

Disciplina: Tópicos de e-science

**Rationale:**

O desenvolvimento de programas de computador tem ocupado cada vez mais o tempo de cientistas no exercício de sua função, seja por conta da recente avalanche de dados astronômicos ou refinamento de simulações e modelos a fim de explicar esses dados. Esse curso se propõe a fazer uma introdução ao desenvolvimento de software e processamento de dados na área de astronomia. Além de aulas expositivas do conteúdo teórico, são feitas sessões com foco em problemas do estudante e seus colegas ou de sub-áreas da astronomia (dependendo do tamanho da turma), com desenvolvimento e apresentação de trabalhos.

**Programa:**

- Engenharia de software: Requerimentos, Design, Construção, Testes e Manutenção
  - Abstração, modularização, e versionamento de códigos são alguns dos tópicos abordados.
- Linguagem de script: python
  - o python é uma linguagem em franca expansão na comunidade astronômica, com criação de vários pacotes de suporte (ex. esutil e astropython). Sua curva de aprendizado é bastante inclinada assim como é curto o tempo para prototipar e produzir nessa linguagem.
- Infra-estruturas de hardware: grids e clusters
  - Com o advento do Big Data, várias infra-estruturas estão sendo desenvolvidas para dar suporte à sua análise. Apresentamos algumas dessas infra-estruturas e que desafios são enfrentados em sua operação.
- Workflows e paralelismo
  - A maneira como os dados fluem em um pipeline, suas dependências entre módulos e diferentes tipos de paralelismo são discutidos.
- Ferramentas do Observatório Virtual
  - O Observatório Virtual desenvolve várias ferramentas que auxiliam o astrônomo em seu trabalho, combinando e filtrando amostras de diferentes tipos, através da inter-operabilidade de bases de dados.
- Uso de Databases
  - Big Data exige maneiras organizadas de manter e acessar seus dados. Vários levantamentos disponibilizam seus dados na internet, sendo um grande exemplo o Sloan Digital Sky Survey. Vamos revisar o que são databases e como acessar esses dados.

**Bibliografia:**

- The Fourth Paradigm: Data-Intensive Scientific Discovery <http://research.microsoft.com/en-us/collaboration/fourthparadigm/> Ed. Tony Hey et al. 2009
- Software Engineering Body of Knowledge (SWEBOK) <http://www.computer.org/portal/web/swebok/htmlformat>
- Python Scripting for Computational Science. Hans Petter Langtangen. 2004.
- Grid Computing: Making the Global Infrastructure a Reality. Ed. Fran Berman et al. 2003
- Astropython <http://www.astropython.org/>
- Scisoft <http://www.eso.org/sci/software/scisoft/>
- US Virtual Astronomical Observatory <http://www.usvao.org/>
- SDSS SQL Tutorial <http://cas.sdss.org/dr4/en/help/howto/search/>