Science: give DueCredit to software and methods developers!

Matteo Visconti di Oleggio Castello, Yaroslav O. Halchenko,
Dept. of Psychological and Brain Sciences, Dartmouth College

http://Neuro.Debian.net

http://CenterForOpenNeuroscience.org

Do you acknowledge the people who created your favorite toolbox?

Often users are not aware of all the methods they're using in their analysis, and thus can't properly acknowledge the software and method developers

Can you imagine your research without Open Source Software?

Without adequate citations for softwares and methods, funding agencies receive the false impression that scientific software is not important enough to be funded

Have you ever created a new project just to get your work cited?

Many times people (e.g., Ph.D. students) "re-implement the wheel" creating short-lived independent projects instead of contributing to existing ones, just because they need to get citations for their methods/software

Solution: a simple framework in Python to embed references in the code

User

```
# A tiny analysis script to demonstrate duecredit

from scipy.cluster.hierarchy import linkage
from scipy.spatial.distance import pdist
from sklearn.datasets import make_blobs

print("I: Simulating 4 blobs")
data, true_label = make_blobs(centers=4)

dist = pdist(data, metric='euclidean')

Z = linkage(dist, method='single')
print("I: Done clustering 4 blobs")
```

http://duecredit.org

"Injected" and supported packages include scipy, scikit-learn, numpy, mdp, dipy, nipype, pymvpa, ...

Devel

1. copy duecredit/stub.py to your codebase as due.py
2. import necessary pieces
   `from .due import due, Doi`
3. cite modules or functions using a doi or a BibTeX entry

```
# module
due.cite(Doi("1.2.3/x.y.z"),
         description="Solves all your problems",
         path="magicpy")

# functions
@due.dcite(Doi("1.2.3/x.y.z"),  description="Finds love")
def find_love():
    ...
```

Run it!

```
$> python -m duecredit examples/example_scipy.py
I: Simulating 4 blobs
I: Done clustering 4 blobs

DueCredit Report:
- Scientific tools library / numpy (v 1.10.4) [1]
- Scientific tools library / scipy (v 0.14) [2]
  - Single linkage hierarchical clustering /
scipy.cluster.hierarchy:linkage (v 0.14) [3]

2 packages cited
0 modules cited
1 function cited

References
----------

[1] Van Der Walt, S., Colbert, S.C. & Varoquaux, G., 2011. The
NumPy array: a structure for efficient numerical computation.
Computing in Science & Engineering, 13(2), pp.22–30.
[2] Jones, E. et al., 2001. SciPy: Open source scientific tools
for Python.
[3] Sibson, R., 1973. SLINK: an optimally efficient algorithm for
the single-link cluster method. The Computer Journal, 16(1),
pp.30–34.
```

BibTeX?

```
$> duecredit summary --format=bibtex
@article{van2011numpy,
  title={The NumPy array: a structure for efficient numerical computation},
  author={Van Der Walt, Stefan and Colbert, S Chris and Varoquaux, Gael},
  journal={Computing in Science \& Engineering},
  volume={13},
  number={2},
  pages={22--30},
  year={2011},
  publisher={AIP Publishing}
}
@Misc{JOP+01,
  author={Eric Jones and Travis Oliphant and Pearu Peterson and others},
  title={{SciPy}: Open source scientific tools for {Python}},
  year={2001--},
  url="http://www.scipy.org/",
  note={[Online; accessed 2015-07-13]}
}
@article{sibson1973slink,
  title={SLINK: an optimally efficient algorithm for the single-link cluster method},
  author={Sibson, Robin},
  journal={The Computer Journal},
  volume={16},
  number={1},
  pages={30--34},
  year={1973},
  publisher={Br Computer Soc}
}
```

What next?

Extend support to other languages (R, MATLAB)

Integration with DataLad for citation of datasets www.datalad.org

Spread the word and increase adoption!

Centralized system for usage stats of packages and software

Use it and/or contribute at github.com/duecredit

duecredit was born during the 2015 OHBM Hackaton. Thanks to the organizers!

Partially supported by NSF grant #1429999