GenApp: Extensible Tool for Rapid Generation of Web and Native GUI Applications

Alexey Savelyev, Emre H Brookes

Gateways 2016
SDSC, San Diego
3 November 2016

SCHOOL OF MEDICINE
UT HEALTH SCIENCE CENTER®
DEPARTMENT OF BIOCHEMISTRY

Outline

- Background on GenApp
- Dynamic User interface from “static” definition files
- Alexey's perspective

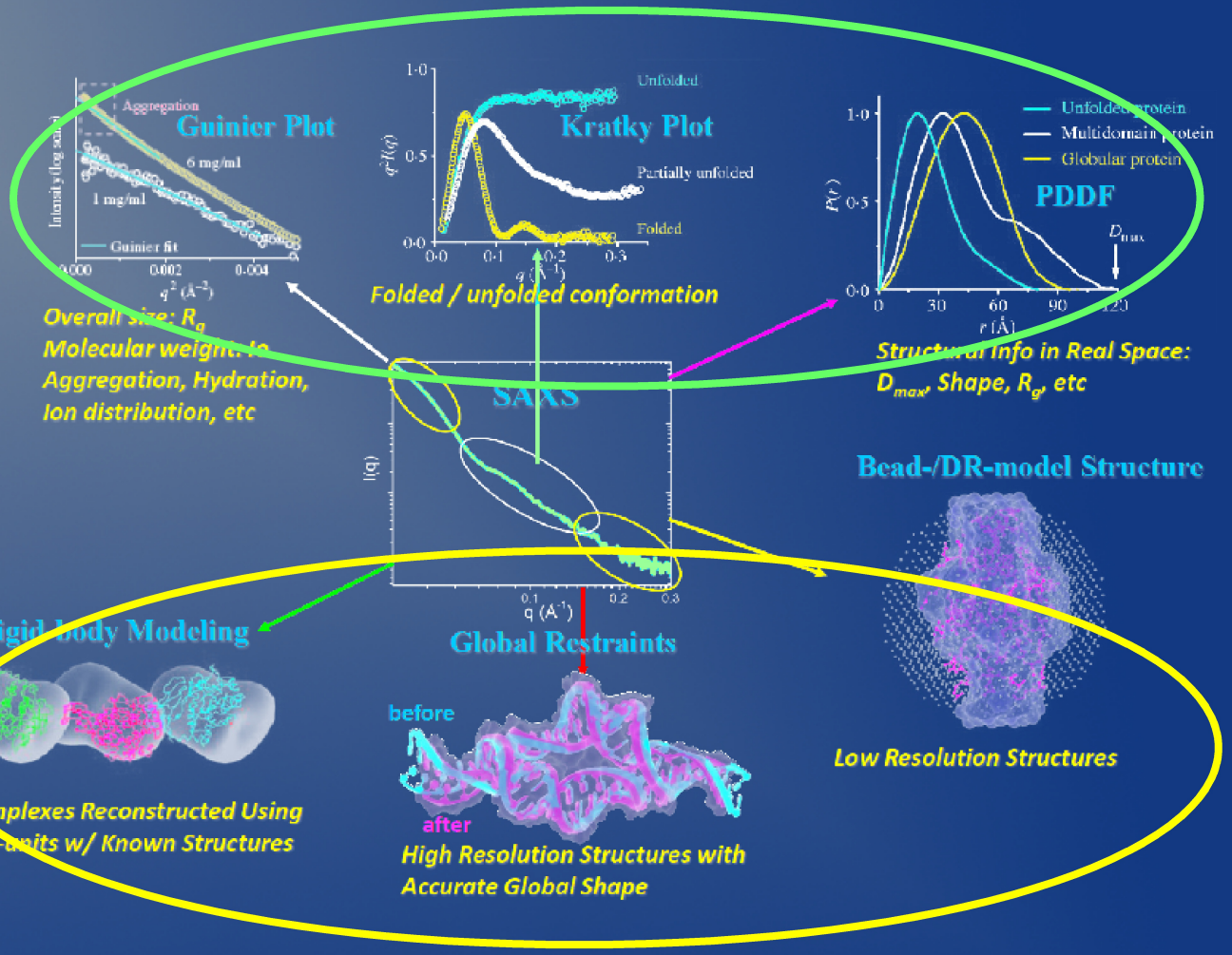
CCP-SAS

- SASSIE http://www.smallangles.net/sassie
  - Joseph Curtis et al.
  - **PYTHON**
  - **includes wrapped binary executables**
- SCT/SCTPL/HYDRO http://www.ucl.ac.uk/smb/perkins
  - Steve Perkins et al.
  - Structural Immunology Group at University College London
  - **FORTRAN**
- US-SOMO http://somo.uthscsa.edu

  - Emre Brookes et al.
  - **C++/Qt**
  - **includes wrapped binary executables**
- attract others …

Bio-SAS – a variety of time scales and computational requirements

sub-second to minutes
single core



hours to days on many cores (tens to hundreds)

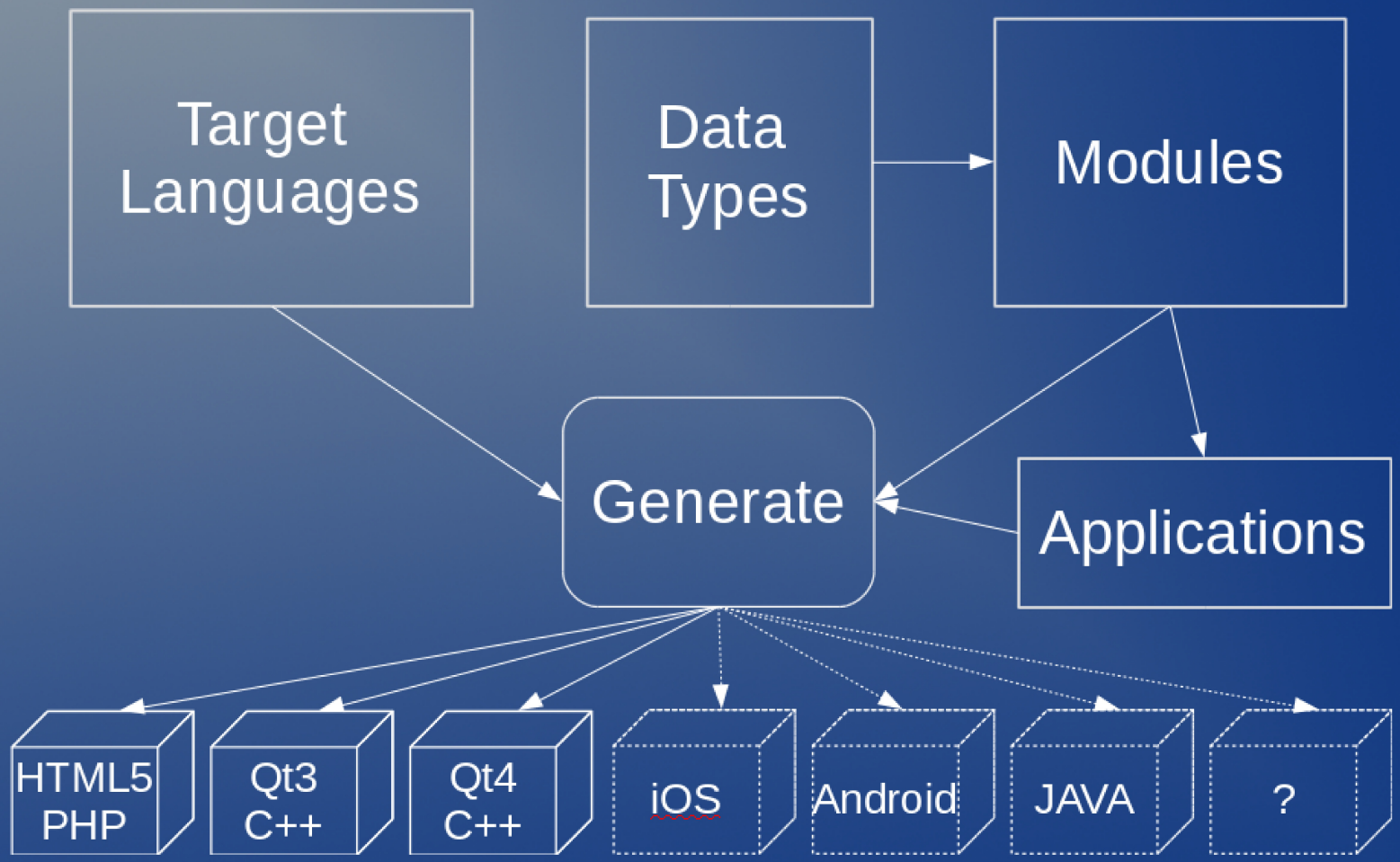
Considerations

- **Ease of deployment** in an ever-evolving software environment landscape
- Legacy and frequently specific lab developed codes
- Labs frequently can not afford a dedicated software team nor the cost in time and funding

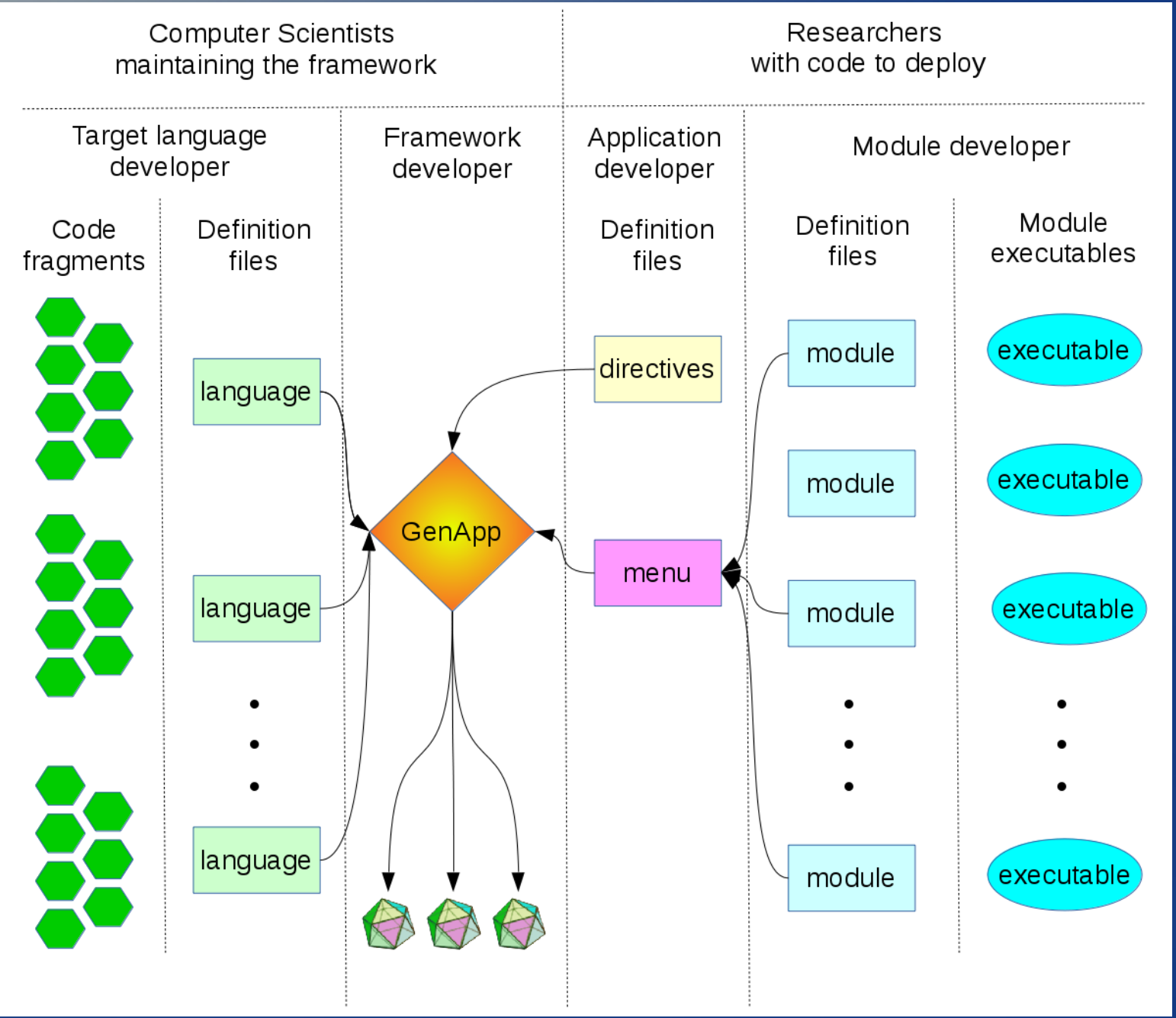
"Dark code"

An Open Extensible Multi-Target Application Generation Tool for Simple Rapid Deployment of Multi-Scale Scientific Codes

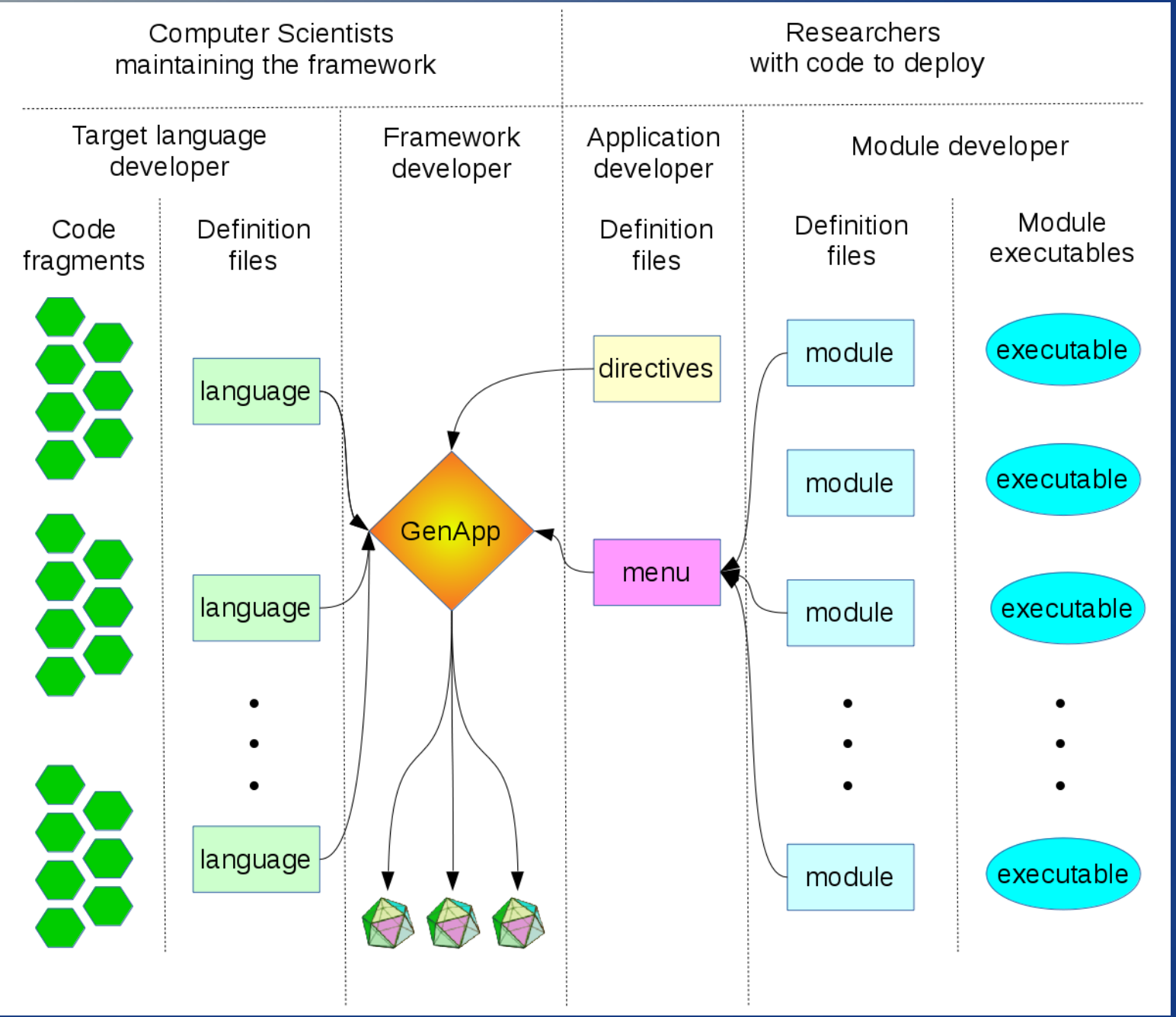
Brookes, E. H. XSEDE 14 Atlanta



Definition file driven generation



Insure preservation in two ways



GenApp from the “Researcher” perspective

- Given an existing executable:
  - Create a definition file describing inputs and outputs
  - Wrap or modify that executable to accept inputs and outputs as defined
  - Run the GenApp “compiler”
  - →

→ A fully functional Science Gateway

- Users, user management and statistics
- Job management, reattach
- “Cloud” file system
- Optional messaging for “live” updates
- Caching
- Multiple execution models
  - local to remote HPC and cloud
    - Airavata integration
    - OpenStack
- Can also simultaneously create “GUI” applications over the same modules
  - e.g. QT, JAVA
- Extensible!
  - features added on an as-needed basis

Some GenApp generated Gateways

In production:

SASSIE-web (J.E. Curtis)	https://sasssie-web.chem.utk.edu/sassie2 *40 modules*
SCT (S. Perkins)	*350+ users, 10,000+ jobs in 2016, 40+ papers*
BayesApp (S. Hansen)	http://genapp.rocks/bayesapp
Denfert (J. Perez)	http://genapp.rocks/denfert
Vortex Shedding (A. Perlstein)	http://genapp.rocks/vortexshedding

Alpha

NAMDrunner (A. Savelyev)	http://genapp.rocks/namdrunner

Dynamic UI's from static definition files

```
{
    "moduleid"     : "energy"
    ,"label"       : "Energy"
    ,"help"        : "help for Energy"
    ,"executable"  : "energy"
    ,"fields"      : [
        {
            "role"        : "input"
            ,"id"         : "m"
            ,"label"      : "mass [kg]"
            ,"type"       : "float"
            ,"required"   : "true"
            ,"help"       : "Enter the mass in kilograms"
        }
        ,{
            "role"        : "input"
            ,"id"         : "c"
            ,"label"      : "Speed of light [m/s]"
            ,"type"       : "float"
            ,"default"    : 299792458
            ,"required"   : "true"
            ,"help"       : "Enter the speed of light in meters/second"
        }
        ,{
            "role"        : "output"
            ,"id"         : "e"
            ,"label"      : "Energy [J]"
            ,"type"       : "text"
        }
    ]
}
```

Dynamic UI's from static definition files

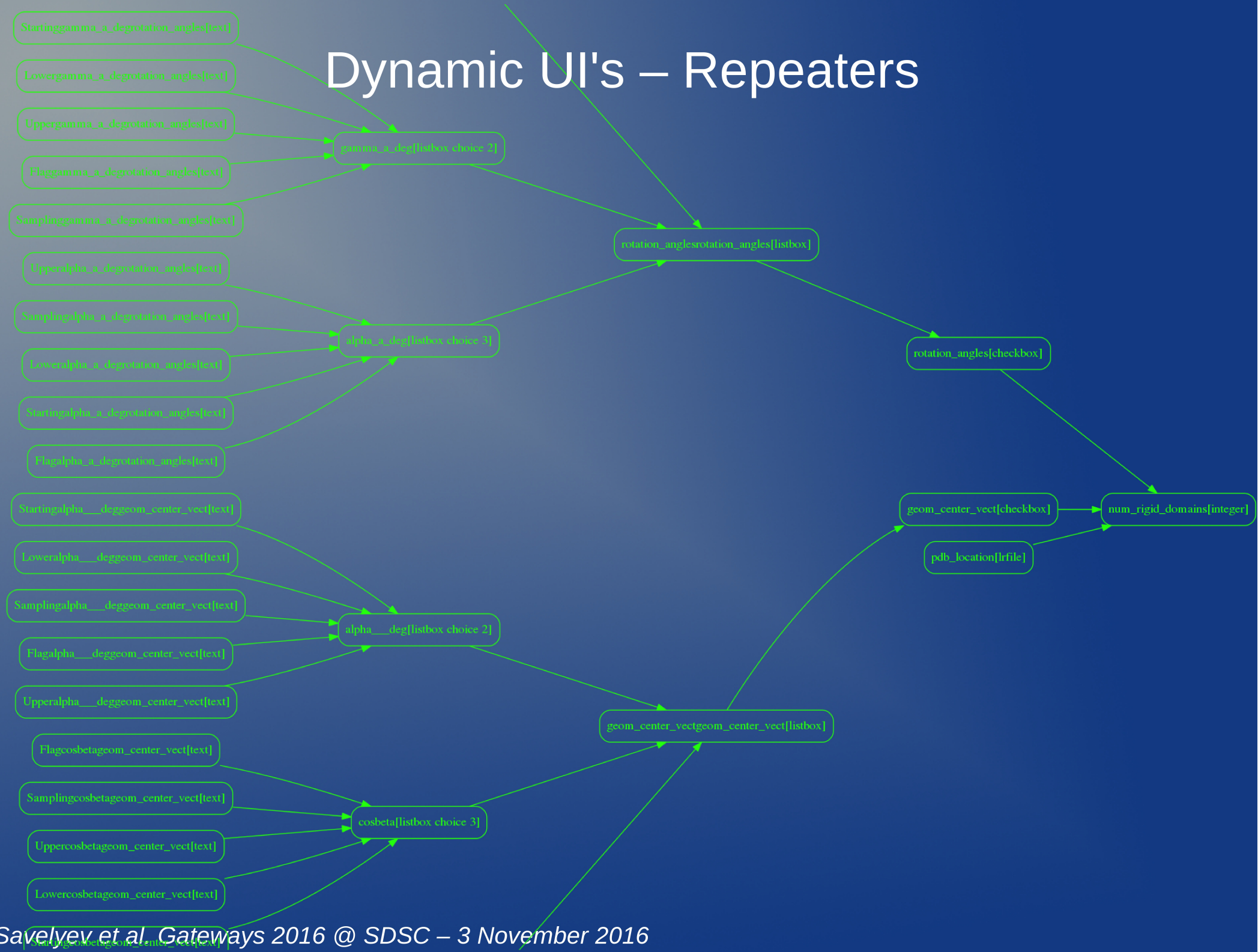
- How to best define interactivity in a single definition file that can be handled within the UI?

Dynamic UI's – repeaters & repeats

```
{
...
    ,"fields"      : [
        {
            "role"         : "input"
            ,"id"          : "count"
            ,"label"       : "How many calculations?"
            ,"type"        : "integer"
            ,"repeater"    : "true"
        }
        ,{
            "role"         : "input"
            ,"id"          : "m"
            ,"label"       : "mass [kg]"
            ,"type"        : "float"
            ,"repeat"      : "count"
        }
        ,{
            "role"         : "input"
            ,"id"          : "c"
            ,"label"       : "Speed of light [m/s]"
            ,"type"        : "float"
            ,"default"     : 299792458
        }
…
    ]
}
```

- Integer
- Listbox
- Checkbox

Dynamic UI's – Repeaters



Dynamic UI's – Calculated fields

```
{
 ...
    ,"fields"      : [
        {
            "role"         : "input"
            ,"id"          : "m"
            ,"label"       : "mass [kg]"
            ,"type"        : "float"
            ,"required"    : "true"
            ,"help"        : "Enter the mass in kilograms"
        }
        ,{
            "role"         : "input"
            ,"id"          : "c"
            ,"label"       : "Speed of light [m/s]"
            ,"type"        : "float"
            ,"default"     : 299792458
            ,"required"    : "true"
            ,"help"        : "Enter the speed of light in meters/second"
        }
        ,{
            "role"         : "input"
            ,"id"          : "e"
            ,"label"       : "Energy [J]"
            ,"type"        : "text"
            ,"calc"        : "m*c^2"
        }
    ]
}
```

- Can be chained
- Can be under repeaters – context is managed

Resources

http://genapp.rocks & esp. ***http://genapp.rocks/wiki***

- Primary host at University of Tennessee Knoxville
  - "Entropy" server 128 core, 256 GB ram, 8 Tesla K20m GPUs, Rocks OS
  - Running HTML5/PHP
- Indiana University Quarry nodes
  - Trac wiki with integrated subversion repository
  - Testing node
- XSEDE Jetstream
  - XSEDE ECSS vortexshedding gateway
  - SASSIE-web instance
  - NAMDrunner
- AWS
- XSEDE TG-MCB140255 *Computational support for small angle scattering for advanced analyses of structural data in chemical biology and soft condensed matter*
- ORNL Titan, UK SCARF (pending)

Resources – Personnel

- Alexey Savelyev
  - Target language developer
  - Application developer
  - Module wrapper
- Emre Brookes
  - Primary developer
- External Application/Module developers
  - Joseph Curtis, PI and personnel
    - SASSIE
  - David Wright
    - SCT
  - Arne Perlstein, PI's assigned personnel (currently Josef Sabuda)
    - Vortexshedding

Future

- Ongoing
  - More application & module wrapping
    - Training others to wrap
      (and making it as easy as possible)
    - We could wrap a lot given sufficient resources
  - Easy install of applications for setup of web-servers and standalone
- Near future
  - Instance generator / VM's and/or simple JSON
  - Identity management
  - Module portal
  - API access
  - Containerization
- Further out
  - Apache membership
  - Automatic regression testing

GenApp Based Science Gateways

- **US-SOMO** https://somo.chem.utk.edu/somo
- **WillItFit** https://somo.chem.utk.edu/willitfit
- **QuaFit** https://somo.chem.utk.edu/quafit
- **Mulch** https://somo.chem.utk.edu/mulch
- **ParamMD** https://somo.chem.utk.edu/parammd
- **NAMDrun** http://js-170-47.jetstream-cloud.org/namdrunner

(Nov 2015 - present)

- **SASSIE** https://somo.chem.utk.edu/sassie2
- **Denfert** https://somo.chem.utk.edu/denfert
- **BayesApp** https://somo.chem.utk.edu/bayesapp
- **Vortexshedding** http://js-172-198.jetstream-cloud.org/vortexshedding

Encountering GenApp (... as a researcher ...)

Underlying codes written in different languages, can be GUI or command line applications:

- C++, Python, Fortran etc.
- C++ wrapped in Python (WillItFit); C++ bundled with Qt (US-SOMO)

“Divorce” GUI from the computational component:

- Create a “command-line” analogue: [./program < {arguments}]

Wrap a command-line application:

- Application arguments and output are JSON strings of key-value pairs, {“binsize_id”: 0.2, “filename_id”: “trajectory.dcd” etc.}
- These key-value pair are described in the application (module) definition file
- Different languages can be used to parse JSON input and transfer it to the underlying application (Python, Perl, C++)

“GenApp”, i.e. compile application and generate a web site.

Encountering GenApp (... as a researcher ...)

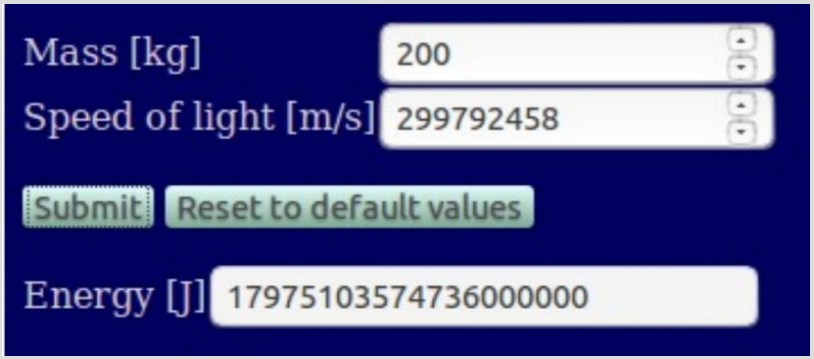
JSON parsers – different languages:

PYTHON

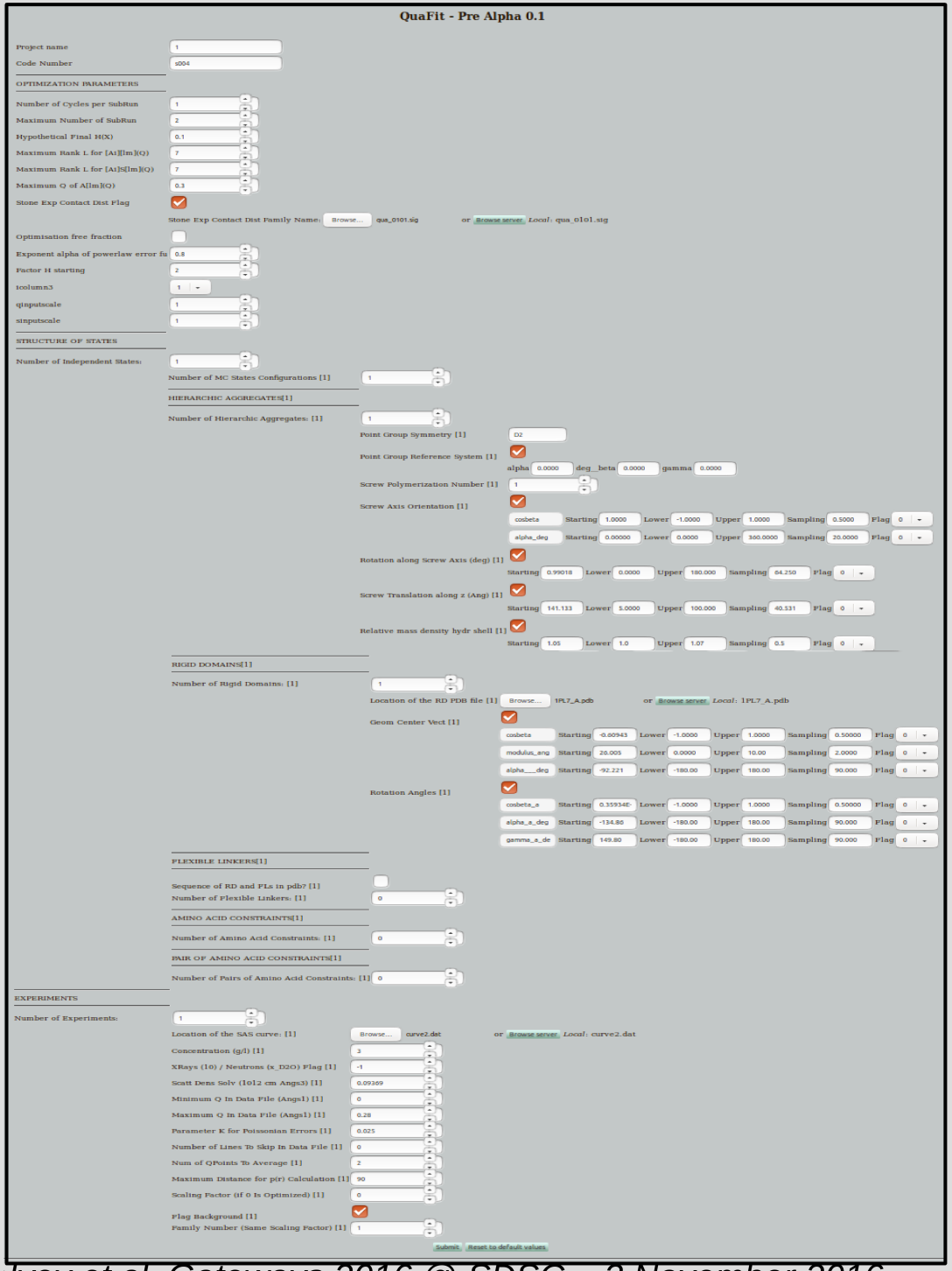
```python
#!/usr/bin/python

import json
from StringIO import StringIO

if (len(sys.argv) < 1):
    print "\{\"error\":\"called with no arguments\"\}\n"
    exit()

json_variables = " "

argv_io_string = StringIO(sys.argv[1])
json_variables = json.load(argv_io_string)

mass        = json_variables['mass']
speed_light = json_variables['speed_of_light']

#.... Energy Calculation .... #

output_res = {}
output_res[ 'energy' ] = energy

print json.dumps( output_res )
```

PERL

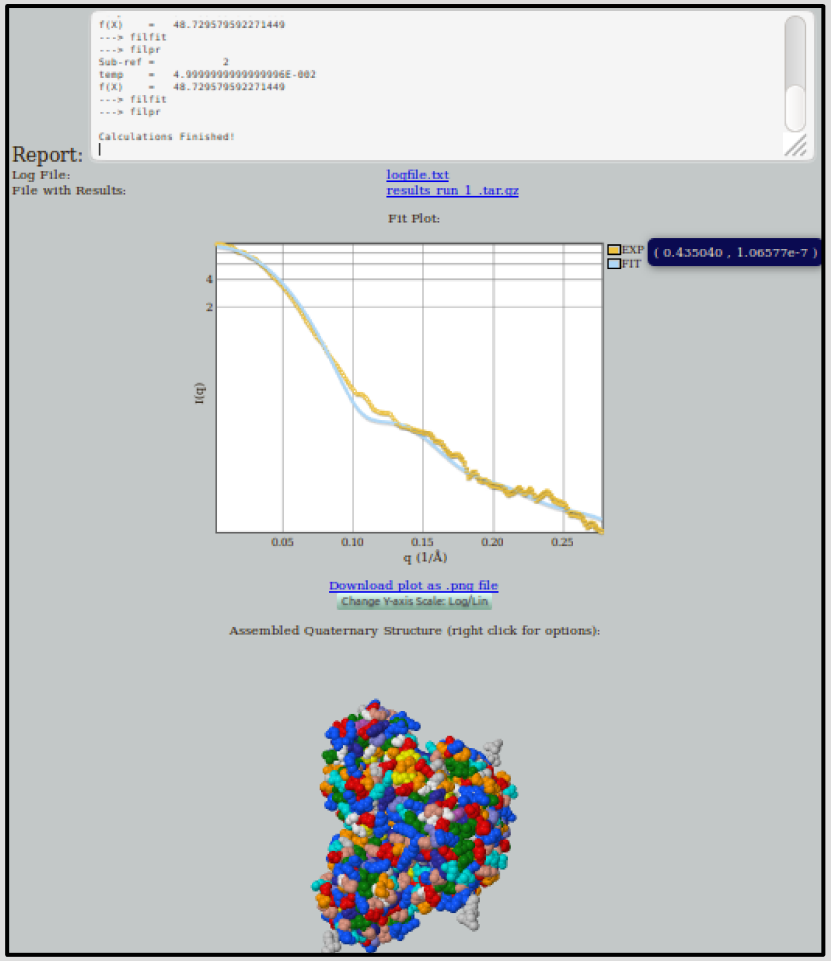
```perl
#!/usr/bin/perl

use JSON;

if ( !@ARGV )
{
    print "\{\"error\":\"called with no arguments\"\}\n";
    exit;
}

$json = shift;
$json_variables = decode_json( $json );

$mass = $$json_variables{ "mass" };
$speed_light = $$json_variables{ "speed_of_light" };

#.... Energy Calculation .... #

$output_res = {};
$$output_res{ "energy" } = energy;

print encode_json( $output_res ). "\n" ;
```



Input menu for 'QuaFit' module



Output of the 'QuaFit' module



QuaFit - Beta:

- Underlying code: Fortran
- Wrapped in PERL (~2000 lines)

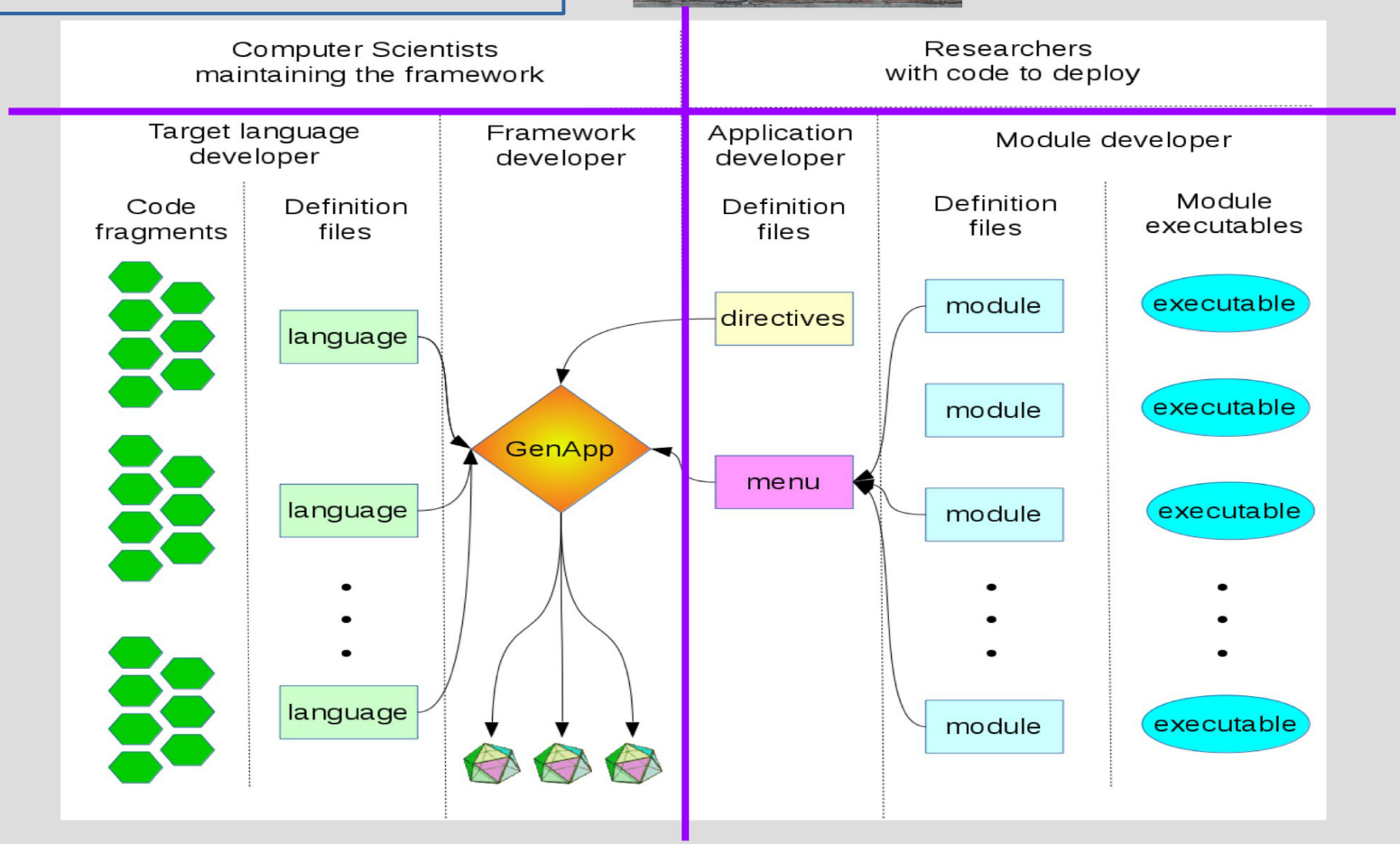
Encountering GenApp: Extending Horizon

Cool stuff:

- Modifying fragments of code
- Adding new /Extending types
- Conditional code generation
- Browser compatibility
- Admin utilities
- ….

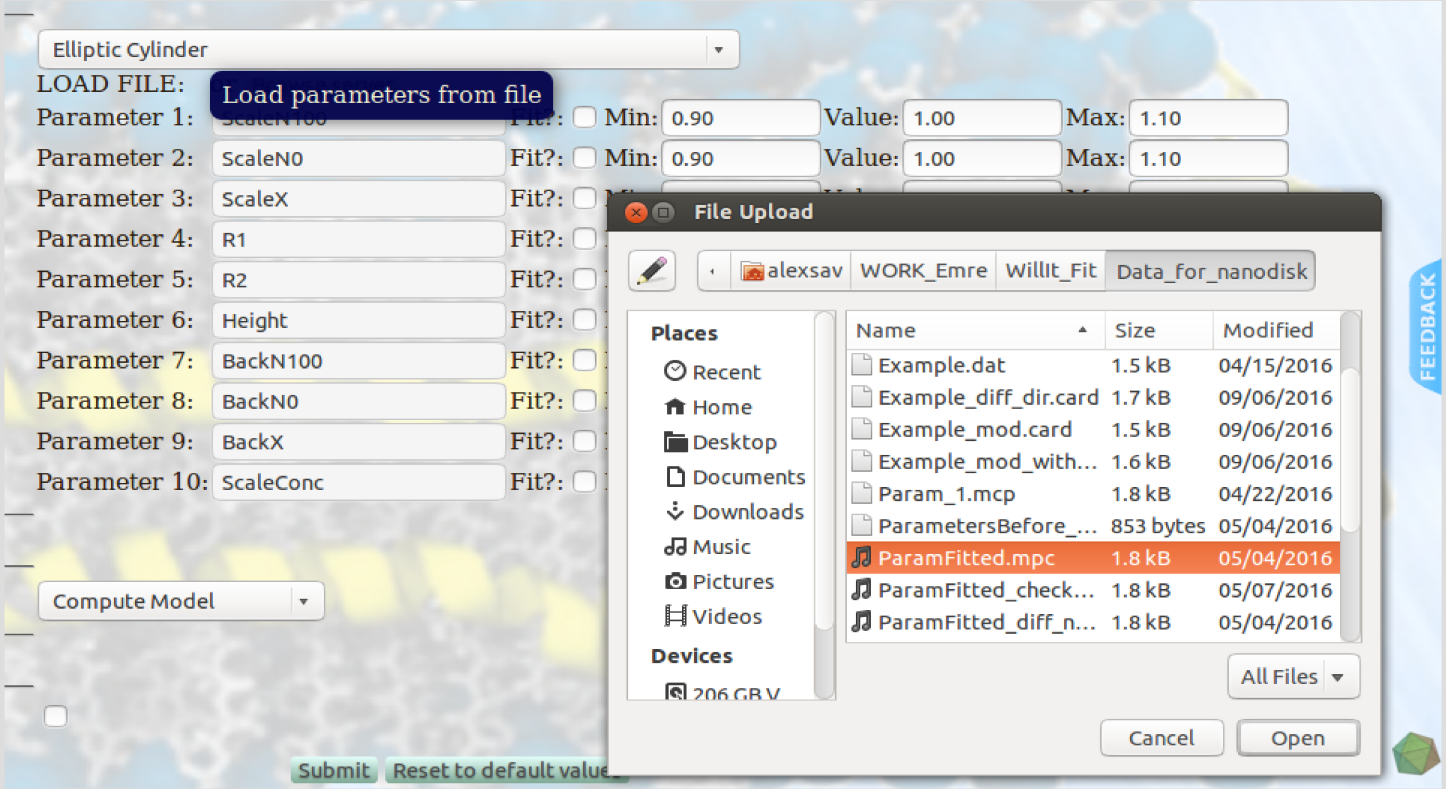
Web Development:

HTML5
CSS
JavaScript (jQuery)
PHP
MongoDB



Advanced “GenApping”: *WillItFit (1)*

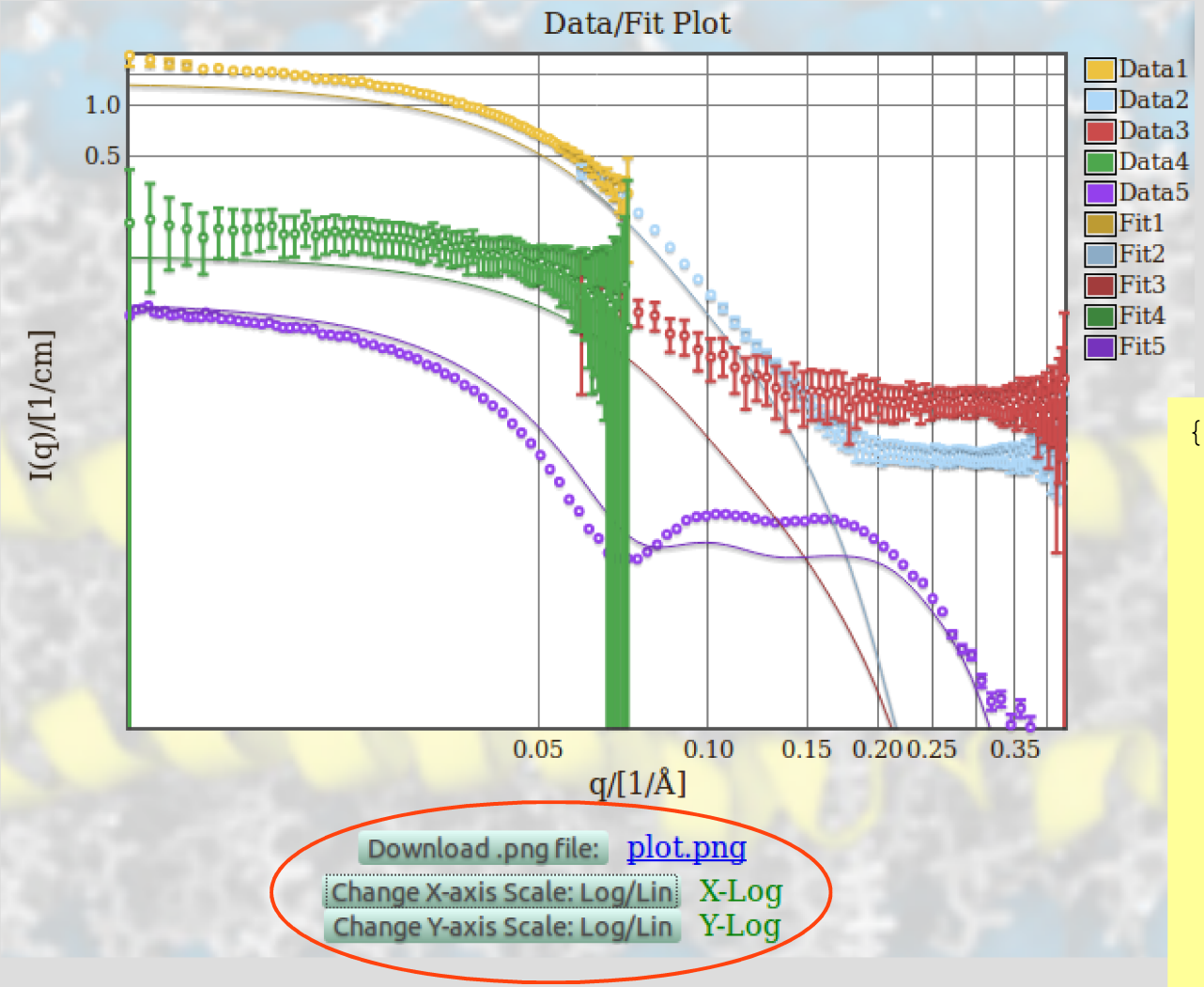
- Uploading parameter file, `lrfile.input` (numerical values, whitespace_separated etc.)




```
{
  "role"  :  "input",
  "id"    :  "param_choice_file_anis_core_shell_micelles",
  "label" :  "LOAD FILE: ",
  "type"  :  "lrfile",
  "setinputfromfile"    : "whitespaceseparated_reverselogic",
  "setinputfromfileids": "..list_of_ids..",
  "repeat":  "model_list_box:anis_core_shell_micelles",
  "help"  :  "Load parameters from file"
}
```

Advanced "GenApping": *WillItFit (2)*

- Advanced `Plot2D.output` options (save to file, change X-, Y-axis scales independently)



jQuery FLOT package
http://www.flotchart.org

```
{
  "role"  : "output",
  "id"    : "fit_plot",
  "label" : "Fit Plot:",
  "type"  : "plot2d",
  "height": "400px",
  "width" : "700px",
  "pan"   : "false",
  "zoom"  : "false",
  "backgroundcolor" : "white",
  "selzoom"         : "true",
  "changescalex"    : "true",
  "changescaley"    : "true",
  "savetofile"      : "true",
  "rotatedylabel"   : "true",
  "hover"           : "true"
}
```

Advanced “GenApping”: MULCh *(1)*

- Reading complex input files, `lrfile.input` (nested repeaters, non-numerical fields)

```
KinA:SDA 2:2 complex in 200mM NaCl, 50mM Tris, 150mM imidazole
7                                   # Number of contrast points
0.00  0.541   0.002 11.9      # D2O fraction, I(0), I(0) error, concentration (arb units.)
0.10  0.352   0.003 11.9
0.20  0.223   0.002 11.9
0.40  0.0745  0.002 26.9
0.80  0.19    0.001 11.9
0.90  0.332   0.002 11.9
1.00  0.537   0.001 11.9
3                                   # Number of things disolved in water
0.200 M NaCl            0.0         # Conc., M=molecule, formula, volume
0.050 M C4H11NO3        0.0         # "       " Tris http://www.jtbaker.com/msds/englishhtml/t7111.htm
0.150 M C3H4N2          0.0         # "       " Imidazole
0.95                                # % of the exchangables that are accessible by the solvent
0.0                                 # % of the non-exch. protons in fragment 1 replaced by deuterons
1                                   #number of components, (next line) number of molecules, P=protein, sequence, volume
2 P GSHMTEELMLKSEKLSIAGQLAAGIAHEIRNPLTAIKGFLQLMKPTMEGNEHYFDIVFSELSRIELILSELLMLAKVK
EYLNLKKLIGEVSALLETQANLNGIFIRTSYEKDSIYINGDQNQLKQVFINLIKNAVESMPDGGTVDIIITEDEHSVHVTVKDE
GEGIPEKVLNRIGEPFLTTKEKGTGLGLHPEKGTAFKISFPKK  0.0
1.00                            # % of the exchangables that are accessible by the solvent
0.85                            # % of the non-exch. protons in fragment 2 replaced by deuterons
1
2 P GSMRKLSDELLIESYFKATEMNLNRDFIELIENEIKRRSLGHIISVSS 0.0
```

Advanced “GenApping”: MULCh *(1)*

- Reading complex input files, `lrfile.input` (nested repeaters, non-numerical fields)

Project Title: KinA:SDA 2:2 complex in 200mM NaCl, 50mM Tris, 150mM imidazole

Upload Contrast File: or Browse server *Local*: contrast.txt

CONTRAST POINTS

Use Contrast Points? ☑

Number of Contrast Points: 7

f_{D_2O}(0-1)	$I(0)$	$\sigma[I(0)]$	Protein Conc.
0.00	0.541	0.002	11.9
0.10	0.352	0.003	11.9
0.20	0.223	0.002	11.9
0.40	0.0745	0.002	26.9
0.80	0.19	0.001	11.9
0.90	0.332	0.002	11.9
1.00	0.537	0.001	11.9

```
{
  "role"    : "input",
  "id"      : "contrast_file",
  "label"   : "Upload Contrast File",
  "type"    : "lrfile",
  "setinputfromfile":
"whitespace_formulchcontrast",
  "setinputfromfileids": "title,
                  contrast_points_num,
                  num_species_solv,
                  proton_frac,
                  deuter_level,
                  …."
}
```

DISSOLVED SPECIES

Number of Dissolved Species in the Solvent: 3

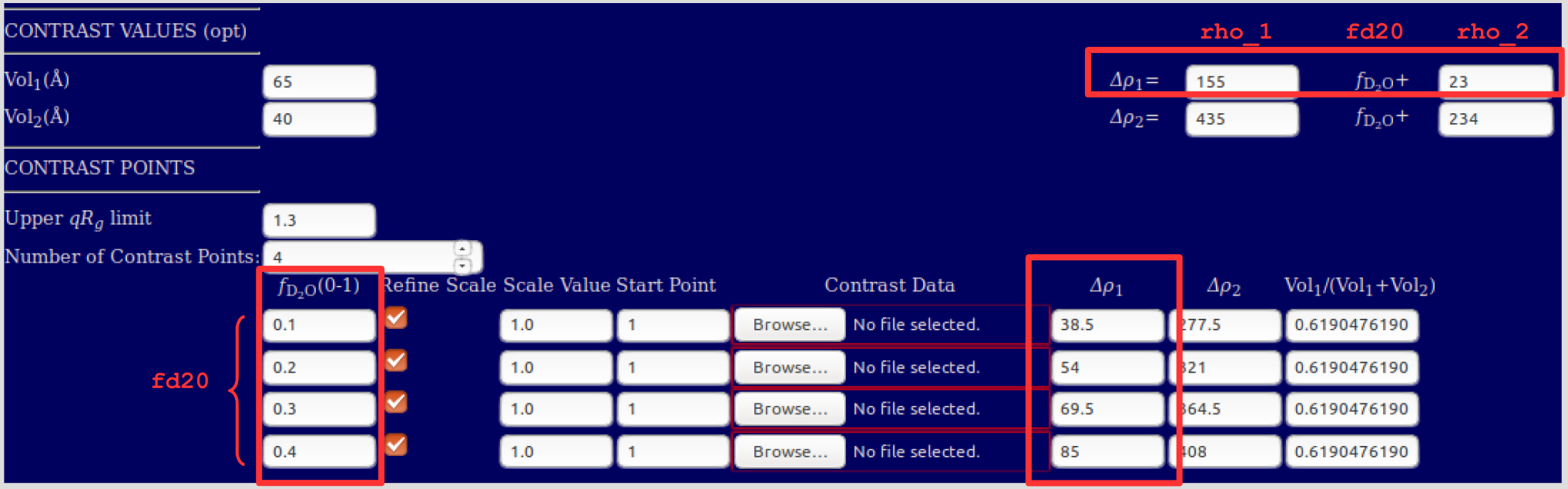
Substance Type: [1] Molecule

Formula: [1] NaCl Concentration: [1] 0.200 Volume: [1] 0.0

Substance Type: [2] Molecule

Advanced “GenApping”: MULCh *(2)*

- Calculated fields (inter-dependent numerical fields)

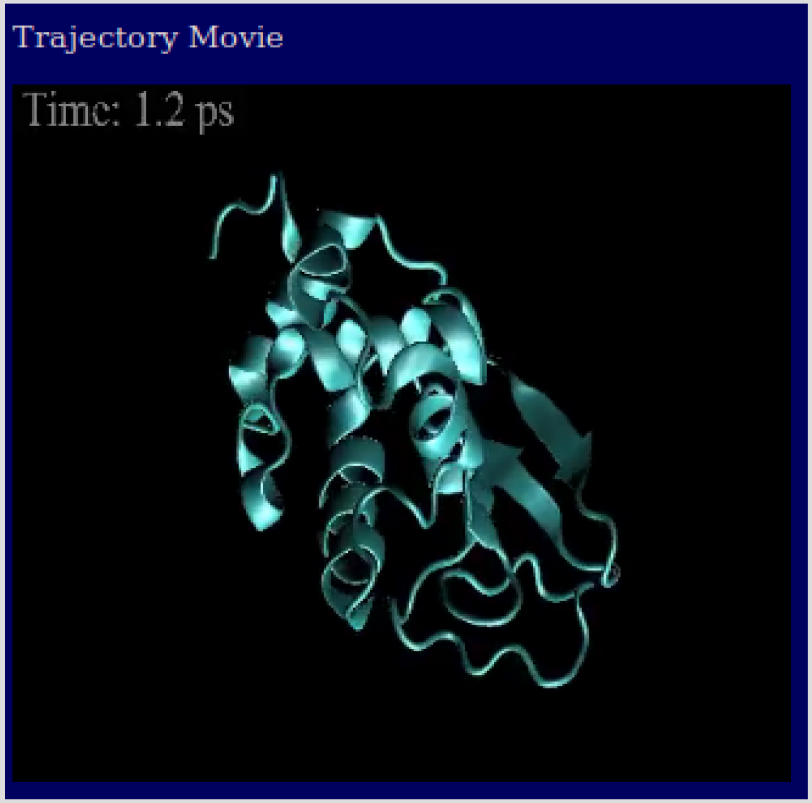



```
{
     "role"     : "input",
     "id"       : "delta_rho_1",
     "label"    : "Delta rho",
     "type"     : "float",
     "repeat"   : "contrast_points",
     "required" : "true",
     "calc"     : "rho_1*fd2o + rho_2",
 }
```

Advanced “GenApping”: ParamMD *(1,2)*

- Trajectory Movie generation [.mp4, .webm]

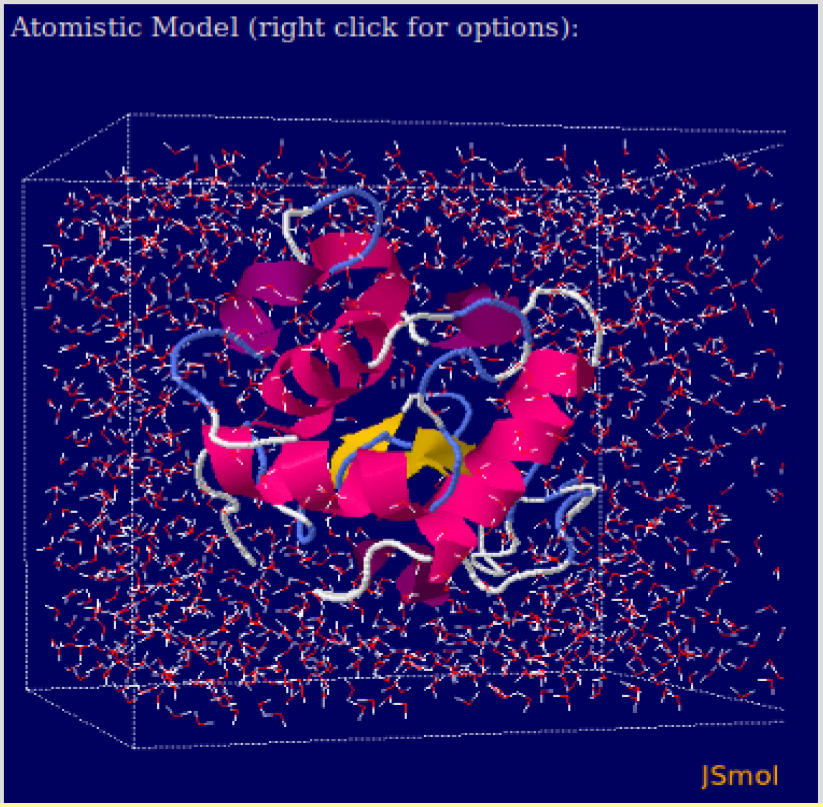
`video.output` (new type added)




```
{
  "role"   : "output",
  "id"     : "vid",
  "type"   : "video",
  "label"  : "Trajectory Movie",
  "width"  : 400,
  "height" : 450
}
```

- Structure visualization, JSmol

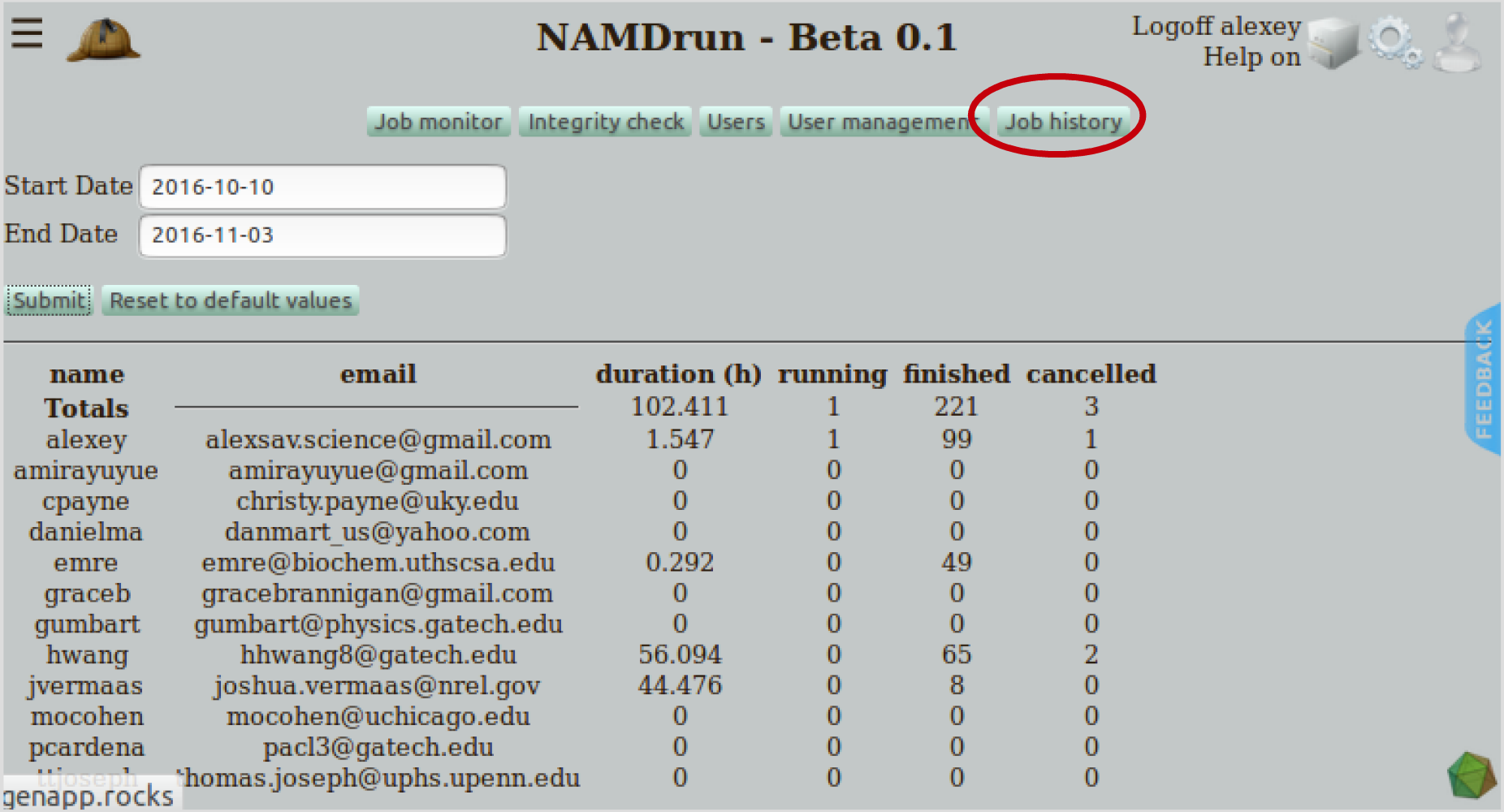
`atomicstructure.output`




```
{
   "role"     : "output",
   "id"       : "outputpdb_view",
   "label"    : "Atomistic Model",
   "type"     : "atomicstructure",
   "jsmoladd" : "hide HOH;spin on",
   "width"    : 450,
   "height"   : 450
}
```

Advanced “GenApping”: Admin Utilities

- Job History, jobs information within specified time frame (PHP, MongoDB)



name	email	duration (h)	running	finished	cancelled
Totals		102.411	1	221	3
alexey	alexsav.science@gmail.com	1.547	1	99	1
amirayuyue	amirayuyue@gmail.com	0	0	0	0
cpayne	christy.payne@uky.edu	0	0	0	0
danielma	danmart_us@yahoo.com	0	0	0	0
emre	emre@biochem.uthscsa.edu	0.292	0	49	0
graceb	gracebrannigan@gmail.com	0	0	0	0
gumbart	gumbart@physics.gatech.edu	0	0	0	0
hwang	hhwang8@gatech.edu	56.094	0	65	2
jvermaas	joshua.vermaas@nrel.gov	44.476	0	8	0
mocohen	mocohen@uchicago.edu	0	0	0	0
pcardena	pacl3@gatech.edu	0	0	0	0
ttjoseph	thomas.joseph@uphs.upenn.edu	0	0	0	0

In closing

- GenApp produces working science gateways & local GUI apps
- Easily extensible
- Advancements are requirements driven
- So let us know your requirements!

Special Thanks

- Gateways 2016 organizing committee and the Gateways Institute

Thanks for attending

Questions: ask now or email me at emre@biochem.uthscsa.edu

Acknowledgments

- Joseph Curtis
- To E. Brookes
  - NIH NIGMS K25GM090154 / PI
  - NSF CHE-1265817 / PI
- NSF XSEDE TG-MCB140255
- UTHSCSA BCF
- To N. Anjum, A. Kapoor, P. Patra
  - Google Summer of Code

Funded by a grant from the National Institute of General Medical Sciences of the National Institutes of Health

SCHOOL OF MEDICINE
UT HEALTH SCIENCE CENTER®
DEPARTMENT OF BIOCHEMISTRY