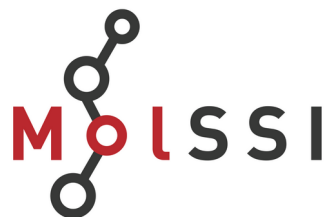


MolSSI Best Practices

Benjamin Pritchard

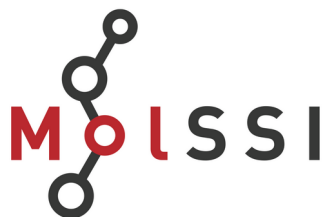
MolSSI Software Scientist

October 31, 2019



Computational Molecular Sciences

- Very old codebases
- Language diversity (Fortran, C/C++, Python)
- Fragmented development (100s of packages)
- Potentially wide userbase/overlap with other fields



OUT OF SORTS —

Researchers find bug in Python script may have affected hundreds of studies



"Willoughby-Hoye" scripts used OS call that caused Mojave

SEAN GALLAGHER - 10/15/2019, 10:17 AM

Science

Chemists bitten by Python scripts: How different OSes produced different results during test number-crunching

Boffins claim code was fine... when they wrote it

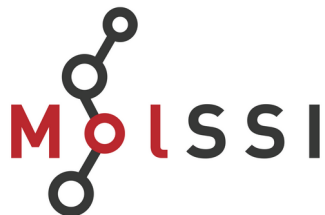
By Thomas Claburn in San Francisco 15 Oct 2019 at 04:56 157  SHARE 

MOTHERBOARD
TECHBYVICE


A Code Glitch May Have Caused Errors in More Than 100 Published Studies



The discovery is a reminder that science is collaborative and ideally self-correcting, but that nothing can be taken for granted.



The Basis Set Exchange

 **Basis Set Exchange** [Download](#) [GitHub](#) [Feedback](#) [About](#) [Help](#) [Request a Basis set](#)

All roles ▼ All ▼

- 2ZaPa-NR
- 2ZaPa-NR-CV
- 3-21G
- 3ZaPa-NR
- 3ZaPa-NR-CV
- 4-31G
- 4ZaPa-NR
- 4ZaPa-NR-CV
- 5-21G
- 5ZaPa-NR
- 5ZaPa-NR-CV
- 6-21G
- 6-31++G
- 6-31++G*
- 6-31++G**
- 6-31++G**-J
- 6-31+G
- 6-31+G*
- 6-31+G*-J
- 6-31+G**
- 6-311++G
- 6-311++G(2d,2p)
- 6-311++G(3df,3pd)
- 6-311++G*

search basis sets...

References for selected basis

Plain Text ▼ [Get References](#)

Total found: 559 basis sets [Select All](#) [Reset Selection](#)

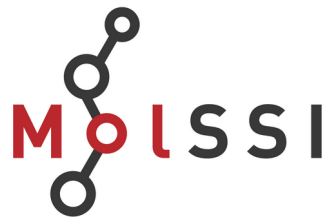
1																	2															
H																	He															
3	4											5	6	7	8	9	10															
Li	Be											B	C	N	O	F	Ne															
11	12											13	14	15	16	17	18															
Na	Mg											Al	Si	P	S	Cl	Ar															
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36															
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr															
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54															
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe															
55	56											72	73	74	75	76	77	78	79	80	81	82	83	84	85	86						
Cs	Ba											Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn						
87	88											104	105	106	107	108	109	110	111	112	113	114	115	116	117	118						
Fr	Ra											Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og						
																		57	58	59	60	61	62	63	64	65	66	67	68	69	70	71
																		La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
																		89	90	91	92	93	94	95	96	97	98	99	100	101	102	103
																		Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

Download basis set

Format NWChem ▼ [Get Basis Set](#) [Advanced](#)

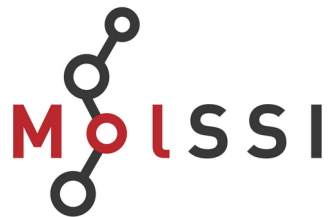
Best Practices

- Project Management
- Testing & CI
- Coding style & documentation



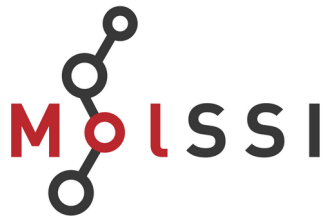
Preface

*Don't let the perfect be the enemy of
the good*



Project Planning

- Have a clear, *written* idea of the project and its scope
 - What is the minimum viable product (MVP)?
 - Is this worth creating a new project?
 - Who should be involved?
 - What happens when I leave or the project is discontinued?



Project Planning and Documents

This page contains the plans and outlines the new Basis Set Exchange project.

Goal

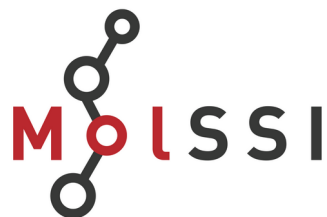
A rewrite of the current EMSL Basis Set Exchange database and website (<https://www.basissetexchange.org>)

Public Code Repository

The main code for this project is located on GitHub at https://github.com/MoISSI-BSE/basis_set_exchange. Comments and bugs/issues should be reported there.

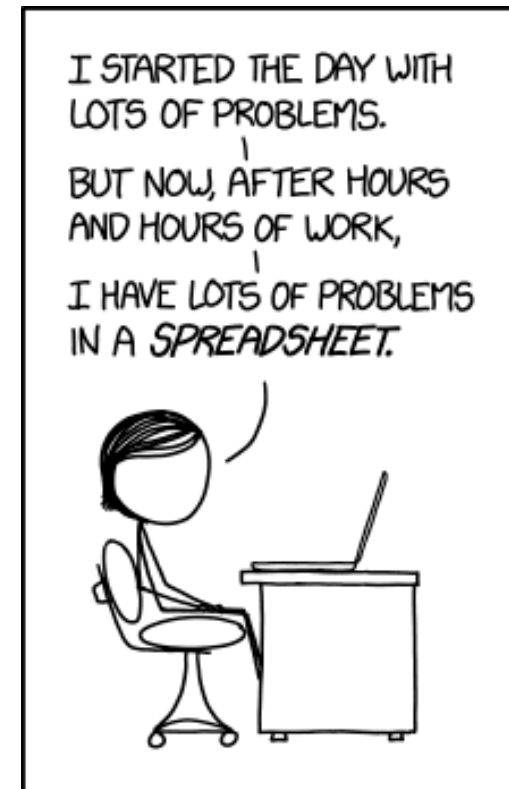
The main documentation is located at https://molssi-bse.github.io/basis_set_exchange

Additional related projects are located under the MoISSI-BSE GitHub organization at <https://github.com/MoISSI-BSE>



Project Management

- Keep track of issues, requests, bugs
- Roughly prioritize and estimate time required
- Let users know when something is fixed. Keep a changelog file
- Regularly review the current status



<https://xkcd.com/1906>



Basis Set Exchange

List Board Timeline Calendar Progress More...



On Track



Share

Search



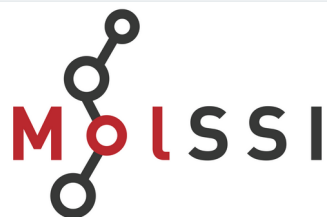
Upgrade



+ Add Task

Send feedback Incomplete tasks Filter Sort Rules Fields

Task name	Assignee	Issue/PR #	Due date	Est. Hours o...	+
▼ Papers					
▼ BSE Feedback					
✓ Rest of UGBS 1					
✓ TX93/PQS format 1					
✓ ANO-DK3 by T. Tsuchiya, M. Abe, T. N. and K. Hirao					
✓ CRYSTAL output format					
✓ Pseudopotential library from QMCPack group					
✓ 6-31G(2d,2p)					
✓ def2-SV and def2-TZV					
✓ 6-21+G* and 6-21G*					
✓ Consistent Gaussian basis sets for solid-state calculations:					



Changelog

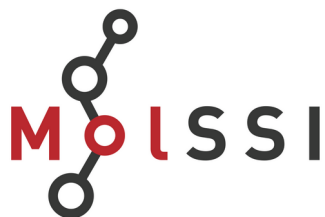
- Don't just rehash commit messages
- Organize and give context to your users

v0.8.X

-
- * Adds PQS output converter
 - * Adds Endnote reference converter
 - * Adds v1 basis sets
 - * Zorilla SB04-SZ-3G and SB04-DZ(d)-3G
 - * DFO of Porezag/Pederson
 - * ANO-DK3

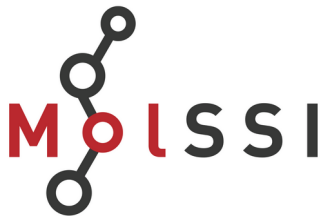
v0.8.10

-
- * Adds output converters
 - * Orca
 - * BDF
 - * AcesII
 - * Fixes issue in turbomole output (AM ordering)
 - * Fixes issue in gaussian output (missing blank line)
 - * Adds v0 basis sets
 - * Dunning PP fitting
 - * aug-cc-pwCVXZ-PP (X=D,T,Q,5)
 - * (aug)-cc-pwCVXZ-RIFIT (X=D,T,Q,5)
 - * Dunning OPTPT (including 512 and pp)



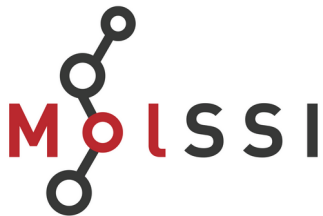
Testing/CI

- How can I ensure that...
 - ...my changes don't break anything
 - ...contributors don't break anything
 - ...my software works on different OSs
 - ...my software compiles with different compilers



Testing

- *Unit Tests* – Test a small piece of code (function, class)
- *Integration Tests* – Test composition of several units
- *Regression Test* – Ensure that a bug reported by a user remains fixed

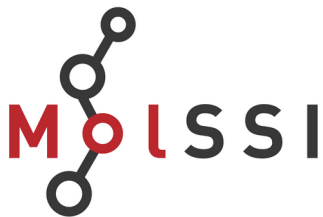


Testing

- Test failures as well as successes
- Add a (regression) test for bugs that are reported
- Don't aim for 100% coverage
- Have separate quick tests and longer tests

```
PASSED [ 99%]
tests/test_valid.py::test_valid_component[/home/ben/programming/bse_ng/basis_set_exchange/basis_set_exchange/data/truhlar/bse_v0/base/cc-pVDZ.0.json] PASSED
[ 99%]
tests/test_valid.py::test_valid_component[/home/ben/programming/bse_ng/basis_set_exchange/basis_set_exchange/data/crenb/CRENBL-ECP.0.json] PASSED [ 99%]
tests/test_valid.py::test_valid_component[/home/ben/programming/bse_ng/basis_set_exchange/basis_set_exchange/data/crenb/CRENBS.0.json] PASSED [ 99%]
tests/test_valid.py::test_valid_component[/home/ben/programming/bse_ng/basis_set_exchange/basis_set_exchange/data/crenb/CRENBL.0.json] PASSED [ 99%]
tests/test_valid.py::test_valid_component[/home/ben/programming/bse_ng/basis_set_exchange/basis_set_exchange/data/crenb/CRENBS-ECP.0.json] PASSED [ 99%]
tests/test_valid.py::test_valid_component[/home/ben/programming/bse_ng/basis_set_exchange/basis_set_exchange/data/orp/orp.1.json] PASSED [ 99%]
tests/test_valid.py::test_valid_reffile PASSED [100%]

===== 12846 passed in 314.73 seconds =====
```



Badges

MoLSSI-BSE / **basis_set_exchange** Unwatch 9 Star 42 Fork 14

Code Issues 16 Pull requests 1 Security Insights Settings

A repository for quantum chemistry basis sets https://molssi-bse.github.io/basis_se... Edit

Manage topics

838 commits 3 branches 39 releases 1 environment 7 contributors BSD-3-Clause

Branch: master New pull request Create new file Upload files Find file Clone or download

bennybp Update docs for endnote Latest commit 9f0e51c 23 hours ago

basis_set_exchange	Few more fixes for endnote (and add to converters.py)	23 hours ago
devtools	Update README.md	last year
doc	Update docs for endnote	23 hours ago
.codecov.yml	TEST: Updates relative imports	last year
.gitignore	git ignore all code in data dir	last month
.lgtm.yml	Add LGTM config	last year
.travis.yml	Add QCSchema to travis	5 months ago
CHANGELOG	Update changelog	yesterday
LICENSE	Update license	8 months ago
MANIFEST.in	Adds a manifest file	last year
README.md	Add link to DOI in README	yesterday
setup.cfg	Fix PytestUnknownMarkWarning (again)	4 months ago
setup.py	graphviz required for docs, too	4 months ago
versioneer.py	Update URLs [skip ci]	27 days ago

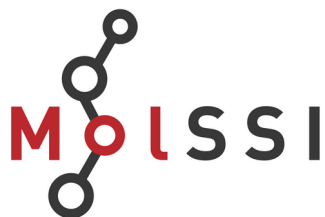
README.md Edit

build passing codecov 93% code quality: python A pypi v0.8.10

Basis Set Exchange Website

build passing codecov 93% code quality: python A pypi v0.8.10

https://github.com/MoLSSI-BSE/basis_set_exchange



Use a VCS

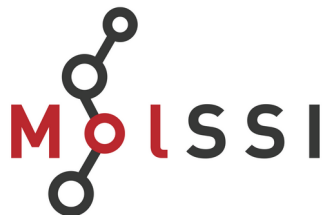
- Code needs to be kept in a *version control system*
 - git, svn, mercurial
- Tracks who made changes and what those changes were
- Conflict resolution – enables distributed development

```
commit 909189b8ee68b612be9f0e2d3904ec69b425930d (HEAD -> master)
Author: Benjamin Pritchard <ben@bennyp.org>
Date:   Wed Oct 30 11:59:18 2019 -0400

    Prune useless primitive from ANO-DK3

commit 9f0051cb08c4598b617b07db90b33d4eaad89e30 (origin/master)
Author: Benjamin Pritchard <ben@bennyp.org>
Date:   Mon Oct 28 15:00:36 2019 -0400

    Update docs for endnote
```



Coding Style/Documentation

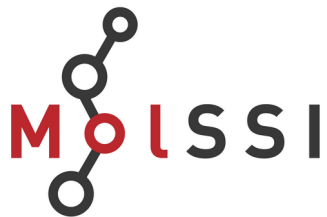
- Consistent style and documentation is important
- Generally no single 'right' way to do it



<https://xkcd.com/1695>

What to document

- Document how to use your code
- Also document error conditions
- Include examples, particularly common use cases
 - And test the examples!
- Keep documentation up-to-date



Enforcement

Don't get too hung up on enforcement

```
def factorial(n):  
    '''Calculate the factorial of an integer  
  
    Parameters  
    -----  
    n : int  
        The number to take the factorial of  
  
    Returns  
    -----  
    int  
        The factorial of the integer n  
    '''
```

Closing Thoughts

- Best practices depend on
 - Programming language
 - Project size/environment
 - Project scope
 - Target audience
- Take these as guidelines, adapt as necessary
- Take the time to learn your tools

