AstroLabNet

Status

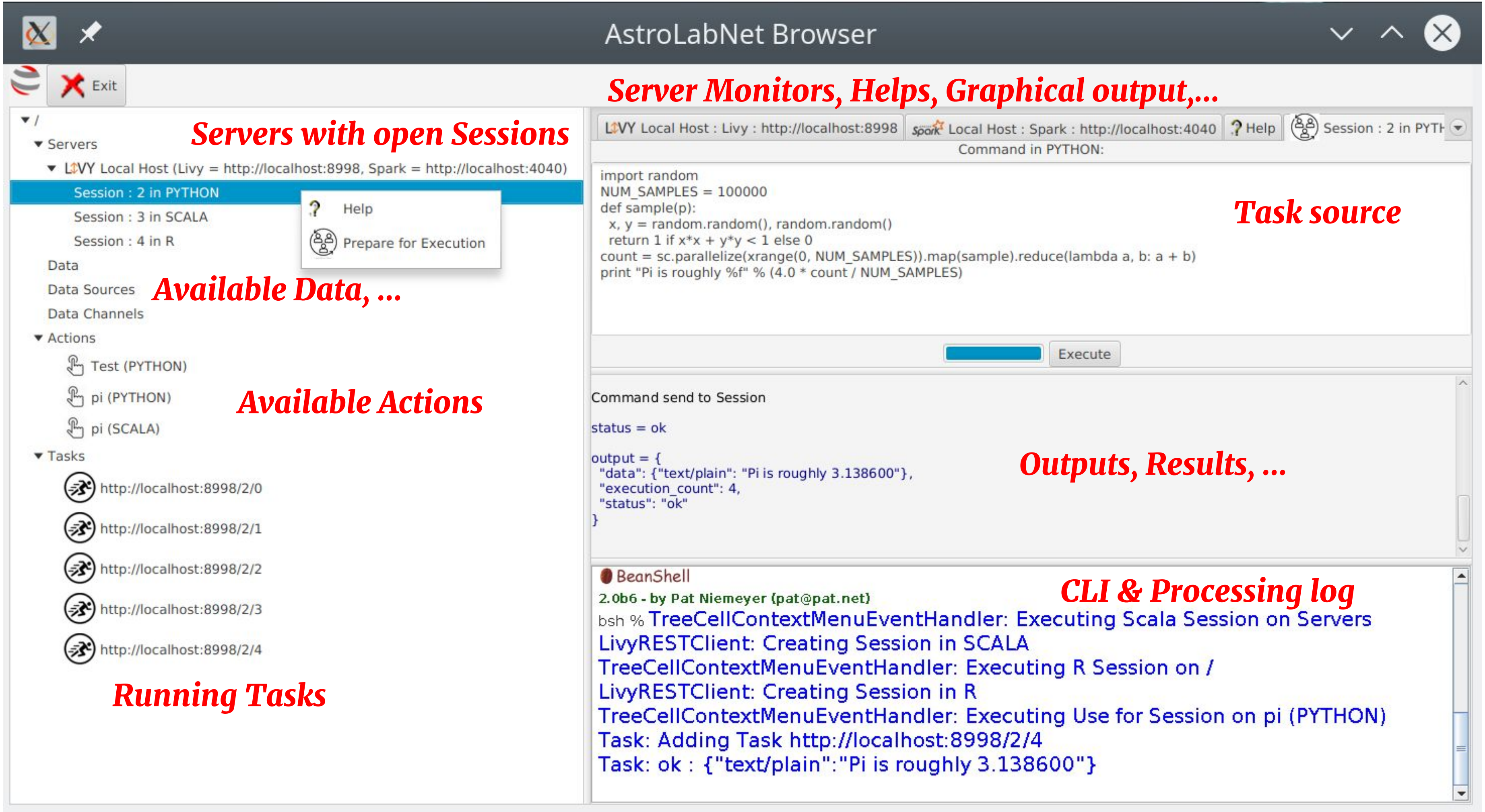
How to give everyone the AstroLab functionality !

AstroLabNet is a front-end to the archipel of *Spark* services.
It allows to

- Access distributed data.
- Deploy jobs to data.
- Move data between servers.
- Arrange data streaming and updating.

Full functionality will be available via CLI, GUI and WS.

Julius Hrivnac
19/2/19, LAL, Orsay

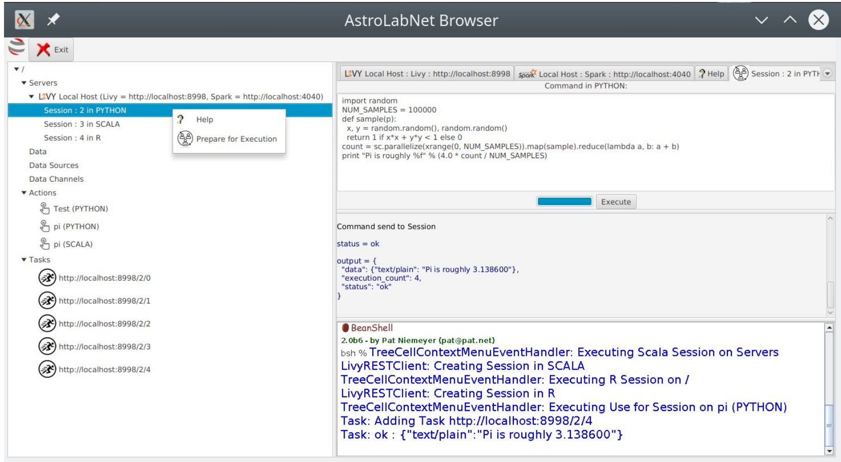

AstroLabNet Browser
Exit
Server Monitors, Helps, Graphical output,...
Servers with open Sessions
/
Servers
LIVY Local Host (Livy = http://localhost:8998, Spark = http://localhost:4040)
Session : 2 in PYTHON
Session : 3 in SCALA
Session : 4 in R
Help
Prepare for Execution
Data
Data Sources
Data Channels
Available Data, ...
Actions
Test (PYTHON)
pi (PYTHON)
pi (SCALA)
Available Actions
Tasks
http://localhost:8998/2/0
http://localhost:8998/2/1
http://localhost:8998/2/2
http://localhost:8998/2/3
http://localhost:8998/2/4
Running Tasks
LIVY Local Host : Livy : http://localhost:8998
Spark Local Host : Spark : http://localhost:4040
Help
Session : 2 in PYTH
Command in PYTHON:
Task source
Execute
Outputs, Results, ...
CLI & Processing log

```
import random
NUM_SAMPLES = 100000
def sample(p):
  x, y = random.random(), random.random()
  return 1 if x*x + y*y < 1 else 0
count = sc.parallelize(xrange(0, NUM_SAMPLES)).map(sample).reduce(lambda a, b: a + b)
print "Pi is roughly %f" % (4.0 * count / NUM_SAMPLES)
```

```
Command send to Session

status = ok

output = {
  "data": {"text/plain": "Pi is roughly 3.138600"},
  "execution_count": 4,
  "status": "ok"
}
```

```
BeanShell
2.0b6 - by Pat Niemeyer (pat@pat.net)
bsh % TreeCellContextMenuEventHandler: Executing Scala Session on Servers
LivyRESTClient: Creating Session in SCALA
TreeCellContextMenuEventHandler: Executing R Session on /
LivyRESTClient: Creating Session in R
TreeCellContextMenuEventHandler: Executing Use for Session on pi (PYTHON)
Task: Adding Task http://localhost:8998/2/4
Task: ok : {"text/plain":"Pi is roughly 3.138600"}
```

Interface


AstroLabNet Browser
Exit
/
Servers
LIVY Local Host (Livy = http://localhost:8998, Spark = http://localhost:4040)
Session : 2 in PYTHON
Session : 3 in SCALA
Session : 4 in R
Help
Prepare for Execution
Data
Data Sources
Data Channels
Actions
Test (PYTHON)
pi (PYTHON)
pi (SCALA)
Tasks
http://localhost:8998/2/0
http://localhost:8998/2/1
http://localhost:8998/2/2
http://localhost:8998/2/3
http://localhost:8998/2/4
LIVY Local Host : Livy : http://localhost:8998
Spark Local Host : Spark : http://localhost:4040
Help
Session : 2 in PYTH
Command in PYTHON:
Execute

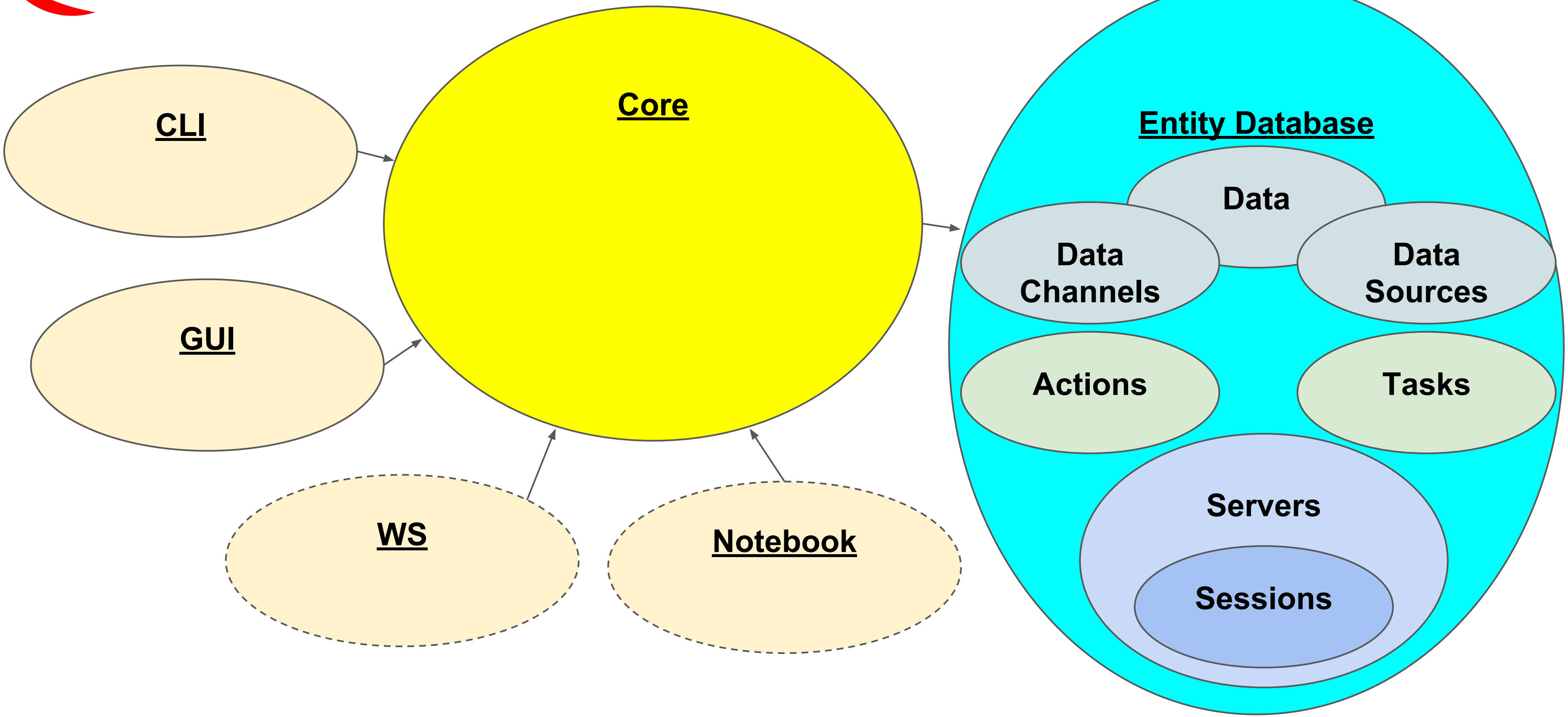
```
import random
NUM_SAMPLES = 100000
def sample(p):
  x, y = random.random(), random.random()
  return 1 if x*x + y*y < 1 else 0
count = sc.parallelize(xrange(0, NUM_SAMPLES)).map(sample).reduce(lambda a, b: a + b)
print "Pi is roughly %f" % (4.0 * count / NUM_SAMPLES)
```

```
Command send to Session

status = ok

output = {
 "data": {"text/plain": "Pi is roughly 3.138600"},
 "execution_count": 4,
 "status": "ok"
}
```

```
BeanShell
2.0b6 - by Pat Niemeyer (pat@pat.net)
bsh % TreeCellContextMenuEventHandler: Executing Scala Session on Servers
LivyRESTClient: Creating Session in SCALA
TreeCellContextMenuEventHandler: Executing R Session on /
LivyRESTClient: Creating Session in R
TreeCellContextMenuEventHandler: Executing Use for Session on pi (PYTHON)
Task: Adding Task http://localhost:8998/2/4
Task: ok : {"text/plain":"Pi is roughly 3.138600"}
```

- **Servers:** All available Livy/Spark servers
  - Servers will have list of other servers so all will be directly reachable
- **Sessions:** Open sessions on all Servers
- **Data:** Existing data
  - Needs a standard Catalog (?)
- **Data Source:** Sources of Data
  - Using Brokers, Alerts
- **Data Channels:** Active connections between Data Sources
- **Actions:** Available actions to be send to Sessions
  - Scala, Python, R, SQL
  - Scripts or JARs
- **Tasks:** Running or finished Actions

Architecture



Status - ToDos

- Functional Framework
  - Graphical Browser
  - CLI (with scripting possibilities) - In Development
  - Web Service (Tomcat) - ToDo
  - Notebook - ToDo
-
  - Can connect to (several) Spark servers via Livy proxies
  - Can send jobs, monitor progress and receive results
    - Even after re-start or from different site
- To Do now:
  - Clean up rough edges & dark corners
  - Connect to data Catalogs and Alerts
  - Send Jobs to data
  - Replicate data
- To Do next:
  - Run as a WS
  - Run in a Notebook
  - Handle Authentication

Command Line

In Development

```
usage: java -jar AstroLabNet.exe.jar
  -b,--browser         open graphical browser (default)
  -c,--cli             start command line
  -h,--help            show help
  -s,--source <file>   source bsh file (init.bsh is also read)
```

```
w.addServer("Local Host", "http://localhost:8998", "http://localhost:4040");
w.addServer("LAL", "http://vm-75222.lal.in2p3.fr:8998", "http://vm-75222.lal.in2p3.fr:20001");
w.servers();
for (Server server : w.servers()) {
  server.sessions();
  }
...
```

Can be extended by

- New BSH Scripts
- New Actions: A Spark job
  - Scala, Python, R, SQL
  - As simple scripts or full jars
- New plugins
  - Interactive graphical view to visualise and manage network of Servers with data flow
  - Content visualisation

Where

- Home:
  - https://hrivnac.web.cern.ch/hrivnac/Activities/Packages/AstroLabNet
- Source:
  - https://github.com/hrivnac/AstroLabNet.git