Aplicações de Software Livre para pesquisa

19 Semana Universitária
Casa de Cultura da América Latina
26/09/2019

Dalton Martins

Faculdade de Ciência da Informação Universidade de Brasília

daltonmartins@unb.br

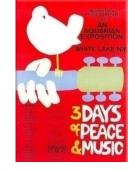








Software nos anos de 1970



Era da Liberdade e Pós Woodstock

Software não era propriedade, era compartilhado.

Grupos de pesquisa produziam software livre.

Empresas distribuíam software livre.

Programadores eram livres para cooperar.



Software nos anos de 1980

Era dos softwares proprietários e Microsoft

Software como propriedade exclusiva de empresas.

O usuário só tinha o direito de executar o programa, e em determinadas maquinas.

Conceito de licença de uso (o software continuava sendo

da empresa)



O MIT, O Guru e A Impressora

Aqui começa a história do software livre....







Renasce a **Filosofia** de Software Livre

Conceito central é aliberdade

"Free as in speech, not as in beer."

"Liberdade de expressão, não cerveja grátis"

Open Source Voices



Software Livre

Definição de Software Livre:

Qualquer programa de computador, onde exista a **liberdade** de ser **compartilhado**, **estudado** e **modificado**.

Software livre é o OPOSTO de software proprietário.

Software livre não OPÕE ao software comercial.

Proprietário x livre

Proprietário

Sem acesso ao código fonte do sistema.

Controle sobre a comercialização e utilização.

Lucro é sobre o software.

Livre

Acesso ao código fonte. Permissão de cópia e alteração.

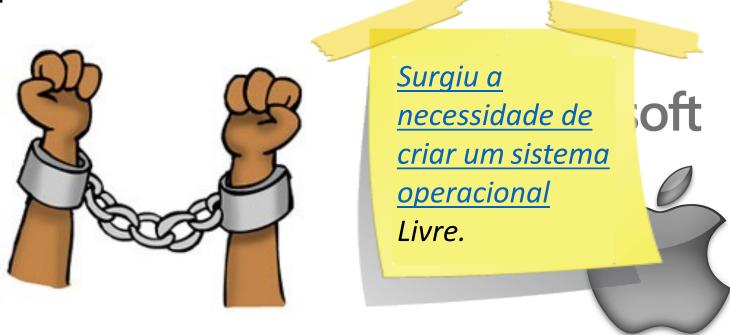
Lucro é no suporte e serviços relacionados.

Libertando-se dos software proprietários

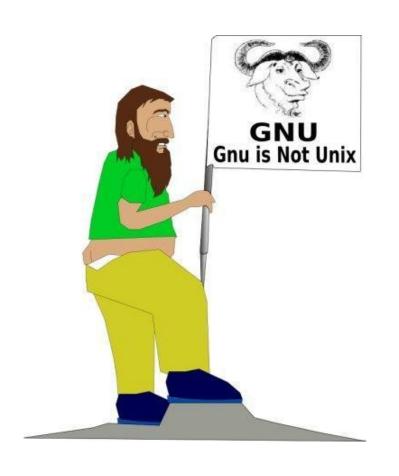
Software Livre é apenas a filosofia, mas como torná-la realidade ?

Como propagar software livre na era dos softwares

proprietários?



Nasce o **Projeto** GNU



Fundado por Richard Stallman em 1984, tem como objetivo concretizar a ideologia de Software Livre

Solidariedade/Liberdade.

Um sistema livre compatível com Unix.

Ficar totalmente livre dos software proprietários.

Ainda Faltava o Kernel

Passado alguns anos, eles já tinha desenvolvido vários softwares, mas ...

CompiladoresGCC (Gnu C Compiler)

Editores Vi, Emacs

Outros Sed, Make, ...

Interpretadores de comando
Bash, Sh

"Nos anos 90, nós tínhamos ou encontrado ou escrito todos os componentes principais, exceto um -- o kernel" (Stallman)

Um pouco longe do MIT



TreinaLinux.com.br,

Surge o Linux no momento Certo

No dia 5 de outubro de 1991, Linus publicou a seguinte mensagem no fórum de discussão, usenet, na WEB

"Você suspira pelos bons tempos do Minix-1.1, quando os homens eram homens e escreviam seus próprios "device drivers?... Você está sem um bom projeto em mãos e deseja trabalhar num S.O. que possa modificar de acordo com as suas necessidades?...".

Como eu mencionei há um mês atrás, estou trabalhando numa versão independente de um S.O. similar ao Minix para computadores AT-386... Ele está na versão 0.02... contudo eu tive sucesso ao executar bash, gcc, gnu-make, gnu-sed, compressão etc. nele.

GNU/LINUX

Linux é um clone livre do kernel do UNIX, escrito a partir do zero por Linus Torvalds, com a ajuda de um grupo de programadores espalhados pela internet.

GNU/Linux é o sistema operacional totalmente livre que usa o Linux como kernel.





O que é software livre?

- Software Livre, ou Free Software, conforme a <u>definição de software</u> <u>livre</u> criada pela <u>Free Software Foundation</u>, é o software que pode ser usado, copiado, estudado, modificado e redistribuído sem restrição. A forma usual de um software ser distribuído livremente é sendo acompanhado por uma licença de software livre (como a GPL ou a BSD), e com a disponibilização do seu código-fonte.
- Software Livre é diferente de software em domínio público. O primeiro, quando utilizado em combinação com licenças típicas (como as licenças GPL e BSD), garante os direitos autorais do programador/organização. O segundo caso acontece quando o autor do software renuncia à propriedade do programa (e todos os direitos associados) e este se torna bem comum.

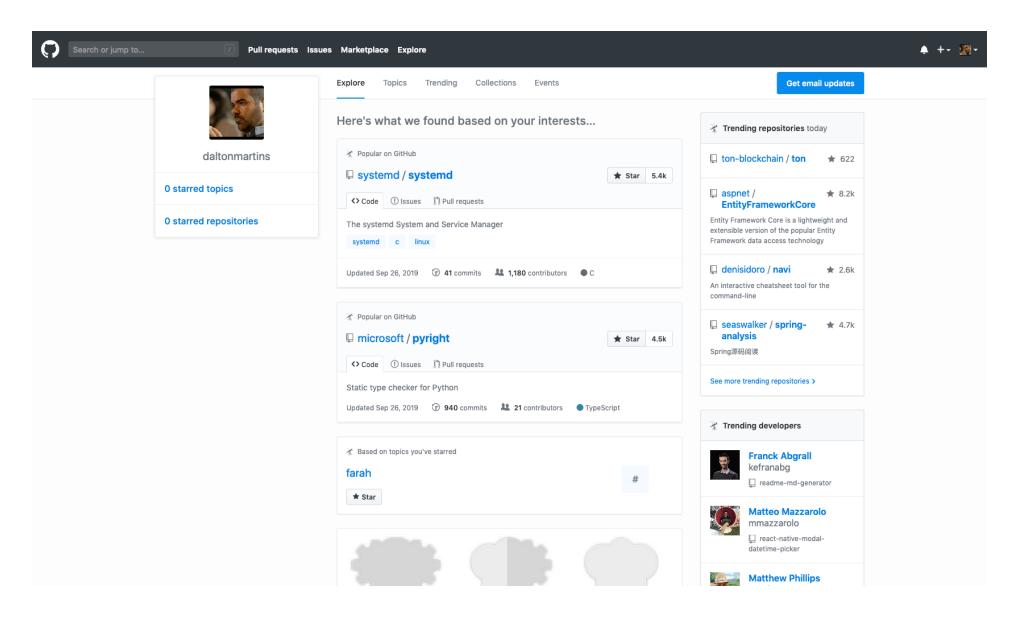
O que é software livre?

- O Software Livre como movimento organizado teve início em 1983, quando Richard Stallman (foto acima) deu início ao <u>Projeto GNU</u> e, posteriormente, à Free Software Foundation.
- Software Livre se refere à existência simultânea de quatro tipos de liberdade para os usuários do software, definidas pela Free Software Foundation. Veja abaixo uma explicação sobre as 4 liberdades, baseada no texto em português da Definição de Software Livre publicada pela FSF:
- As 4 liberdades básicas associadas ao software livre são:
 - A liberdade de executar o programa, para qualquer propósito (liberdade nº 0)
 - A liberdade de estudar como o programa funciona, e adaptá-lo para as suas necessidades (liberdade nº 1). Acesso ao código-fonte é um pré-requisito para esta liberdade.
 - A liberdade de redistribuir cópias de modo que você possa ajudar ao seu próximo (liberdade nº 2).
 - A liberdade de aperfeiçoar o programa, e liberar os seus aperfeiçoamentos, de modo que toda a comunidade se beneficie (liberdade nº 3). Acesso ao código-fonte é um pré-requisito para esta liberdade.

Algumas reflexões...

- Software é uma forma de criação intelectual;
- Uma maneira de expressar ideias, pensamentos e formas de como entendemos que as coisas deveriam funcionar;
- Dominar a forma de construir software é dominar uma forma de expressar o pensamento humano: um grau de liberdade a mais;
- É, sem dúvida, um valor que devemos defender e procurar nos capacitar como modo de expressão no universo da Ciência da Informação...
- Para as funções sociais que devemos desempenhar para a construção de uma sociedade livre, democrática e com menos desigualdade, essa é uma das ferramentas com as quais podemos atuar!

Onde encontrar softwares livres?



Uma caixa de ferramentas de softwares livres para pesquisa...

Pacote de aplicativos de escritório



+Fácil, +Rápido, e + Produtivo!

O LibreOffice é uma poderosa suíte de produtividade de escritório com planilha, editor de texto, editor de apresentação e muito mais.

Sua interface limpa e suas ferramentas avançadas o farão soltar sua criatividade e melhorar sua produtividade.

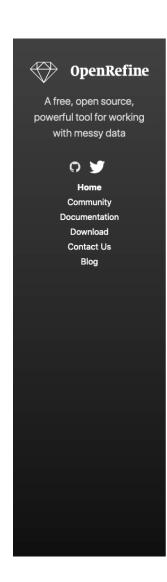
O LibreOffice incorpora várias aplicações que as torna a mais poderosa suíte de produtividade de escritório livre e aberta do mercado.







Tratamento de dados



Welcome!

OpenRefine (previously Google Refine) is a powerful tool for working with messy data: cleaning it; transforming it from one format into another; and extending it with web services and external data.

OpenRefine always keeps your data private on your own computer until YOU want to share or collaborate. Your private data never leaves your computer unless you want it to. (It works by running a small server on your computer and you use your web browser to interact with it)

OpenRefine is available in English, Chinese, Spanish, French, Russian, Portuguese (Brazil), German, Japanese, Italian, Hungarian, Hebrew, Filipino, Cebuano, Tagalog

OpenRefine was supported in 2018 by:



Introduction to OpenRefine

1. Explore Data

OpenRefine can help you explore large data sets with ease. You can find out more about this functionality by watching the video below.



Análise estatística, análise e mineração de dados



Home]

Download

CRAN

R Project

About R Logo Contributors What's New? Reporting Bugs Conferences Search

Get Involved: Mailing Lists Developer Pages

R Blog

R Foundation

Foundation Board Members Donors Donate

Help With R

Getting Help

Documentation

Manuals FAQs The R Journal Books Certification Other

Links

Bioconductor Related Projects GSoC

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To **download R**, please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.

News

- R version 3.6.1 (Action of the Toes) has been released on 2019-07-05.
- useR! 2020 will take place in St. Louis, Missouri, USA.
- R version 3.5.3 (Great Truth) has been released on 2019-03-11.
- The R Foundation Conference Committee has released a call for proposals to host useR! 2020 in North America.
- You can now support the R Foundation with a renewable subscription as a supporting member
- The R Foundation has been awarded the Personality/Organization of the year 2018 award by the professional association of German market and social researchers.

News via Twitter





Mineração de dados e aprendizagem de máquina

Project Software Book Courses Publications People Related

Machine Learning at Waikato University

An exciting and potentially far-reaching development in computer science is the invention and application of methods of machine learning (ML). These enable a computer program to automatically analyse a large body of data and decide what information is most relevant. This crystallised information can then be used to automatically make predictions or to help people make decisions faster and more accurately.

Project Objectives

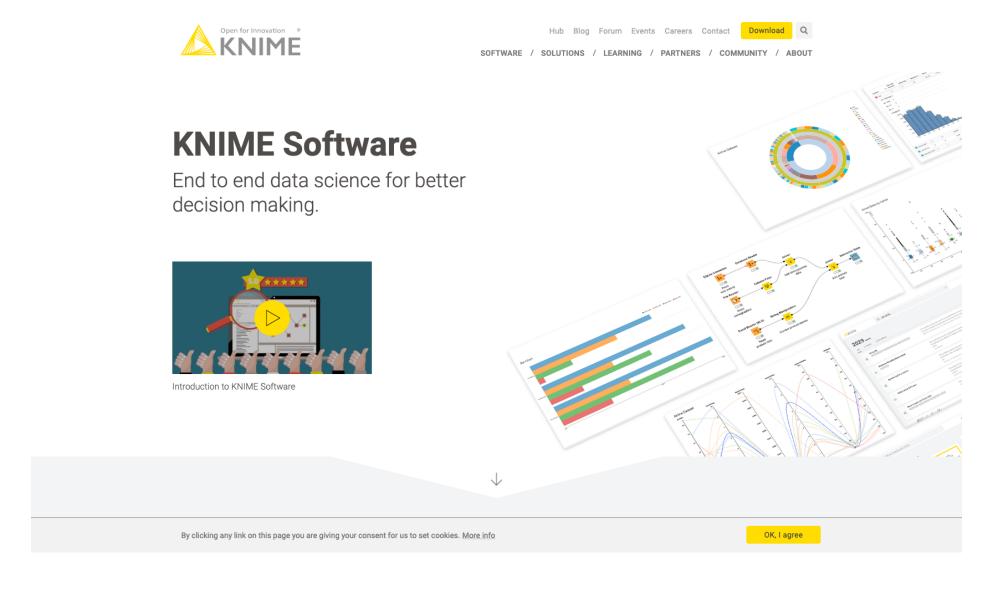
Our objectives are to

- make ML techniques generally available:
- · apply them to practical problems that matter to New Zealand industry;
- · develop new machine learning algorithms and give them to the world;
- · contribute to a theoretical framework for the field.

Software

Our team has incorporated several standard ML techniques into a software "workbench" called Weka, for Waikato Environment for Knowledge Analysis. With it, a specialist in a particular field is able to use ML to derive useful knowledge from databases that are far too large to be analysed by hand. Weka's users are ML researchers and industrial scientists, but it is also widely used for teaching. Recently, our team has also worked on MOA, an environment for mining data streams.

Ciência de dados

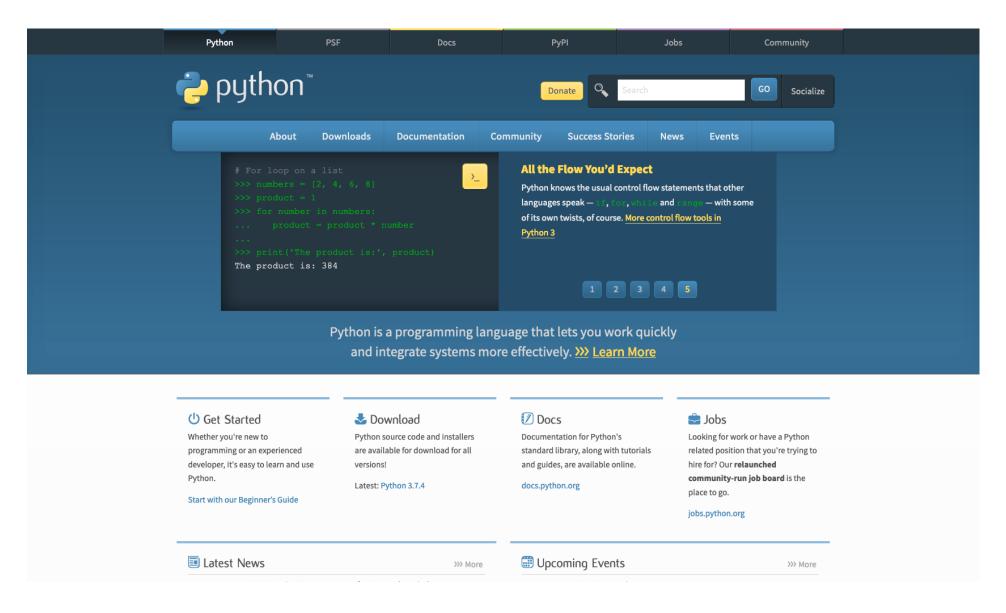


Gestão da informação e Repositórios

- Muitas possibilidades:
 - ATOM https://www.accesstomemory.org/pt-br/
 - KOHA https://koha-community.org/
 - Omeka https://omeka.org/
 - Tainacan http://tainacan.org/
 - Dspace https://duraspace.org/dspace/
 - Dataverse https://dataverse.org/
 - CKAN https://ckan.org/
 - Etc...

Ciência de Dados com Python...

Linguagem de script para ciência de dados: Python



Pandas: biblioteca Python para análise de dados











home // about // get pandas // documentation // community // talks // donate

Python Data Analysis Library

pandas is an open source, BSD-licensed library providing high-performance, easy-touse data structures and data analysis tools for the Python programming language.

pandas is a NumFOCUS sponsored project. This will help ensure the success of development of pandas as a world-class open-source project, and makes it possible to donate to the project.

A Fiscally Sponsored Project of



v0.25.1 Final (August 22, 2019)

This is a minor bug-fix release in the 0.25.x series and includes some regression fixes and bug fixes. We recommend that all users upgrade to this version.

See the full whatsnew for a list of all the changes.

The release can be installed with conda from the defaults and conda-forge channels:

conda install pandas

Or via PyPI:

python -m pip install ---upgrade pandas

VERSIONS

Release

0.25.1 - August 2019 download // docs // pdf

Development

0.26.0 - September 2019 github // docs

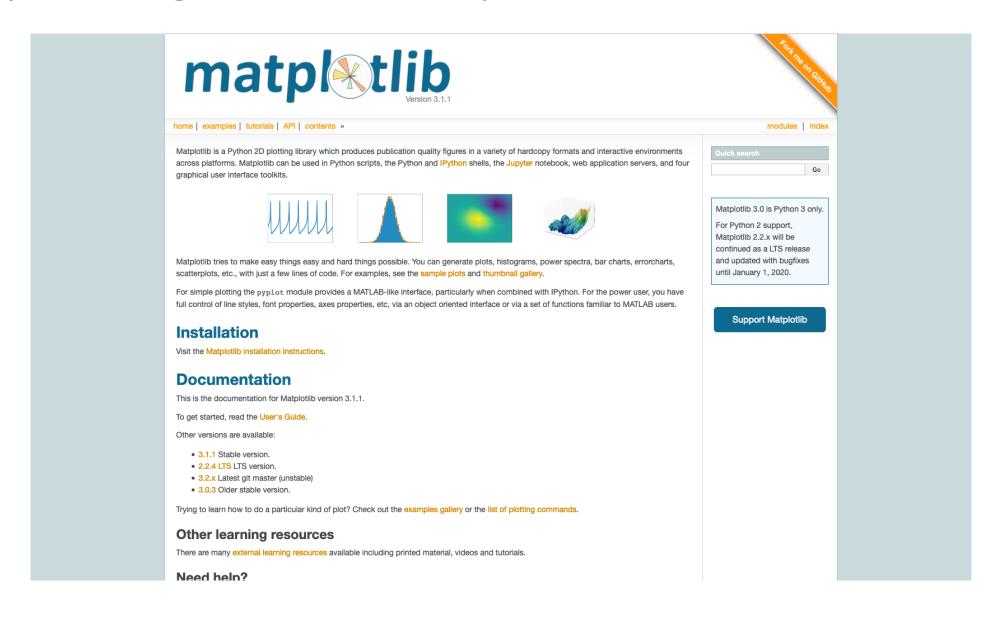
Previous Releases

0.25.0 - download // docs // pdf 0.24.2 - download // docs // pdf 0.24.1 - download // docs // pdf 0.24.0 - download // docs // pdf 0.23.4 - download // docs // pdf 0.23.3 - download // docs // pdf 0.23.2 - download // docs // pdf 0.23.1 - download // docs // pdf 0.23.0 - download // docs // pdf 0.22.0 - download // docs // pdf 0.21.1 - download // docs // pdf 0.21.0 - download // docs // pdf 0.20.3 - download // docs // pdf 0.19.2 - download // docs // pdf 0.18.1 - download // docs // pdf 0.17.1 - download // docs // pdf 0.16.2 - download // docs // pdf 0.15.2 - download // docs // pdf 0.14.1 - download // docs // pdf 0.13.1 - download // docs // pdf 0.12.0 - download // docs // pdf

ABOUT PANDAS

**** - 11 - K

Matplotlib: gráficos com Python



Scikit-learn: aprendizagem de máquina com Python



Classification

Identifying to which category an object belongs to.

Applications: Spam detection, Image recognition.

Algorithms: SVM, nearest neighbors, random forest, ... — Examples

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices. **Algorithms**: SVR, ridge regression, Lasso,

Examples

Clustering

Automatic grouping of similar objects into

Applications: Customer segmentation,
Grouping experiment outcomes
Algorithms: k-Means, spectral clustering,
mean-shift.... — Examples

Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency

Algorithms: PCA, feature selection, nonnegative matrix factorization. — Example:

Model selection

Comparing, validating and choosing parameters and models.

Goal: Improved accuracy via parameter

tuning
Modules: grid search, cross validation,

metrics. — Examples

Preprocessing

Feature extraction and normalization.

Application: Transforming input data such as text for use with machine learning algorithms.

Modules: preprocessing, feature extraction.

Example:

News

On-going development: What's new (Changelog)

Scikit-learn from 0.21 requires Python 3.5 or greater.

July 2019. scikit-learn 0.21.3 (Changelog) and 0.20.4 (Changelog) are available for download.

Community

About us See authors and contributing

More Machine Learning Find related projects

Questions? See FAQ and stackoverflow

Mailing list: scikit-learn@python.org

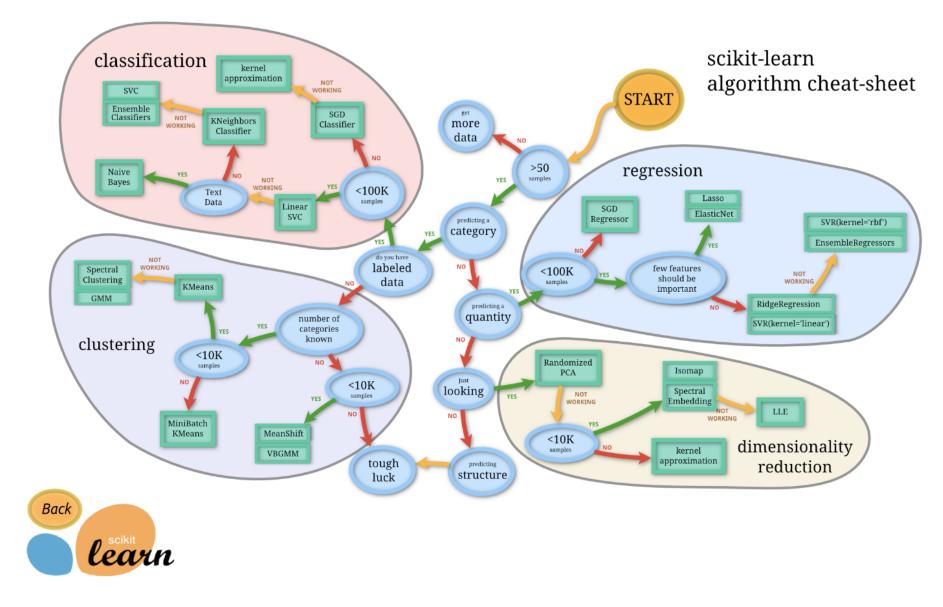
IRC: #scikit-learn @ freenode

Who uses scikit-learn?

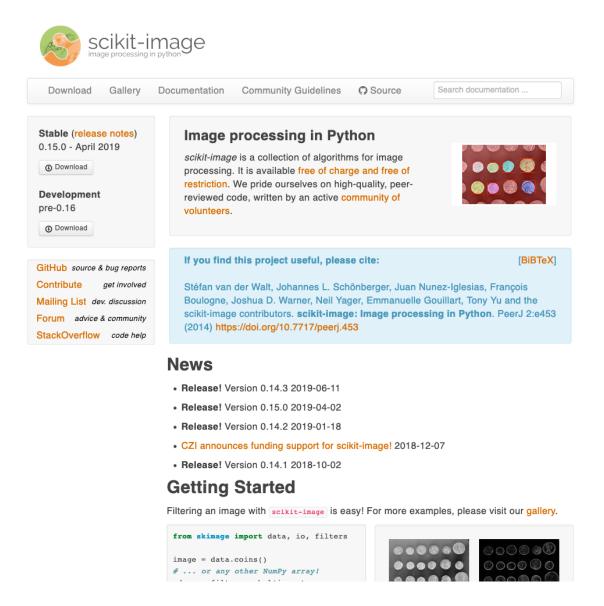


"The great benefit of scikit-learn is its fast

Scikit-learn: aprendizagem de máquina com Python



Scikit-image: processamento de imagem em Python



OPEN CHAT