.N.R.A.S. 434, 2629; arXiv/1306.6664
- *A Systematic Search for Close Supermassive Black Hole Binaries in the Catalina Real-Time Transient Survey*, Graham, M., Djorgovski, S.G., Stern, D., Drake, A., Mahabal, A., Donalek, C., Glikman, E., Larson, S., & Christensen, E. 2015, MNRAS, 453, 1562; arXiv/1507.07603
- *Understanding extreme quasar optical variability with CRTS: I.* Major AGN flares, M. Graham, S. G. Djorgovski, A. Drake, D. Stern, A. Mahabal, E. Glikman, S. Larson, & E. Christensen 2017, MNRAS, 470, 4112; arXiv/1706.0307