

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



Outline

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

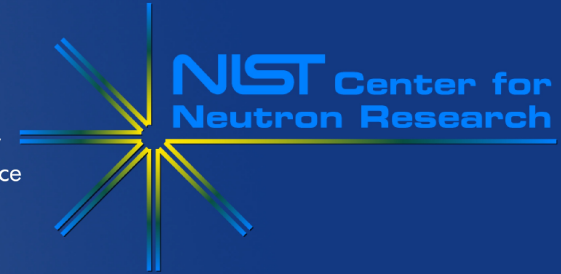
CCP-SAS

- **SASSIE** <http://www.smallangles.net/sassie>

- Joseph Curtis et al.

NIST

National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce



- **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**

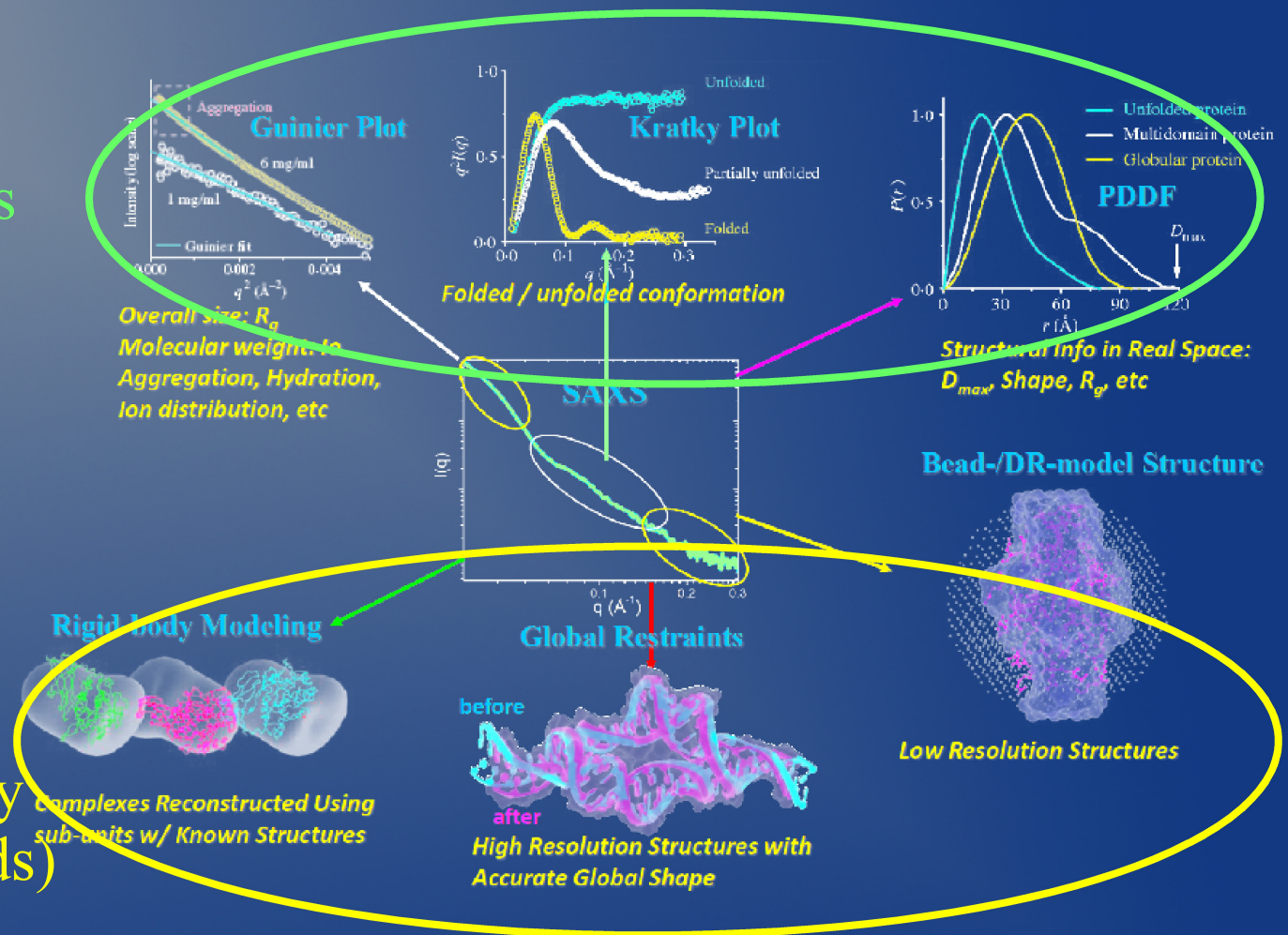
- attract others ...



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)



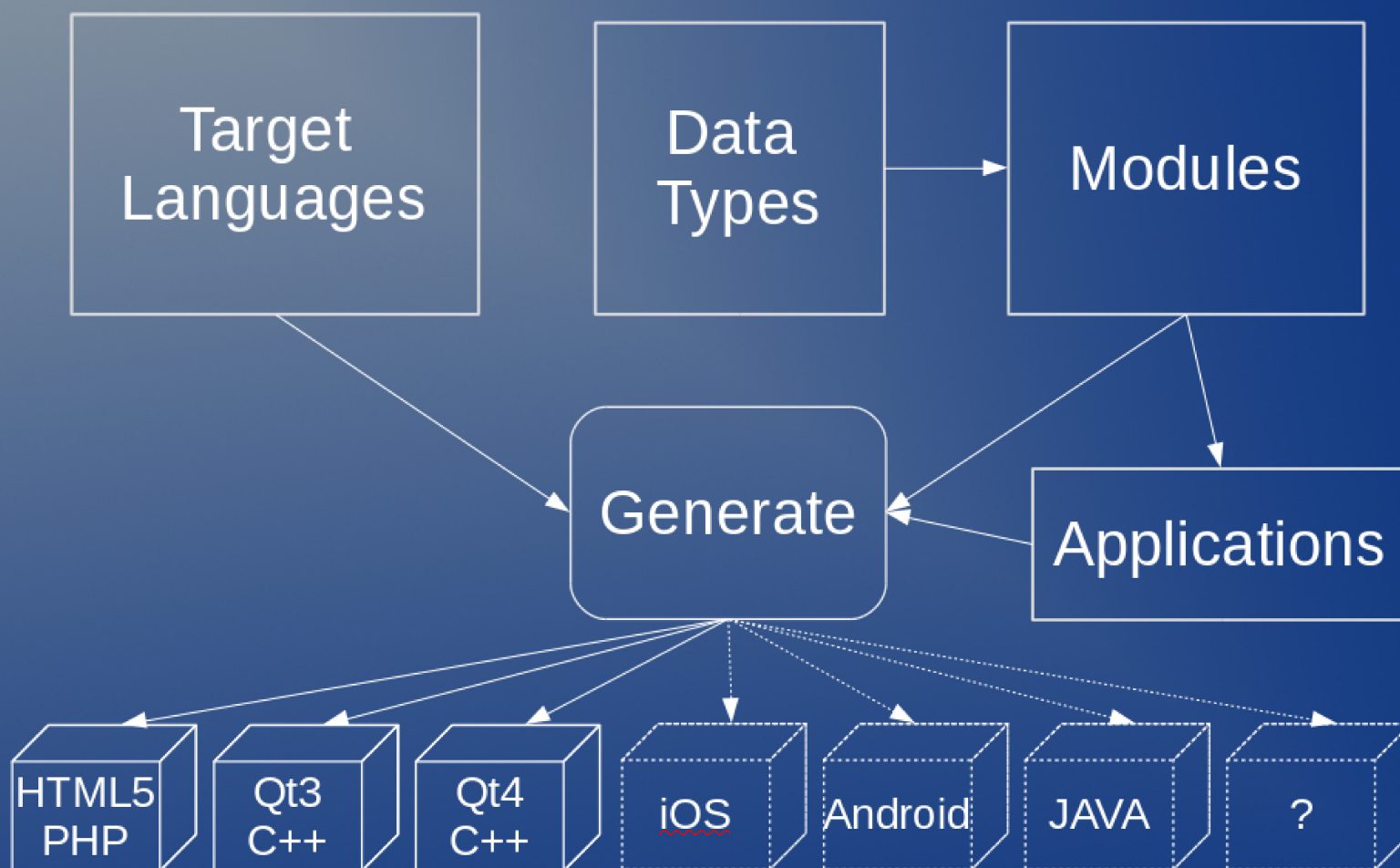
Considerations

- **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

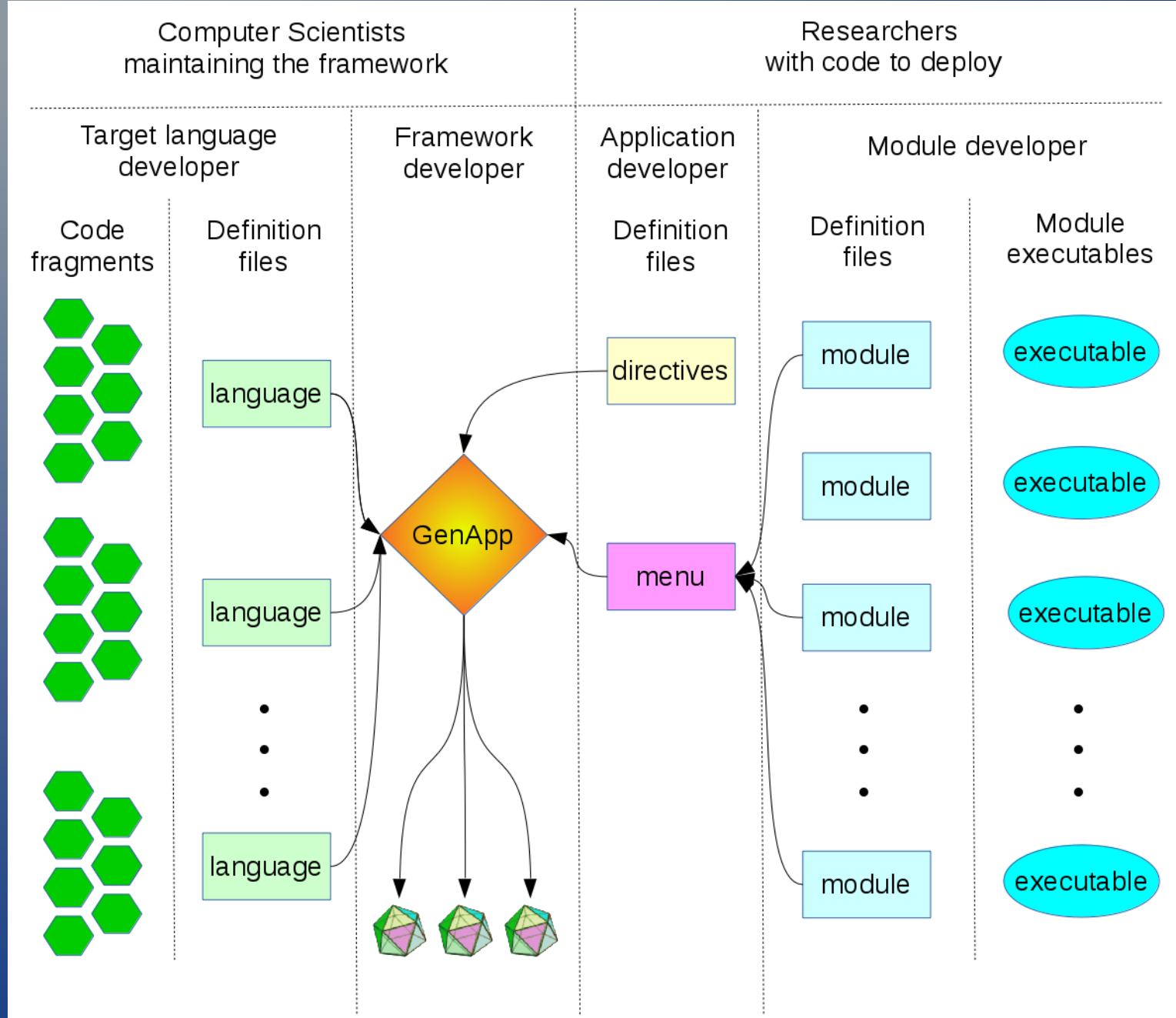
“Dark code”

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

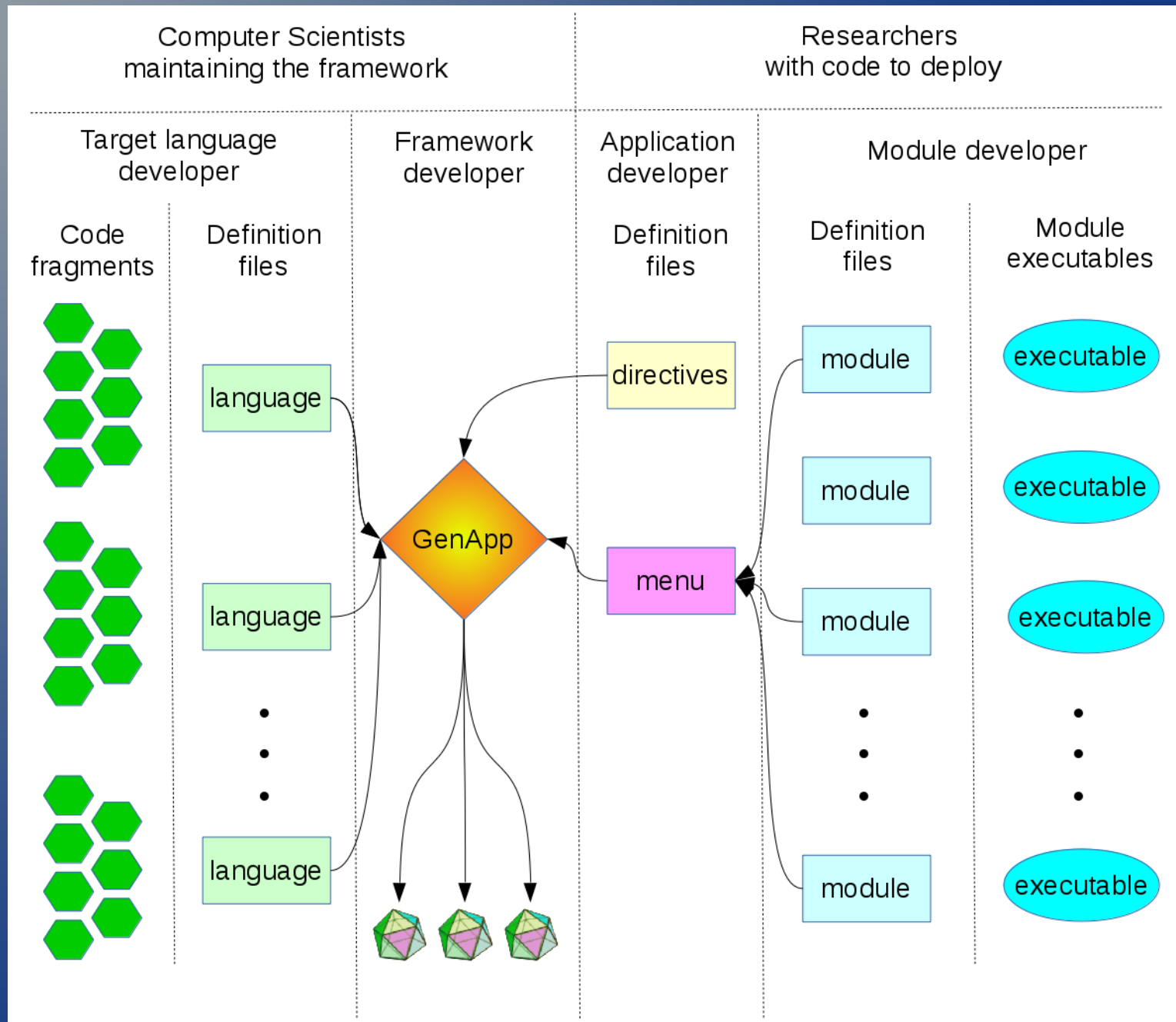
Brookes, E. H. XSEDE 14 Atlanta



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)
SCT (S. Perkins)

<https://sassie-web.chem.utk.edu/sassie2> *40 modules*
350+ users, 10,000+ jobs in 2016, 40+ papers

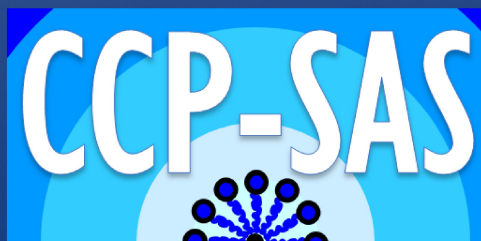
BayesApp (S. Hansen)
Denfert (J. Perez)
Vortex Shedding (A. Perlstein)

<http://genapp.rocks/bayesapp>
<http://genapp.rocks/denfert>
<http://genapp.rocks/vortexshedding>

Alpha

NAMDranner (A. Savelyev)

<http://genapp.rocks/namdranner>



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

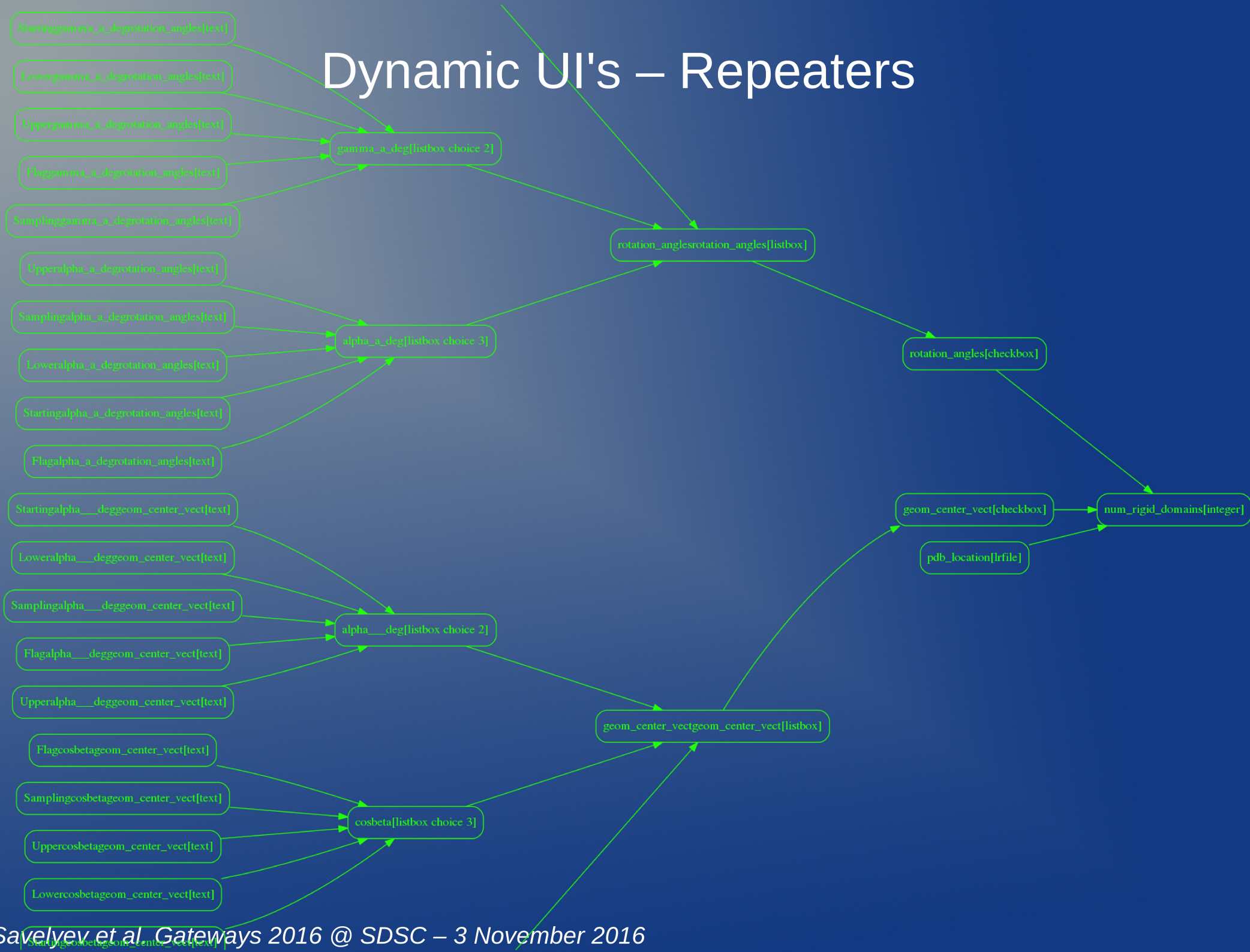
- How to best define interactivity in a single definition file that can be handled within the UI?

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:

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

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 )
```

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

PERL

```
use JSON;

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

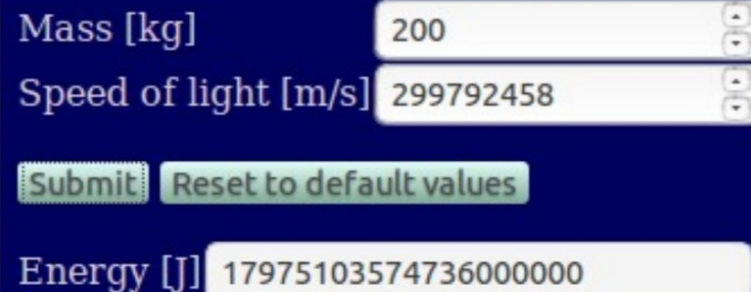
$json = shift;
$json_variables = decode_json( $json );

$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" ;
```



Mass [kg] 200

Speed of light [m/s] 299792458

Energy [J] 17975103574736000000

Input menu for 'QuaFit' module

QuaFit - Pre Alpha 0.1

Project name: 1
Code Number: 0004

OPTIMIZATION PARAMETERS

Number of Cycles per SubRun: 1
Maximum Number of SubRun: 2
Hypothetical Final H(X): 0.1
Maximum Rank L for [A]Im(Q): 7
Maximum Rank L for [A]SIm(Q): 7
Maximum Q of AIm(Q): 0.3
Stone Exp Contact Dist Flag: ☒

Stone Exp Contact Dist Family Name: Browse... qua_0101.sig or Browse server Local: qua_0101.sig

Optimisation free fraction: ☐
Exponent alpha of powerlaw error fu: 0.8
Factor H starting: 2
itoluns5: 1
qunpscale: 1
stnpuscale: 1

STRUCTURE OF STATES

Number of Independent States: 1

Number of MC States Configurations [1]: 1
HIERARCHIC AGGREGATE[1]
Number of Hierarchic Aggregates: [1]: 1

Point Group Symmetry [1]: 02
Point Group Reference System [1]: ☒
alpha: 0.0000 deg_beta: 0.0000 gamma: 0.0000
Screw Polymerization Number [1]: 1
Screw Axis Orientation [1]: ☒
cobeta: Starting: 1.0000 Lower: -1.0000 Upper: 1.0000 Sampling: 0.5000 Flag: 0
alpha_deg: Starting: 0.00000 Lower: 0.0000 Upper: 360.0000 Sampling: 20.0000 Flag: 0
Rotation along Screw Axis (deg) [1]: ☒
Starting: 0.95018 Lower: 0.0000 Upper: 180.0000 Sampling: 64.250 Flag: 0
Screw Translation along z (Ang) [1]: ☒
Starting: 141.133 Lower: 0.0000 Upper: 100.0000 Sampling: 40.131 Flag: 0
Relative mass density hydr shell [1]: ☒
Starting: 1.05 Lower: 1.0 Upper: 1.07 Sampling: 0.5 Flag: 0

RIGID DOMAINS[1]

Number of Rigid Domains: [1]: 1
Location of the RD PDB file [1]: Browse... 1PL7_A.pdb or Browse server Local: 1PL7_A.pdb
Geom Center Vect [1]: ☒
cobeta: Starting: -0.00943 Lower: -1.0000 Upper: 1.0000 Sampling: 0.50000 Flag: 0
modulus_ang: Starting: 20.003 Lower: 0.0000 Upper: 10.00 Sampling: 2.0000 Flag: 0
alpha_deg: Starting: -92.221 Lower: -180.00 Upper: 180.00 Sampling: 90.000 Flag: 0
Rotation Angles [1]: ☒
cobeta_s: Starting: 0.359340 Lower: -1.0000 Upper: 1.0000 Sampling: 0.50000 Flag: 0
alpha_s_deg: Starting: -134.80 Lower: -180.00 Upper: 180.00 Sampling: 90.000 Flag: 0
gamma_s_deg: Starting: 149.80 Lower: -180.00 Upper: 180.00 Sampling: 90.000 Flag: 0

FLEXIBLE LINKER[1]

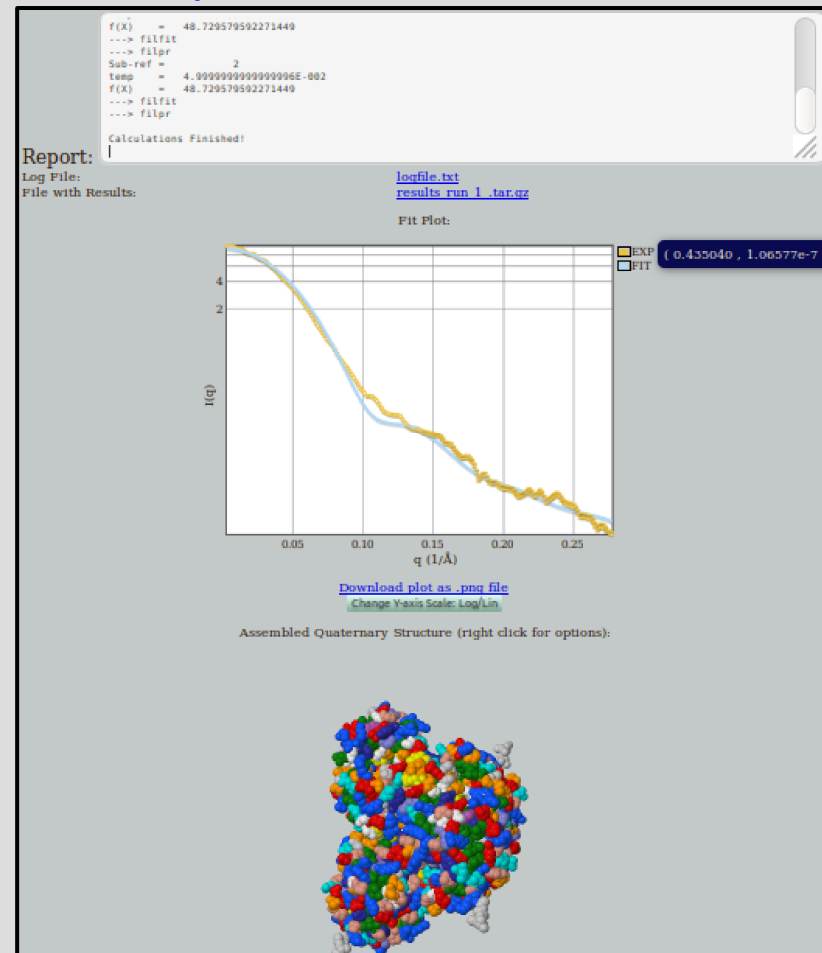
Sequence of RD and PLs in pdb? [1]: ☐
Number of Flexible Linkers: [1]: 0
AMINO ACID CONSTRAINTS[1]
Number of Amino Acid Constraints: [1]: 0
PAIR OF AMINO ACID CONSTRAINTS[1]
Number of Pairs of Amino Acid Constraints: [1]: 0

EXPERIMENTS

Number of Experiments: 1
Location of the SAS curve: [1]: Browse... curve2.dat or Browse server Local: curve2.dat
Concentration (g/dl) [1]: 3
XBays (10) / Neutrons (x_D2O) Flag [1]: -1
Scatt Dens Soly (1012 cm Angs3) [1]: 0.09369
Minimum Q in Data File (Angs1) [1]: 0
Maximum Q in Data File (Angs1) [1]: 0.28
Parameter K for Poissonian Errors [1]: 0.025
Number of Lines To Skip in Data File [1]: 0
Num of QPoints To Average [1]: 2
Maximum Distance for p(r) Calculation [1]: 90
Scaling Factor (if 0 is Optimized) [1]: 0
Flag Background [1]: ☒
Family Number (Same Scaling Factor) [1]: 1

[Submit](#) [Reset to default values](#)

Output of the 'QuaFit' module



QuaFit - Beta:

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

Encountering GenApp: Extending Horizon

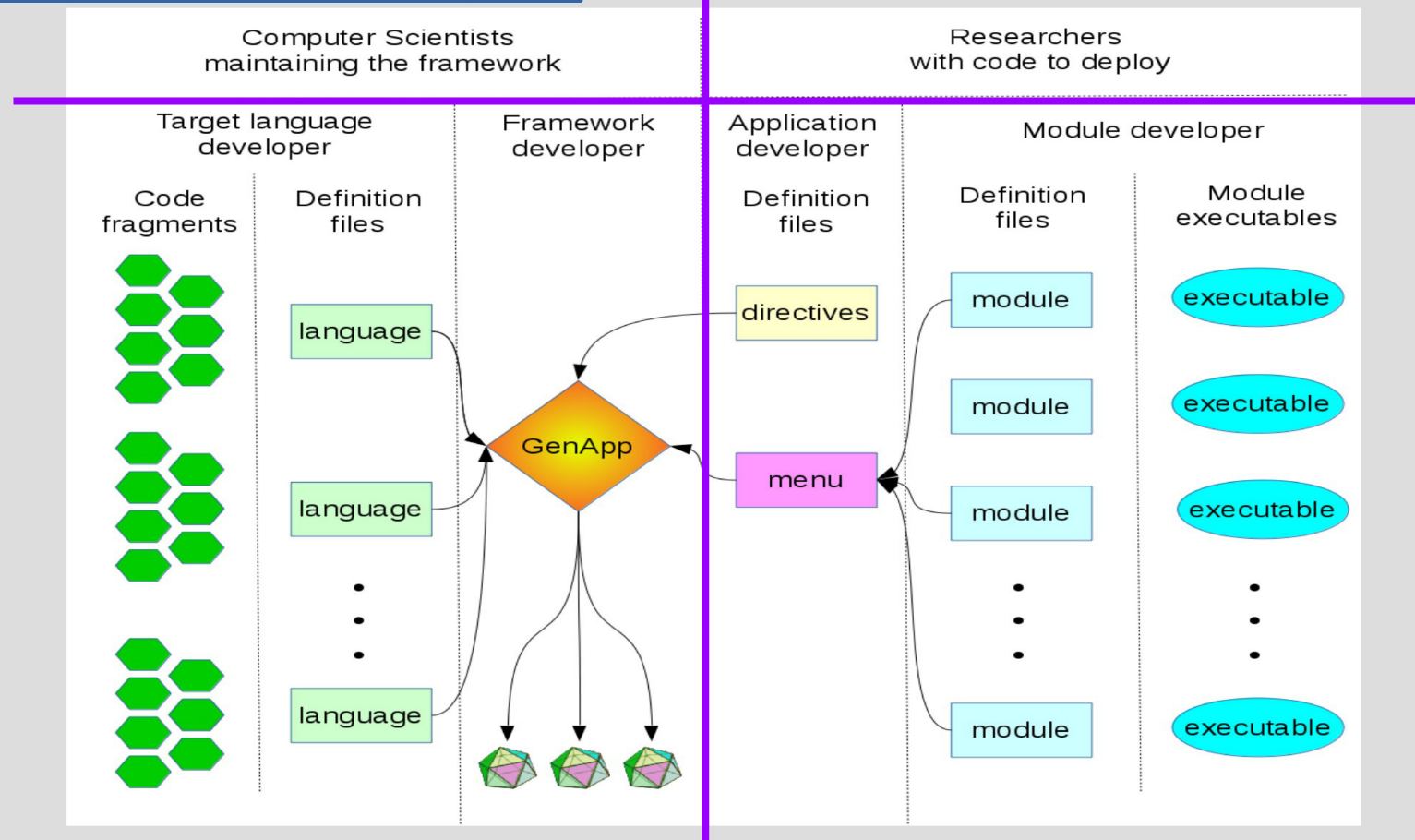
Cool stuff:

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



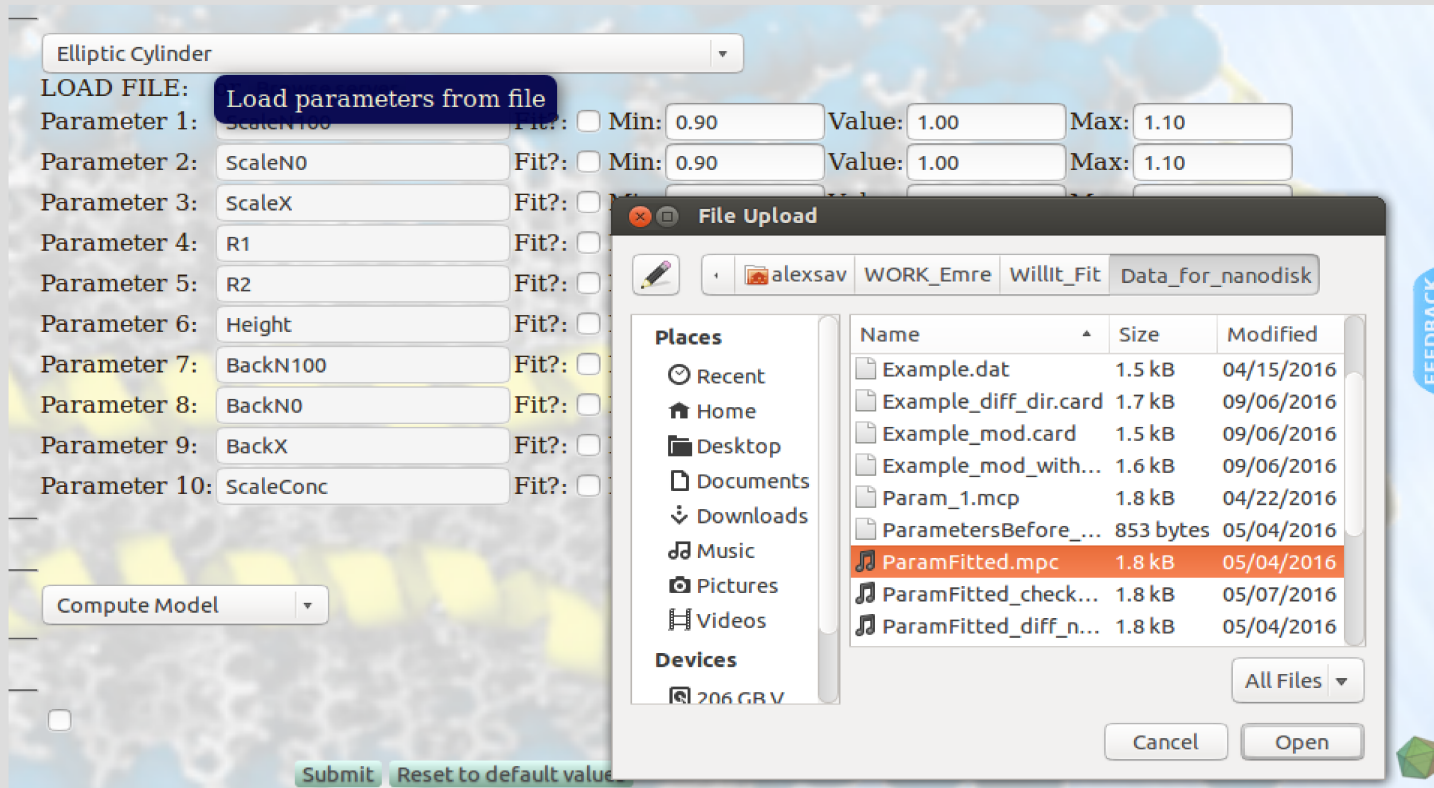
Web Development:

HTML5
CSS
JavaScript (jQuery)
PHP
MongoDB



Advanced “GenApping”: *WillItFit* (1)

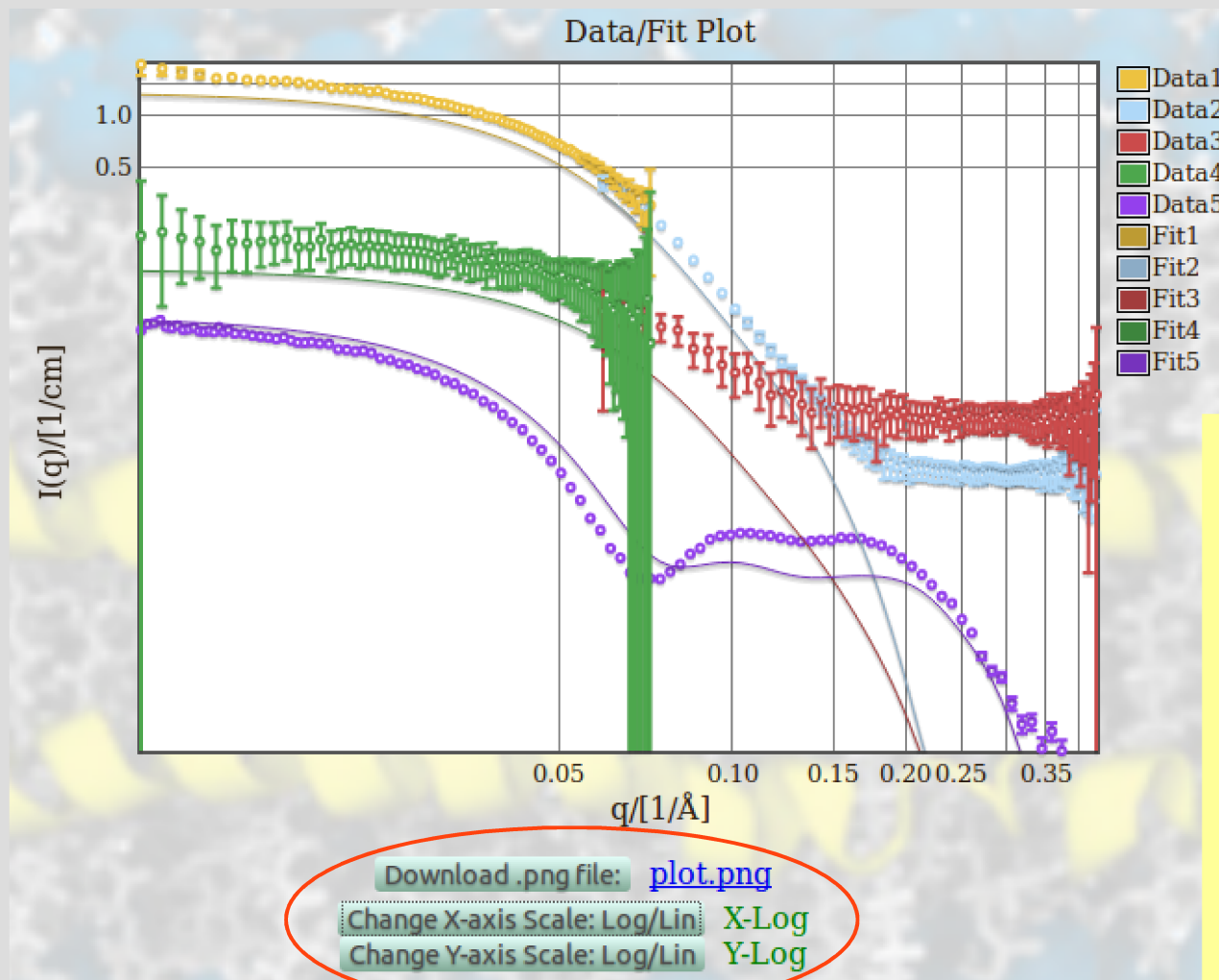
- Uploading parameter file, `lrfile.input` (numerical values, whitespace_separated etc.)



```
{
  "role" : "input",
  "id" : "param_choice_file_anis_core_shell_micelles",
  "label" : "LOAD FILE: ",
  "type" : "lrfile",
  "setinputfromfile" : "whitespaceparated_reversellogic",
  "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 dissolved 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
EYLNLLKKLIGEVSALETQANLNGIFIRTSYEKDSIYINGDQNQLKQVFINLIKNAVESMPDGGTVDIITEDEHSVHVTVKDE
GEGIPEKVLNRIGEPFLTTEKEKGTGLGLHPEKGTAFKISFPKK 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

Upload Contrast File

CONTRAST POINTS

Use Contrast Points? ☒

Number of Contrast Points:

```
{  
  "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,  
                           ..."  
}
```

$f_{D_2O}(0-1)$	$I(0)$	$\sigma[I(0)]$	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

DISSOLVED SPECIES

Number of Dissolved Species in the Solvent

Substance Type: [1]

Formula: [1] Concentration: [1] Volume: [1]

Substance Type: [2]

Advanced “GenApping”: MULCh (2)

- Calculated fields (inter-dependent numerical fields)

CONTRAST VALUES (opt)			rho_1	fd20	rho_2
Vol ₁ (Å)	<input type="text" value="65"/>		$\Delta\rho_1=$ <input type="text" value="155"/>	f_{D_2O+}	<input type="text" value="23"/>
Vol ₂ (Å)	<input type="text" value="40"/>		$\Delta\rho_2=$ <input type="text" value="435"/>	f_{D_2O+}	<input type="text" value="234"/>

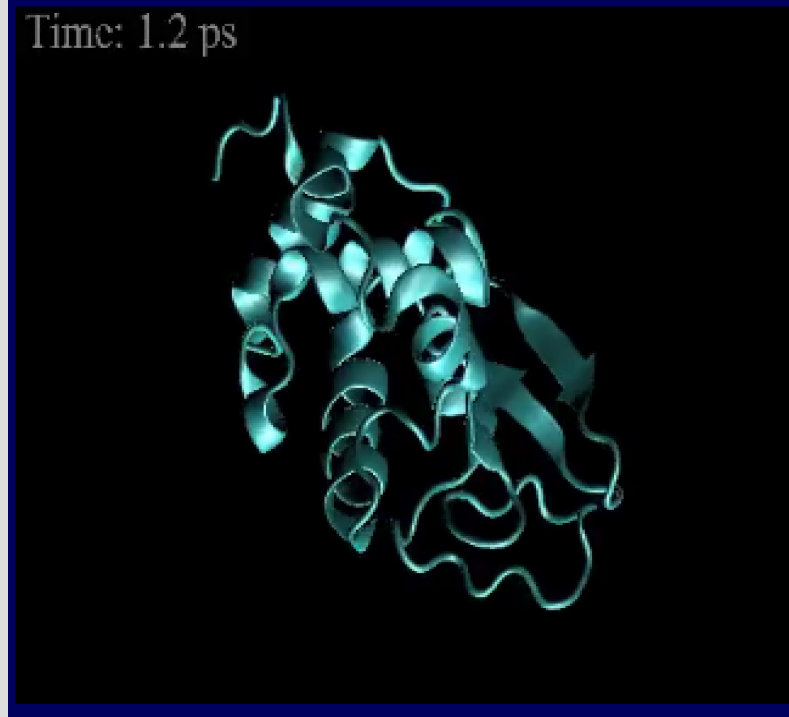
CONTRAST POINTS									
Upper qR_g limit	<input type="text" value="1.3"/>								
Number of Contrast Points:	<input type="text" value="4"/>								
	$f_{D_2O}(0-1)$	Refine	Scale	Value	Start Point	Contrast Data	$\Delta\rho_1$	$\Delta\rho_2$	Vol ₁ /(Vol ₁ +Vol ₂)
fd20 {	<input type="text" value="0.1"/>	<input checked="" type="checkbox"/>	<input type="text" value="1.0"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="button" value="Browse..."/> No file selected.	<input type="text" value="38.5"/>	<input type="text" value="277.5"/>	<input type="text" value="0.6190476190"/>
	<input type="text" value="0.2"/>	<input checked="" type="checkbox"/>	<input type="text" value="1.0"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="button" value="Browse..."/> No file selected.	<input type="text" value="54"/>	<input type="text" value="321"/>	<input type="text" value="0.6190476190"/>
	<input type="text" value="0.3"/>	<input checked="" type="checkbox"/>	<input type="text" value="1.0"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="button" value="Browse..."/> No file selected.	<input type="text" value="69.5"/>	<input type="text" value="364.5"/>	<input type="text" value="0.6190476190"/>
	<input type="text" value="0.4"/>	<input checked="" type="checkbox"/>	<input type="text" value="1.0"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="button" value="Browse..."/> No file selected.	<input type="text" value="85"/>	<input type="text" value="408"/>	<input type="text" value="0.6190476190"/>

```
{
  "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)

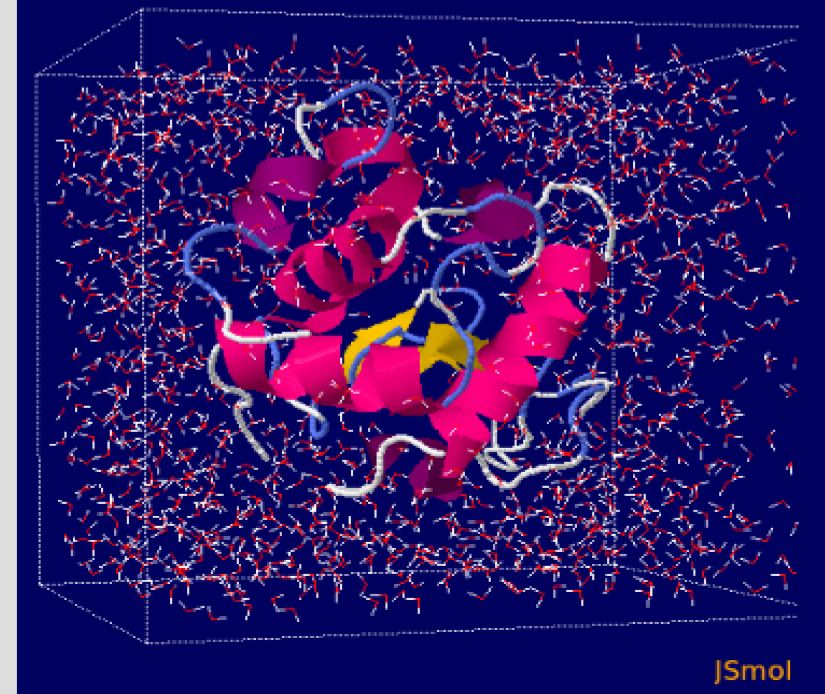
Trajectory Movie



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

- Structure visualization, JSmol
atomicstructure.output

Atomistic Model (right click for options):



```
{  
  "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)

The screenshot shows the NAMDrun - Beta 0.1 web interface. At the top, there is a navigation bar with a menu icon, a hard hat icon, the title "NAMDrun - Beta 0.1", and links for "Logoff alexey" and "Help on". Below the navigation bar, there are several tabs: "Job monitor", "Integrity check", "Users", "User management", and "Job history". The "Job history" tab is highlighted with a red circle. Below the tabs, there are two input fields for "Start Date" (2016-10-10) and "End Date" (2016-11-03). Below these fields are two buttons: "Submit" and "Reset to default values". Below the buttons is a table with the following columns: "name", "email", "duration (h)", "running", "finished", and "cancelled". The table contains data for various users, including "Totals", "alexey", "amirayuyue", "cpayne", "danielma", "emre", "graceb", "gumbart", "hwang", "jvermaas", "mocohen", "pcardena", and "thomas.joseph". A "FEEDBACK" button is visible on the right side of the table. The bottom left corner of the interface shows the text "genapp.rocks".

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
thomas.joseph	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***

