Contributing to the Lapps Grid

Lapps Service Wrapping

Lapps Grid Group

May 26, 2014

Outline

- Introduction
- From Software to Web Service
- From NLP Tool to Lapps Service
  - Java Example
  - Python Example
- Conclusion
- Reference

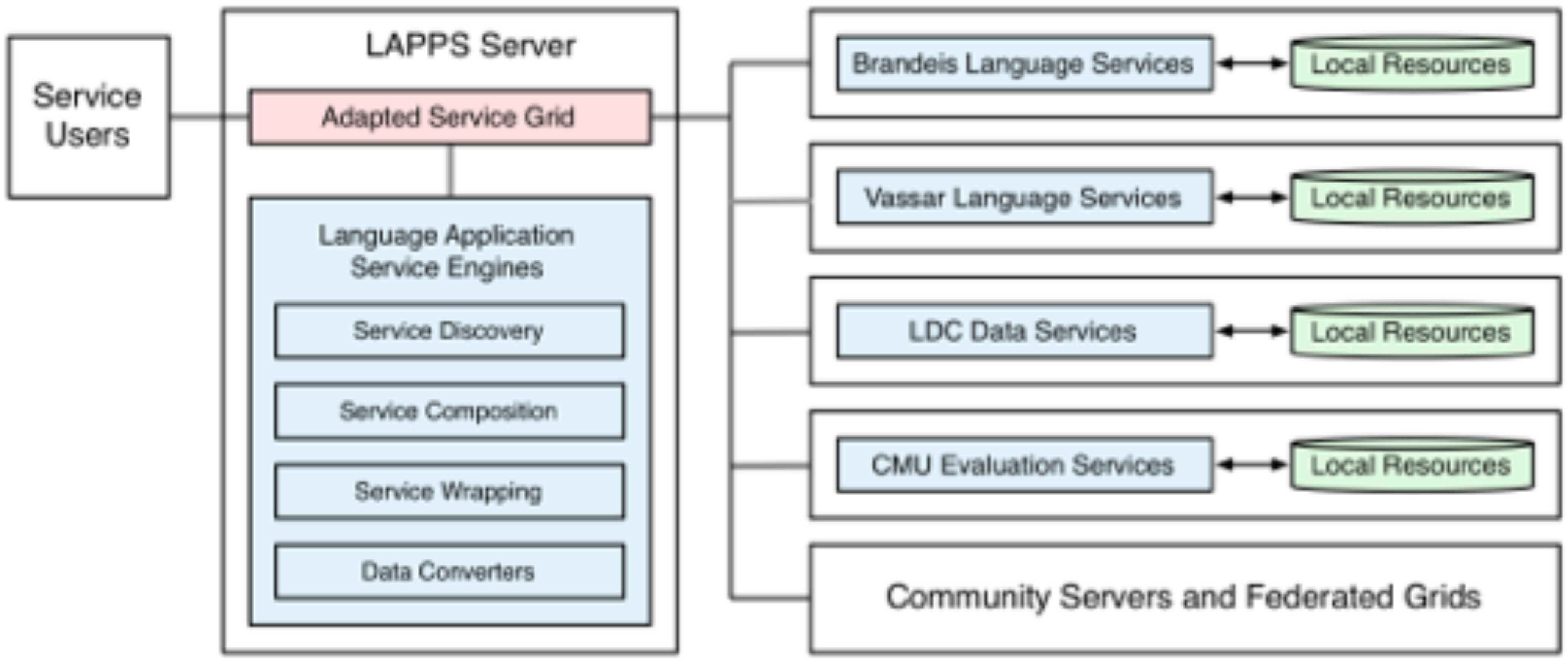
The Language Application Grid

- Availability & Interoperability of NLP Tools
    - Java, Python, tools
    - OpenNLP, Stanford NLP, Gate, NLTK
- Language Application (Lapps) Grid Project
    - Language Service
    - Lapps API Design

Lapps Grid Architecture

Using Composite Lapps Services

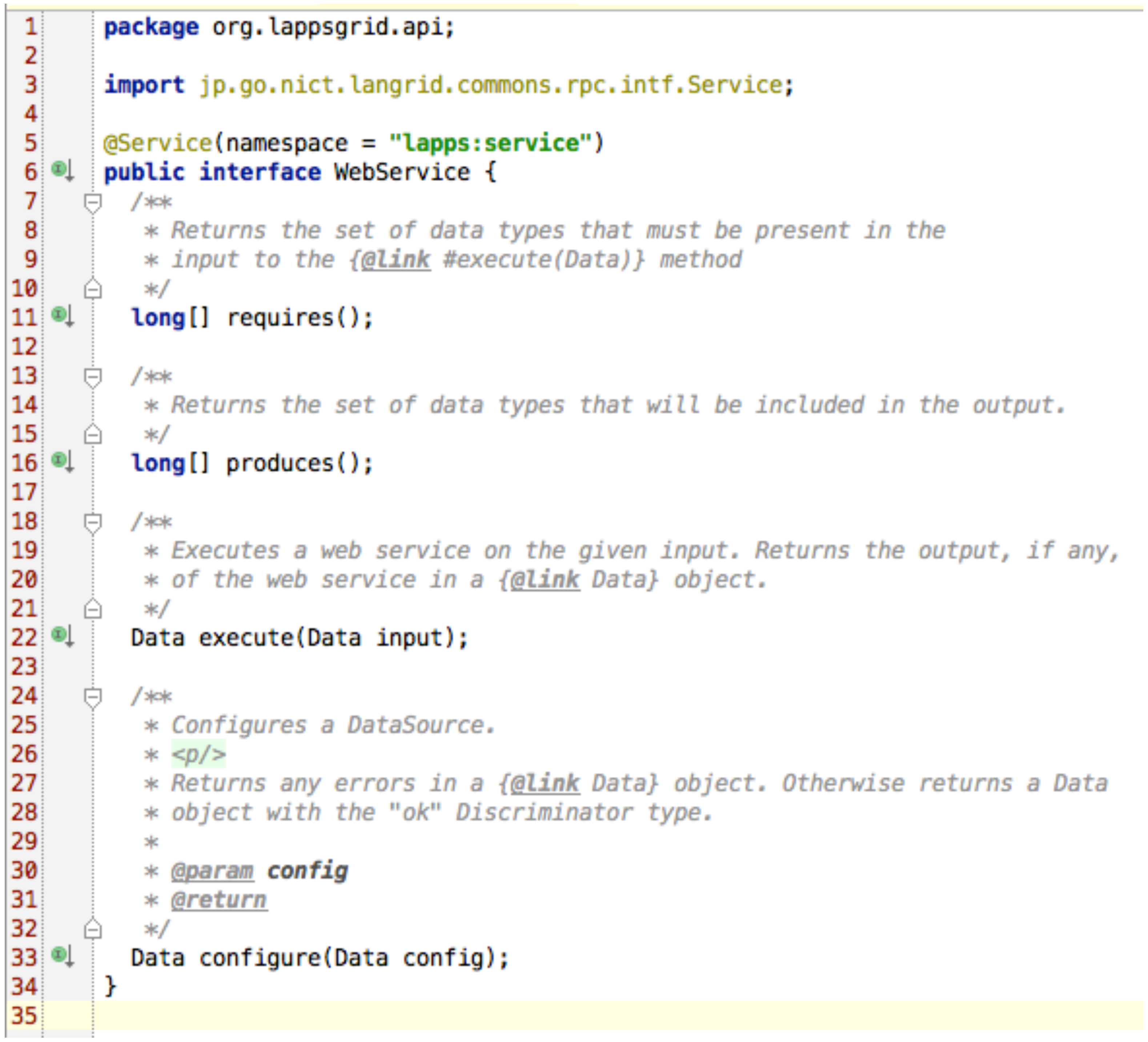
Wrapping Atomic Lapps Services



Lapps API Design

- Consistent Interface
- Discriminator
- JSON Format

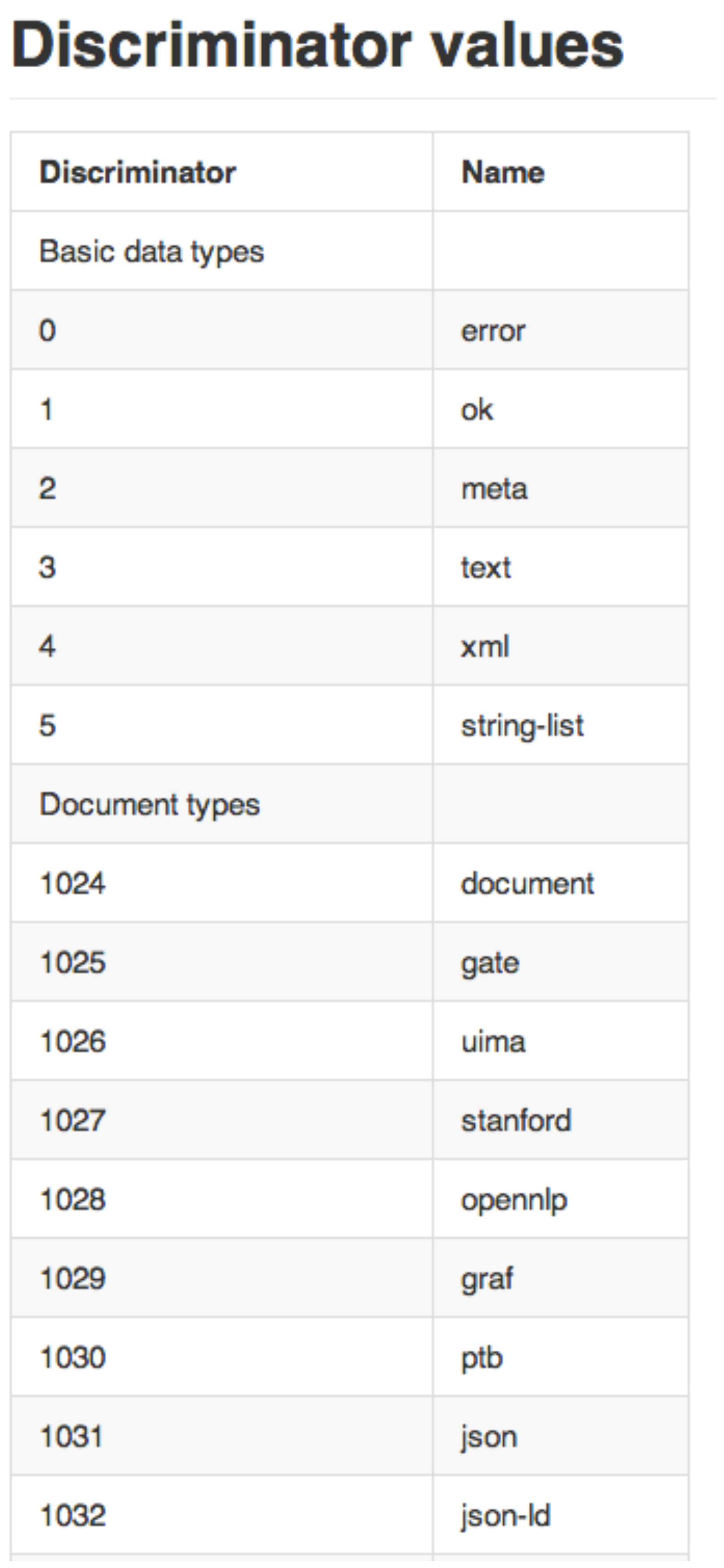
Consistent Interface (Java)



```java
package org.lappsgrid.api;

import jp.go.nict.langrid.commons.rpc.intf.Service;

@Service(namespace = "lapps:service")
public interface WebService {
    /**
     * Returns the set of data types that must be present in the
     * input to the {@link #execute(Data)} method
     */
    long[] requires();

    /**
     * Returns the set of data types that will be included in the output.
     */
    long[] produces();

    /**
     * Executes a web service on the given input. Returns the output, if any,
     * of the web service in a {@link Data} object.
     */
    Data execute(Data input);

    /**
     * Configures a DataSource.
     * <p/>
     * Returns any errors in a {@link Data} object. Otherwise returns a Data
     * object with the "ok" Discriminator type.
     *
     * @param config
     * @return
     */
    Data configure(Data config);
}
```

Discriminator

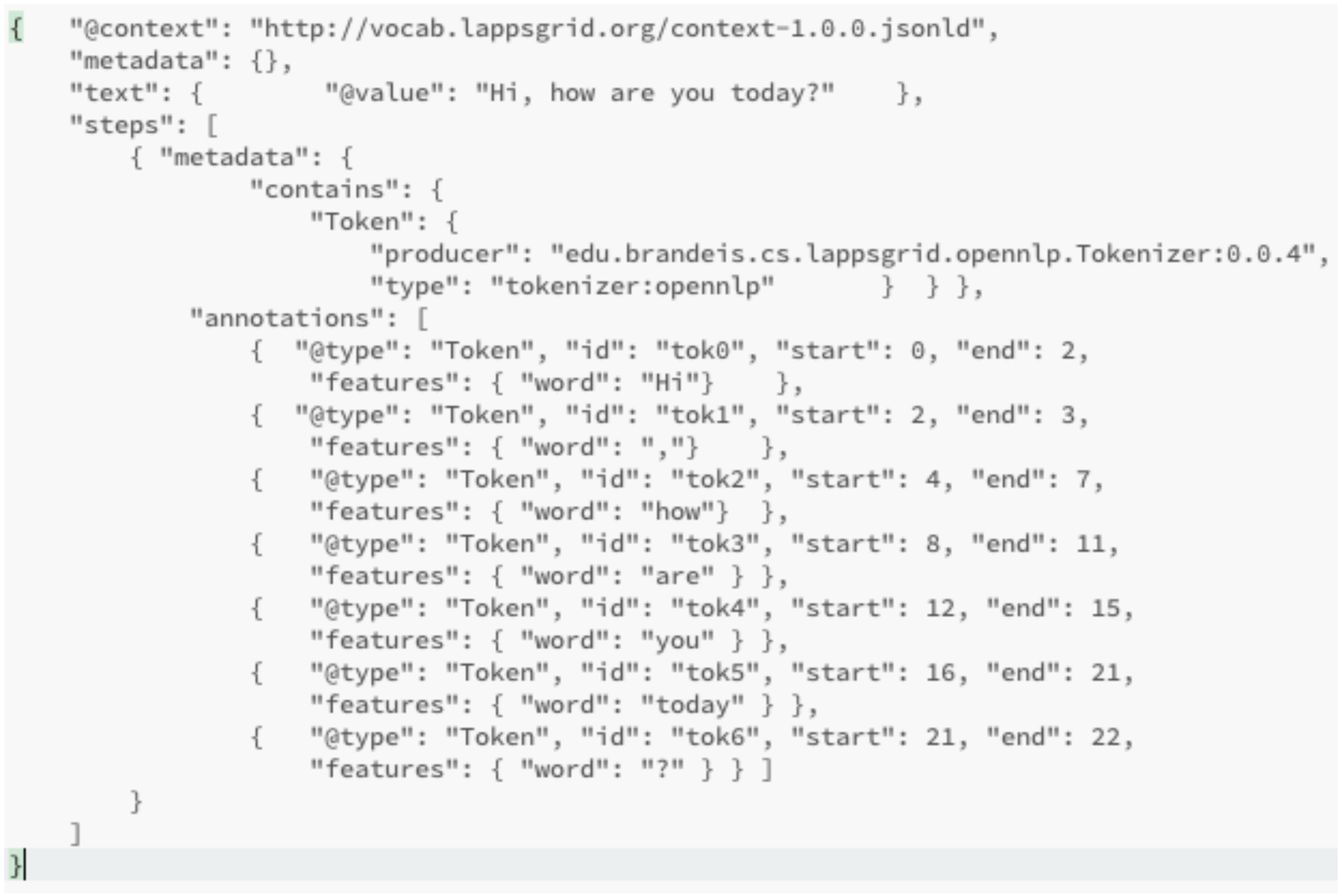
```
long [] requires()
long [] produces()
```



Discriminator values

Discriminator	Name
Basic data types	
0	error
1	ok
2	meta
3	text
4	xml
5	string-list
Document types	
1024	document
1025	gate
1026	uima
1027	stanford
1028	opennlp
1029	graf
1030	ptb
1031	json
1032	json-ld

JSON



```
{   "@context": "http://vocab.lappsgrid.org/context-1.0.0.jsonld",
    "metadata": {},
    "text": {        "@value": "Hi, how are you today?"    },
    "steps": [
        { "metadata": {
                "contains": {
                    "Token": {
                        "producer": "edu.brandeis.cs.lappsgrid.opennlp.Tokenizer:0.0.4",
                        "type": "tokenizer:opennlp"       }  } },
            "annotations": [
                {  "@type": "Token", "id": "tok0", "start": 0, "end": 2,
                    "features": { "word": "Hi"}    },
                {  "@type": "Token", "id": "tok1", "start": 2, "end": 3,
                    "features": { "word": ","}    },
                {   "@type": "Token", "id": "tok2", "start": 4, "end": 7,
                    "features": { "word": "how"}  },
                {   "@type": "Token", "id": "tok3", "start": 8, "end": 11,
                    "features": { "word": "are" } },
                {   "@type": "Token", "id": "tok4", "start": 12, "end": 15,
                    "features": { "word": "you" } },
                {   "@type": "Token", "id": "tok5", "start": 16, "end": 21,
                    "features": { "word": "today" } },
                {   "@type": "Token", "id": "tok6", "start": 21, "end": 22,
                    "features": { "word": "?" } } ]
        }
    ]
}
```

Contributing to Lapps Grid

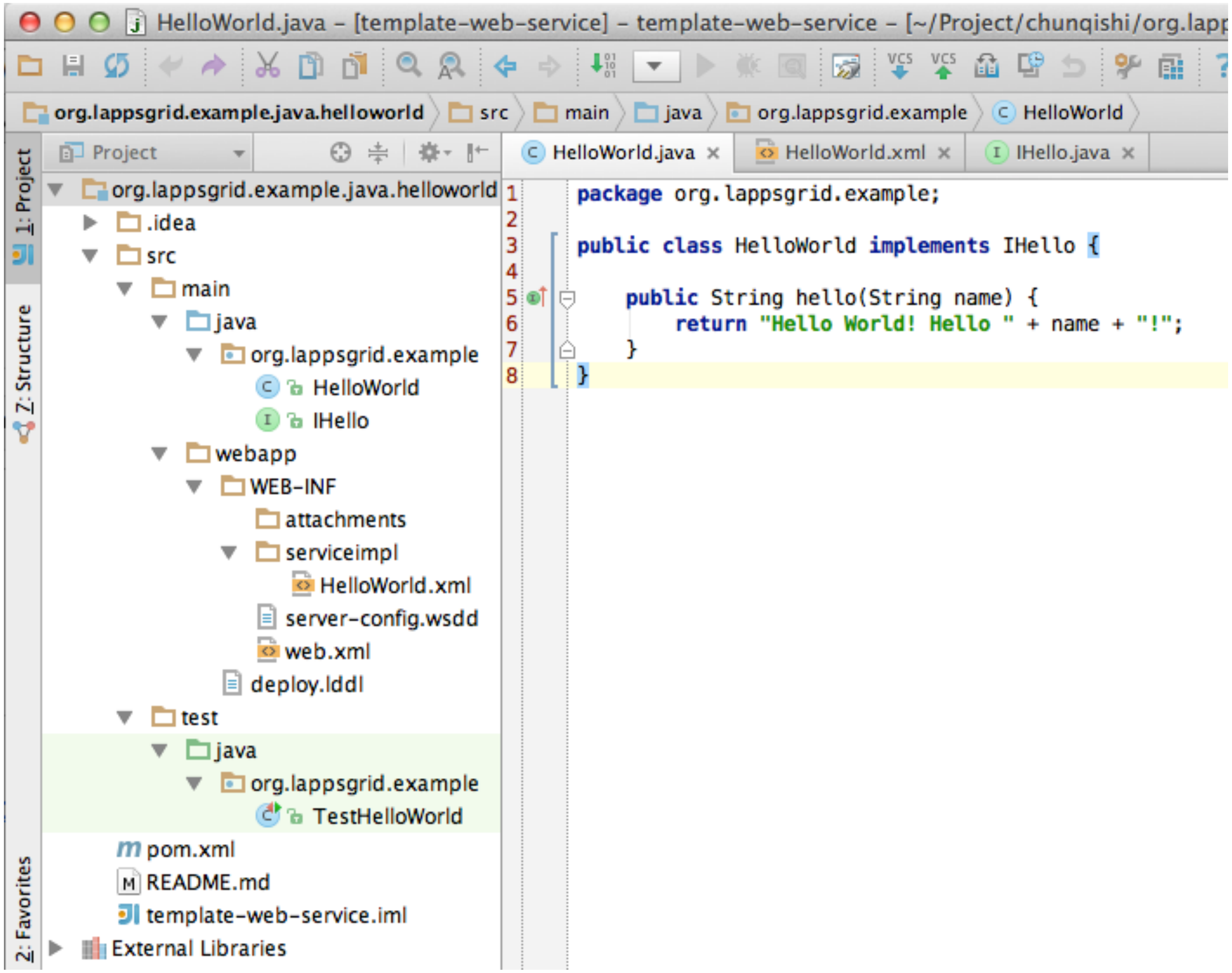
- Wrapping Lapps Service
  - NLP tools + Lapps API to atomic Lapps service
- Registering to Service Manager
  - Atomic Lapps services become available for searching and compositing

Service Wrapping Tutorial

- Web Service: “Hello World!”
  - “Hello World” Program (Java) —> WSDL
- Lapps Service: “Stanford Tagger”
  - Stanford Tagger (Java) + Lapps API —> WSDL
- Lapps Service: “NLTK Tagger”
  - NLTK Tagger (Python) + Lapps API —>WSDL

Web Service Wrapping

Hello World (Java)

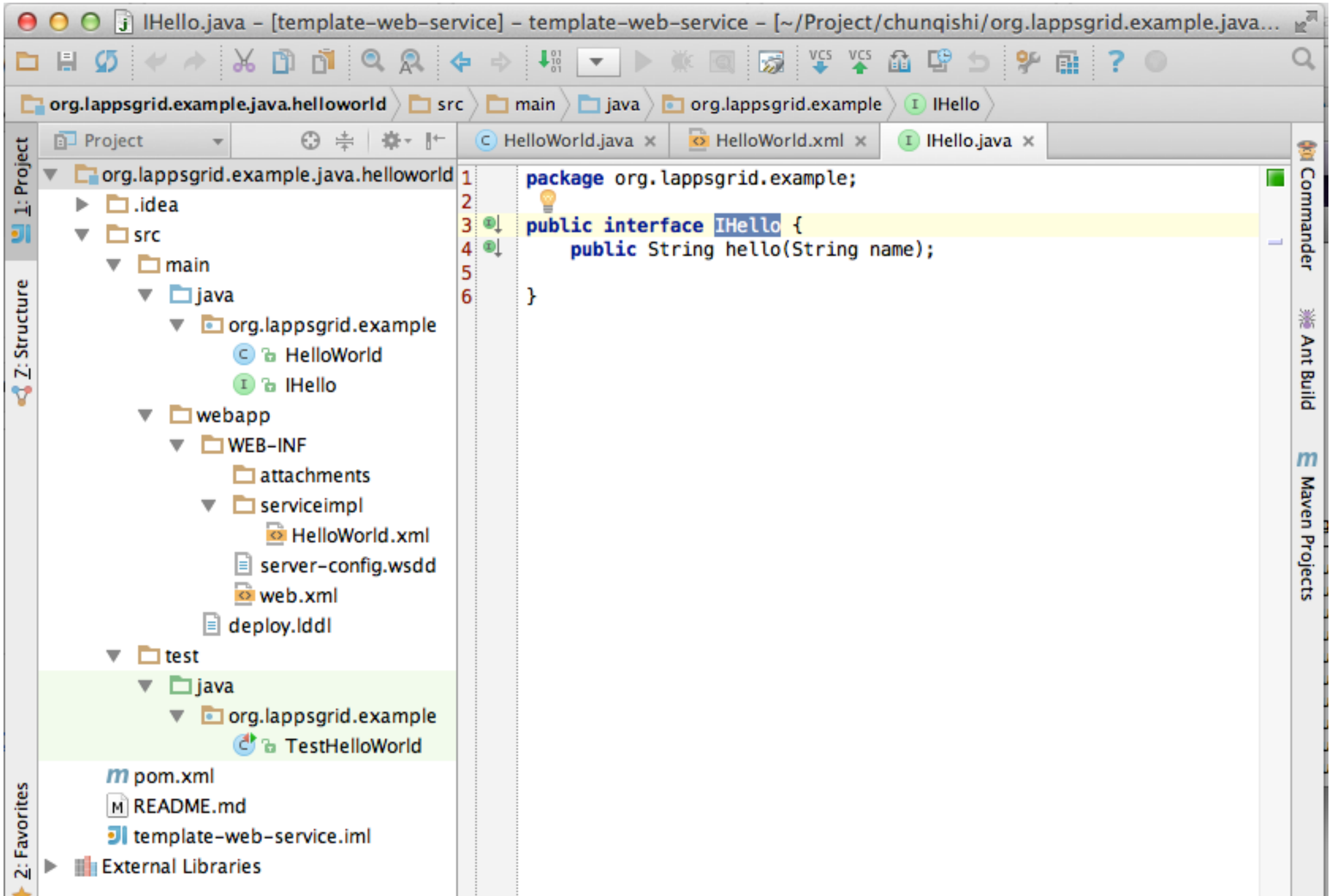

HelloWorld.java - [template-web-service] - template-web-service - [~/Project/chunqishi/org.lapp
org.lappsgrid.example.java.helloworld
src
main
java
org.lappsgrid.example
HelloWorld
Project
HelloWorld.java
HelloWorld.xml
IHello.java
1: Project
7: Structure
2: Favorites

```
org.lappsgrid.example.java.helloworld
    .idea
    src
        main
            java
                org.lappsgrid.example
                    HelloWorld
                    IHello
            webapp
                WEB-INF
                    attachments
                    serviceimpl
                        HelloWorld.xml
                    server-config.wsdd
                    web.xml
                deploy.lddl
        test
            java
                org.lappsgrid.example
                    TestHelloWorld
    pom.xml
    README.md
    template-web-service.iml
External Libraries
```

```java
package org.lappsgrid.example;

public class HelloWorld implements IHello {

    public String hello(String name) {
        return "Hello World! Hello " + name + "!";
    }
}
```

https://github.com/chunqishi/org.lappsgrid.example.java.helloworld

Interface Design



IHello.java - [template-web-service] - template-web-service - [~/Project/chunqishi/org.lappsgrid.example.java...

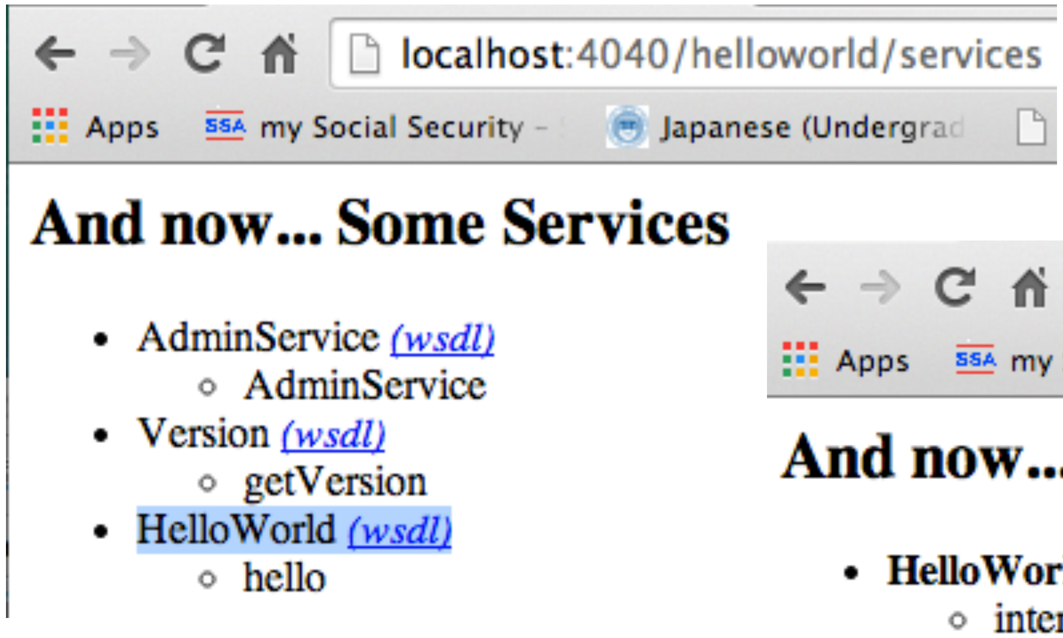
```
org.lappsgrid.example.java.helloworld
    .idea
    src
        main
            java
                org.lappsgrid.example
                    HelloWorld
                    IHello
            webapp
                WEB-INF
                    attachments
                    serviceimpl
                        HelloWorld.xml
                    server-config.wsdd
                    web.xml
                deploy.lddl
        test
            java
                org.lappsgrid.example
                    TestHelloWorld
    pom.xml
    README.md
    template-web-service.iml
External Libraries
```

```java
package org.lappsgrid.example;

public interface IHello {
    public String hello(String name);

}
```

Developing Template

- Developing Template
    - Maven for Dependency Library Management
    - Github Repository
        - https://github.com/chunqishi/org.lappsgrid.example.java.helloworld
- Local Test
    - Maven Compile/Package & Jetty Server based Testing
    - Command: *mvn clean package jetty:run*

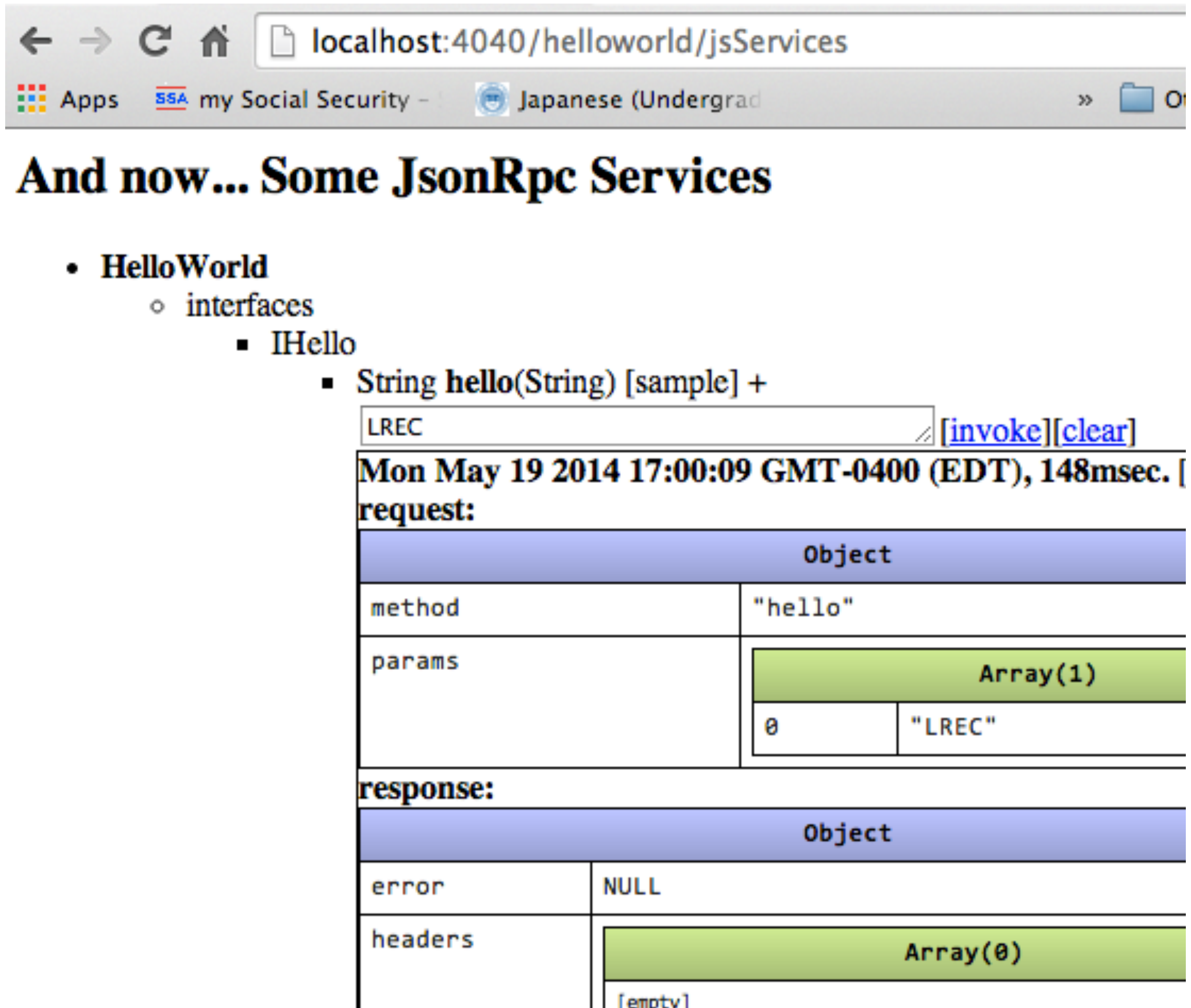
Web Service WSDL


localhost:4040/helloworld/services

Apps | my Social Security – | Japanese (Undergrad

And now... Some Services

- AdminService *(wsdl)*
  - AdminService
- Version *(wsdl)*
  - getVersion
- HelloWorld *(wsdl)*
  - hello


localhost:4040/helloworld/jsServices

Apps | my Social Security – | Japanese (Undergrad | » | O

And now... Some JsonRpc Services

- **HelloWorld**
  - interfaces
    - IHello
      - String **hello**(String) [sample] +

LREC [invoke][clear]

Mon May 19 2014 17:00:09 GMT-0400 (EDT), 148msec. [

request:

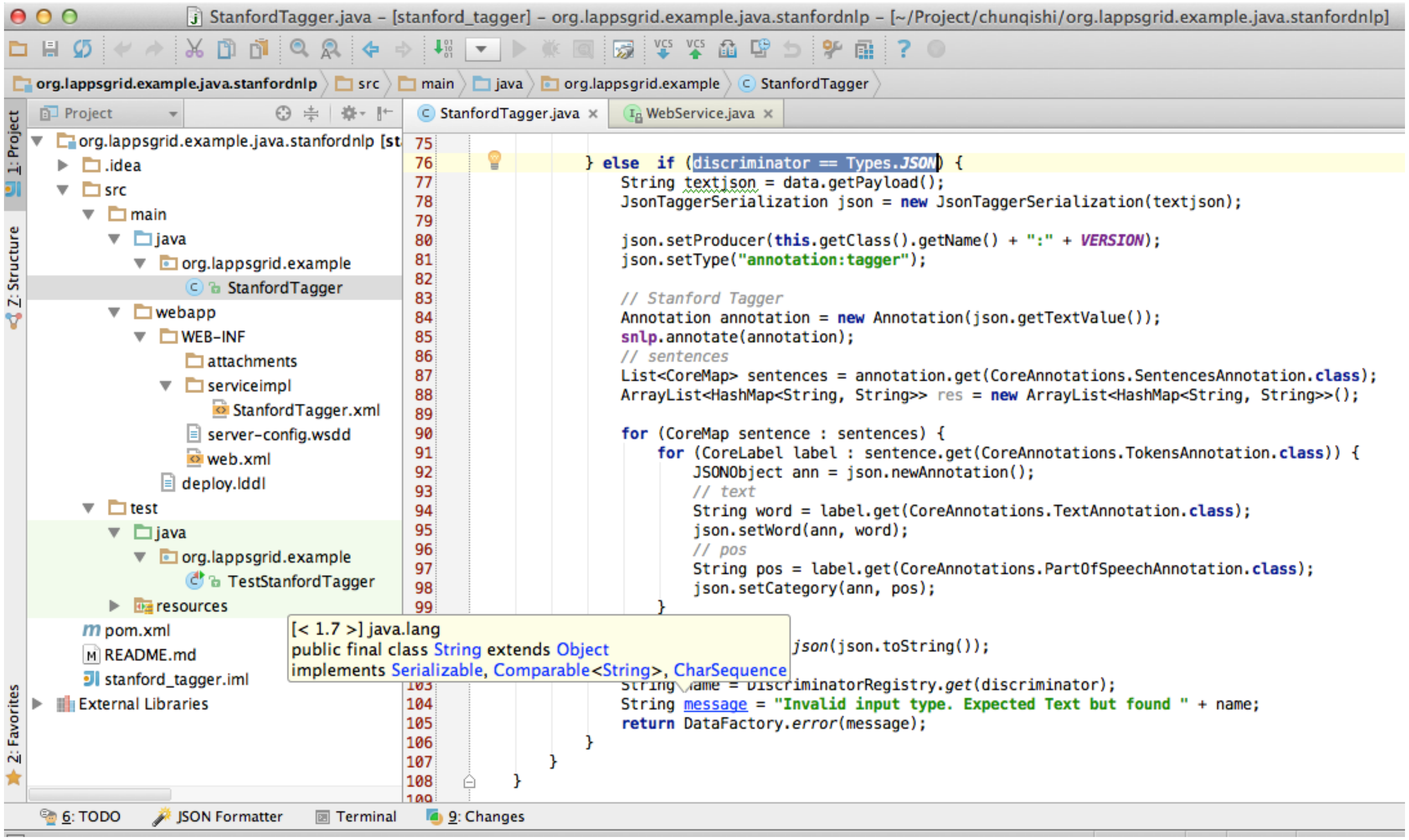
```
Object
method    "hello"
params    Array(1)
          0    "LREC"
```

response:

```
Object
error     NULL
headers   Array(0)
          [empty]
```

Lapps Service Wrapping (Java)

Developing Template


StanfordTagger.java – [stanford_tagger] – org.lappsgrid.example.java.stanfordnlp – [~/Project/chunqishi/org.lappsgrid.example.java.stanfordnlp]

```
org.lappsgrid.example.java.stanfordnlp [st
    .idea
    src
        main
            java
                org.lappsgrid.example
                    StanfordTagger
            webapp
                WEB-INF
                    attachments
                    serviceimpl
                        StanfordTagger.xml
                    server-config.wsdd
                    web.xml
                deploy.lddl
        test
            java
                org.lappsgrid.example
                    TestStanfordTagger
            resources
    pom.xml
    README.md
    stanford_tagger.iml
External Libraries
```

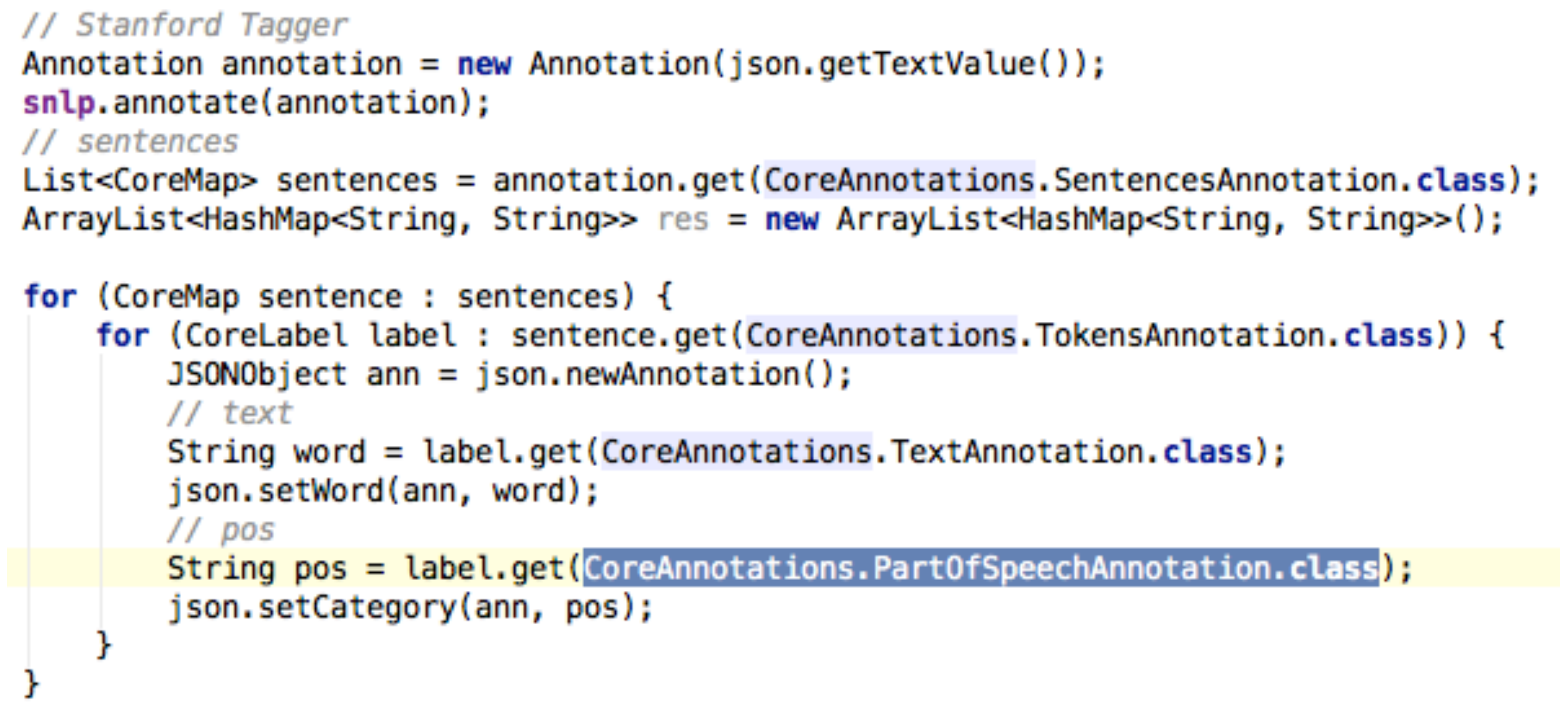
```java

        } else  if (discriminator == Types.JSON) {
            String textjson = data.getPayload();
            JsonTaggerSerialization json = new JsonTaggerSerialization(textjson);

            json.setProducer(this.getClass().getName() + ":" + VERSION);
            json.setType("annotation:tagger");

            // Stanford Tagger
            Annotation annotation = new Annotation(json.getTextValue());
            snlp.annotate(annotation);
            // sentences
            List<CoreMap> sentences = annotation.get(CoreAnnotations.SentencesAnnotation.class);
            ArrayList<HashMap<String, String>> res = new ArrayList<HashMap<String, String>>();

            for (CoreMap sentence : sentences) {
                for (CoreLabel label : sentence.get(CoreAnnotations.TokensAnnotation.class)) {
                    JSONObject ann = json.newAnnotation();
                    // text
                    String word = label.get(CoreAnnotations.TextAnnotation.class);
                    json.setWord(ann, word);
                    // pos
                    String pos = label.get(CoreAnnotations.PartOfSpeechAnnotation.class);
                    json.setCategory(ann, pos);
                }
                                  json(json.toString());
            String name = DiscriminatorRegistry.get(discriminator);
            String message = "Invalid input type. Expected Text but found " + name;
            return DataFactory.error(message);
        }
    }
}
```

```
[< 1.7 >] java.lang
public final class String extends Object
implements Serializable, Comparable<String>, CharSequence
```

https://github.com/chunqishi/org.lappsgrid.example.java.stanfordnlp

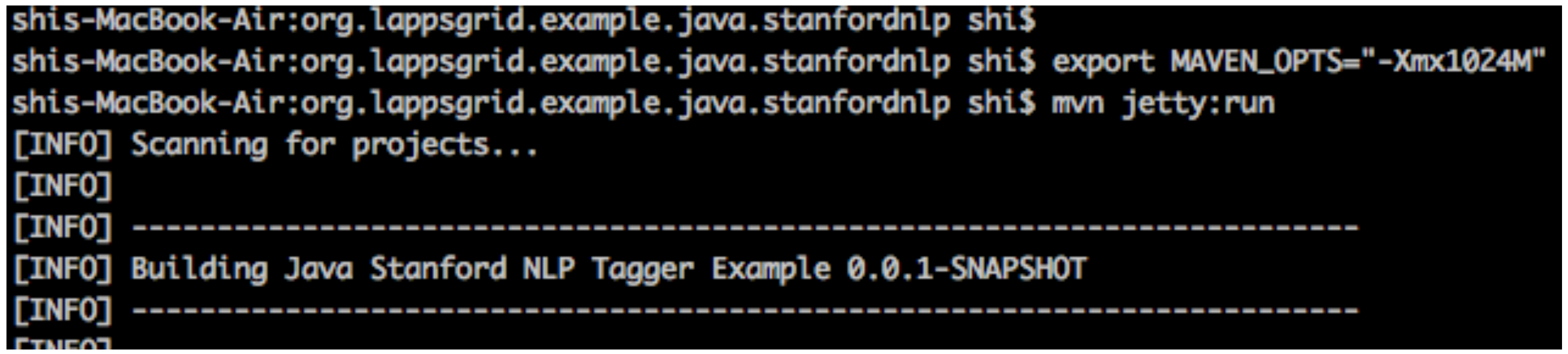
Stanford Tagger Wrapping

- Java Wrapping



```
// Stanford Tagger
Annotation annotation = new Annotation(json.getTextValue());
snlp.annotate(annotation);
// sentences
List<CoreMap> sentences = annotation.get(CoreAnnotations.SentencesAnnotation.class);
ArrayList<HashMap<String, String>> res = new ArrayList<HashMap<String, String>>();

for (CoreMap sentence : sentences) {
    for (CoreLabel label : sentence.get(CoreAnnotations.TokensAnnotation.class)) {
        JSONObject ann = json.newAnnotation();
        // text
        String word = label.get(CoreAnnotations.TextAnnotation.class);
        json.setWord(ann, word);
        // pos
        String pos = label.get(CoreAnnotations.PartOfSpeechAnnotation.class);
        json.setCategory(ann, pos);
    }
}
```

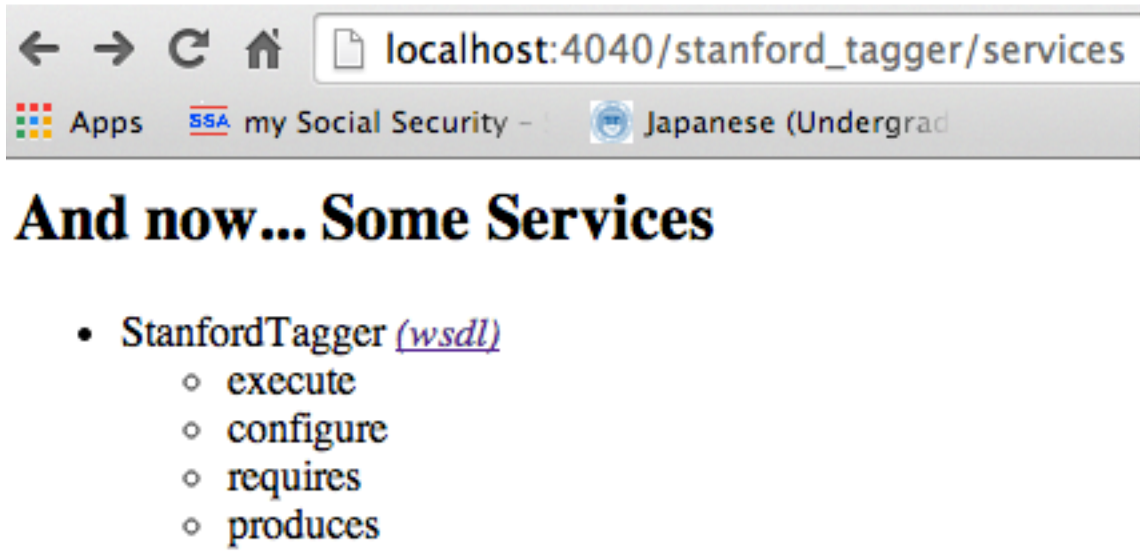
- Jetty Running



```
shis-MacBook-Air:org.lappsgrid.example.java.stanfordnlp shi$
shis-MacBook-Air:org.lappsgrid.example.java.stanfordnlp shi$ export MAVEN_OPTS="-Xmx1024M"
shis-MacBook-Air:org.lappsgrid.example.java.stanfordnlp shi$ mvn jetty:run
[INFO] Scanning for projects...
[INFO]
[INFO] ------------------------------------------------------------------------
[INFO] Building Java Stanford NLP Tagger Example 0.0.1-SNAPSHOT
[INFO] ------------------------------------------------------------------------
```

Stanford Tagger Testing

- Local Service

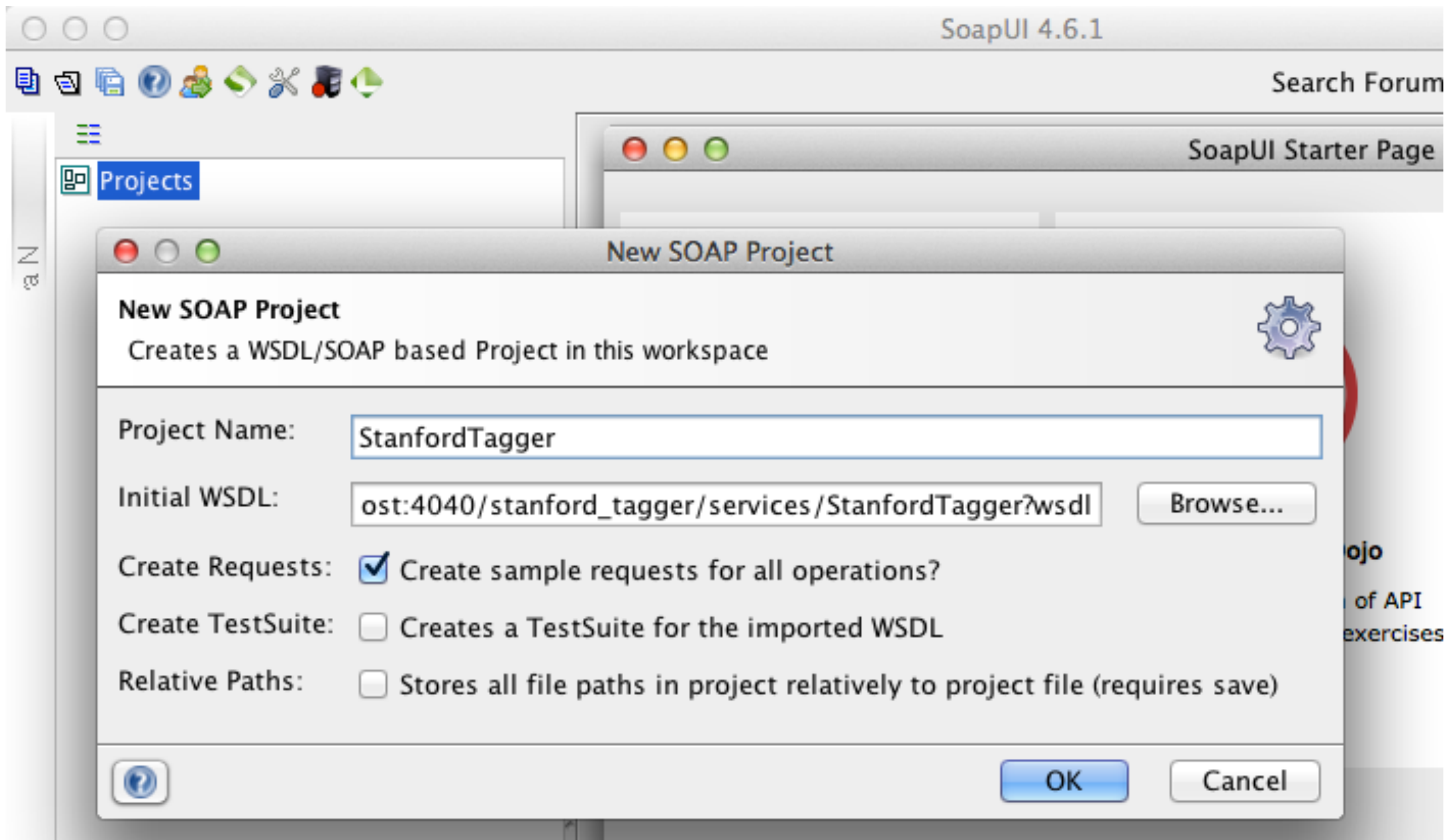


localhost:4040/stanford_tagger/services

And now... Some Services

- StanfordTagger *(wsdl)*
  - execute
  - configure
  - requires
  - produces

- SoapUI Testing

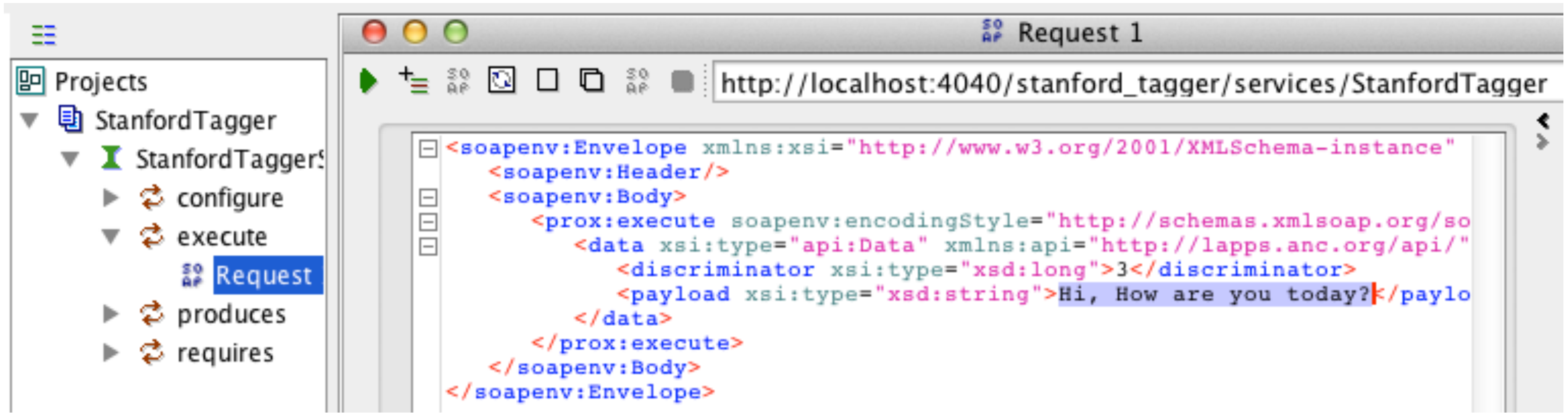


New SOAP Project
Creates a WSDL/SOAP based Project in this workspace

Project Name: StanfordTagger
Initial WSDL: ost:4040/stanford_tagger/services/StanfordTagger?wsdl
Create Requests: ☑ Create sample requests for all operations?
Create TestSuite: ☐ Creates a TestSuite for the imported WSDL
Relative Paths: ☐ Stores all file paths in project relatively to project file (requires save)

Stanford Tagger Testing Result

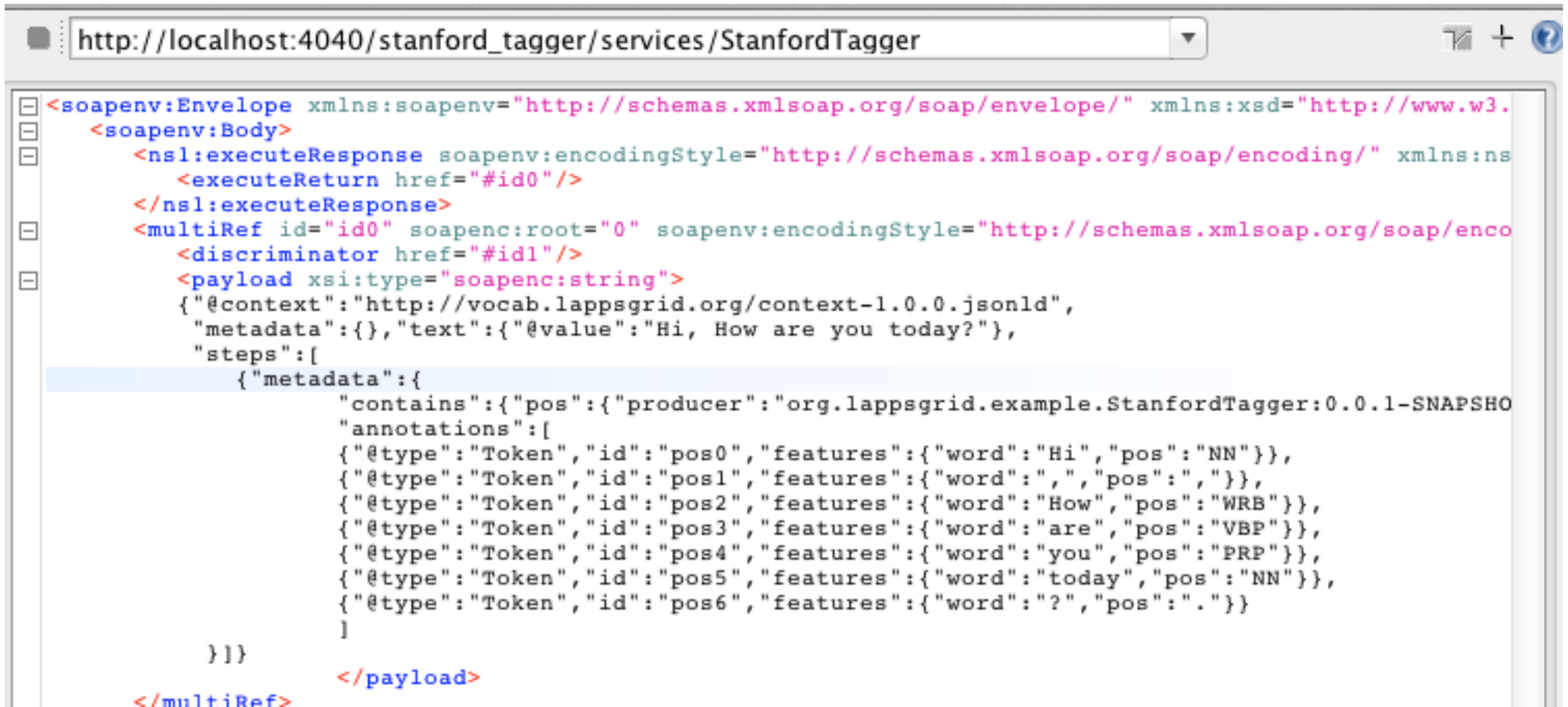
- Request



http://localhost:4040/stanford_tagger/services/StanfordTagger

```
<soapenv:Envelope xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   <soapenv:Header/>
   <soapenv:Body>
      <prox:execute soapenv:encodingStyle="http://schemas.xmlsoap.org/so
         <data xsi:type="api:Data" xmlns:api="http://lapps.anc.org/api/"
            <discriminator xsi:type="xsd:long">3</discriminator>
            <payload xsi:type="xsd:string">Hi, How are you today?</paylo
         </data>
      </prox:execute>
   </soapenv:Body>
</soapenv:Envelope>
```

- Response

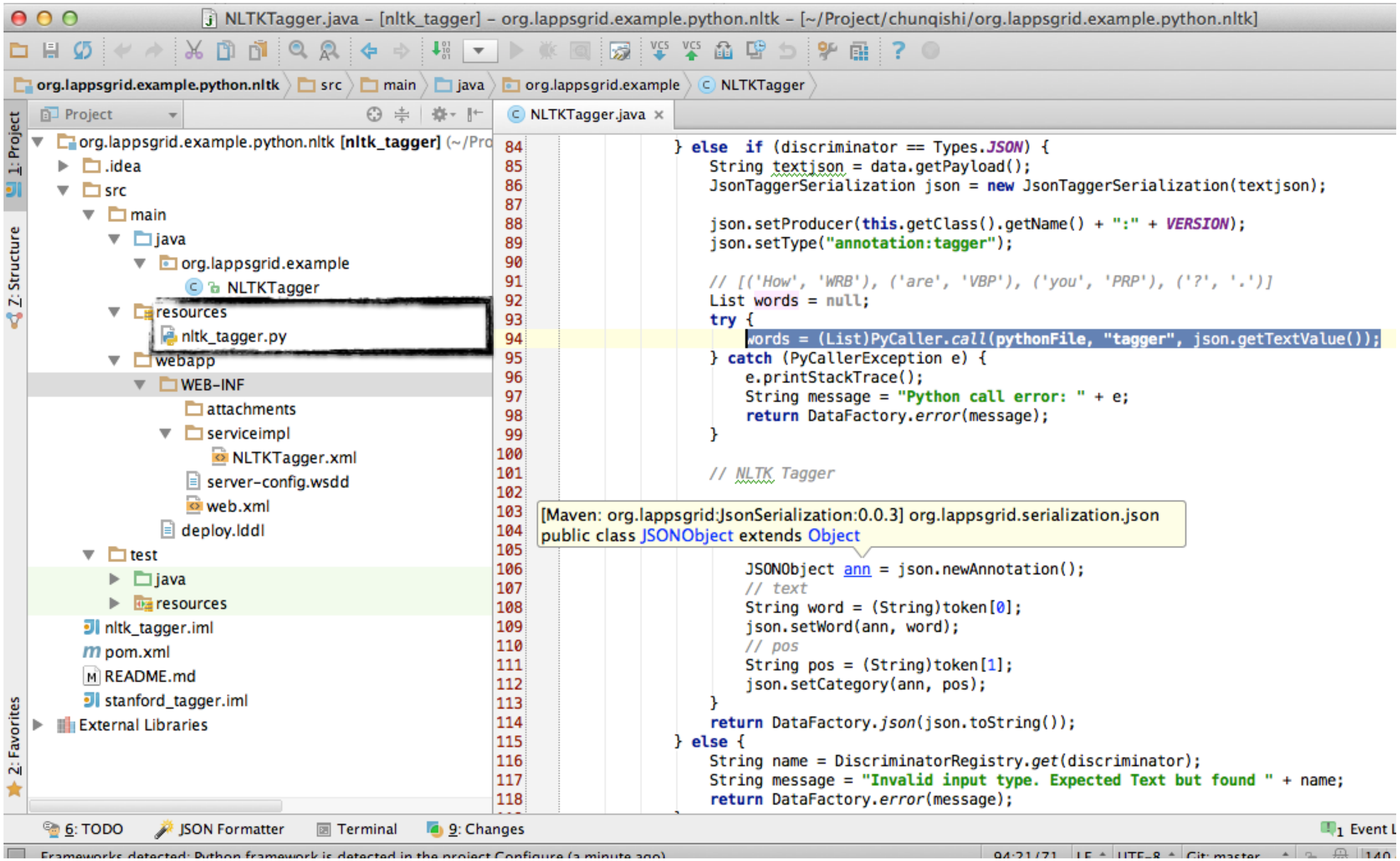


http://localhost:4040/stanford_tagger/services/StanfordTagger

```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/" xmlns:xsd="http://www.w3.
   <soapenv:Body>
      <ns1:executeResponse soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/" xmlns:ns
         <executeReturn href="#id0"/>
      </ns1:executeResponse>
      <multiRef id="id0" soapenc:root="0" soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/enco
         <discriminator href="#id1"/>
         <payload xsi:type="soapenc:string">
         {"@context":"http://vocab.lappsgrid.org/context-1.0.0.jsonld",
          "metadata":{},"text":{"@value":"Hi, How are you today?"},
          "steps":[
            {"metadata":{
                        "contains":{"pos":{"producer":"org.lappsgrid.example.StanfordTagger:0.0.1-SNAPSHO
                        "annotations":[
                        {"@type":"Token","id":"pos0","features":{"word":"Hi","pos":"NN"}},
                        {"@type":"Token","id":"pos1","features":{"word":",","pos":","}},
                        {"@type":"Token","id":"pos2","features":{"word":"How","pos":"WRB"}},
                        {"@type":"Token","id":"pos3","features":{"word":"are","pos":"VBP"}},
                        {"@type":"Token","id":"pos4","features":{"word":"you","pos":"PRP"}},
                        {"@type":"Token","id":"pos5","features":{"word":"today","pos":"NN"}},
                        {"@type":"Token","id":"pos6","features":{"word":"?","pos":"."}}
                        ]
           }]}
                        </payload>
      </multiRef>
```

Lapps Service Wrapping (Python)

Developing Template



```java
        } else  if (discriminator == Types.JSON) {
            String textjson = data.getPayload();
            JsonTaggerSerialization json = new JsonTaggerSerialization(textjson);

            json.setProducer(this.getClass().getName() + ":" + VERSION);
            json.setType("annotation:tagger");

            // [('How', 'WRB'), ('are', 'VBP'), ('you', 'PRP'), ('?', '.')]
            List words = null;
            try {
                words = (List)PyCaller.call(pythonFile, "tagger", json.getTextValue());
            } catch (PyCallerException e) {
                e.printStackTrace();
                String message = "Python call error: " + e;
                return DataFactory.error(message);
            }

            // NLTK Tagger
```

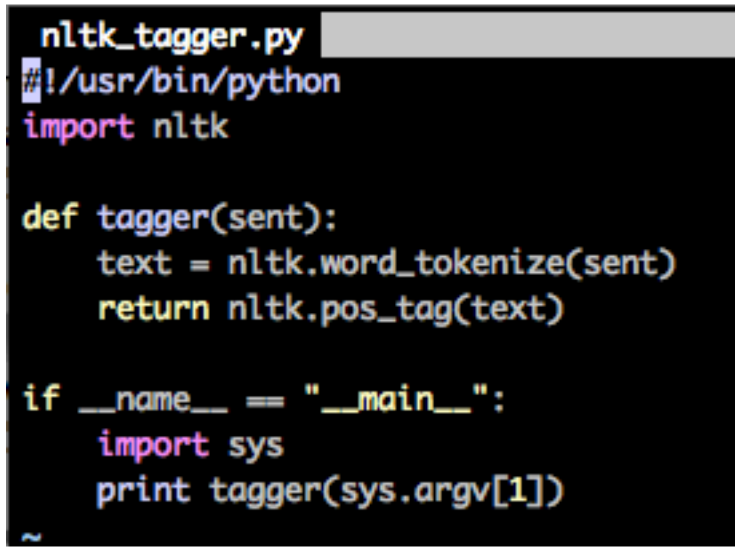
[Maven: org.lappsgrid:JsonSerialization:0.0.3] org.lappsgrid.serialization.json
public class JSONObject extends Object

```java
                JSONObject ann = json.newAnnotation();
                // text
                String word = (String)token[0];
                json.setWord(ann, word);
                // pos
                String pos = (String)token[1];
                json.setCategory(ann, pos);
            }
            return DataFactory.json(json.toString());
        } else {
            String name = DiscriminatorRegistry.get(discriminator);
            String message = "Invalid input type. Expected Text but found " + name;
            return DataFactory.error(message);
```

Frameworks detected: Python framework is detected in the project Configure (a minute ago)   94:21/71   LF   UTF-8   Git: master   140

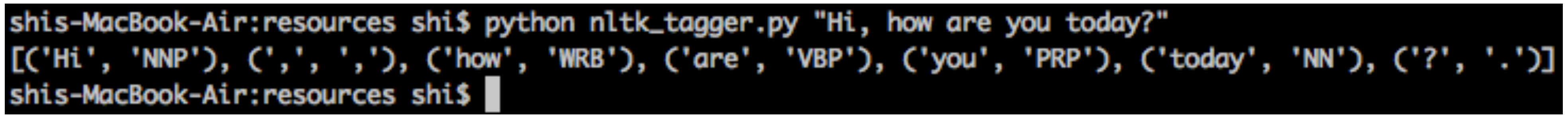
NLTK Python

- Python Program



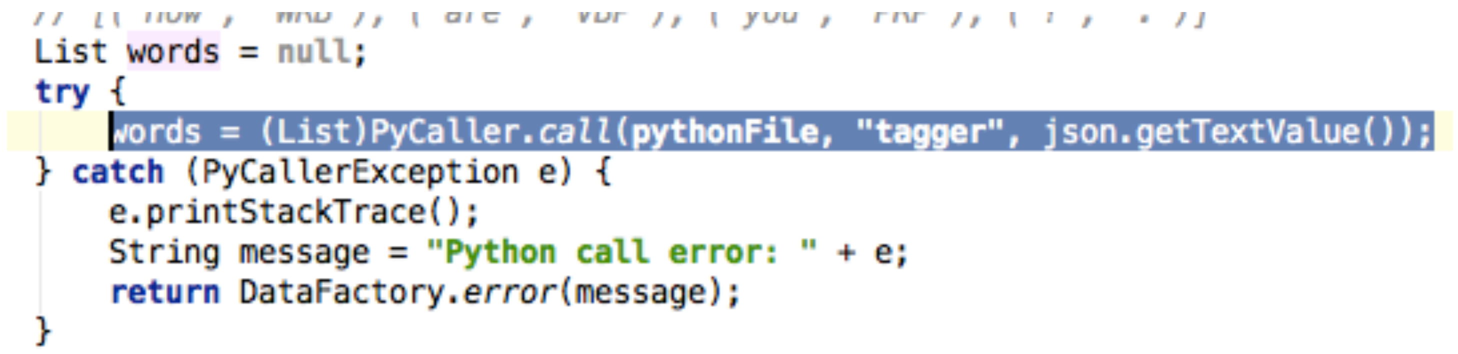
```
nltk_tagger.py
#!/usr/bin/python
import nltk

def tagger(sent):
    text = nltk.word_tokenize(sent)
    return nltk.pos_tag(text)

if __name__ == "__main__":
    import sys
    print tagger(sys.argv[1])
~
```

- Python Result



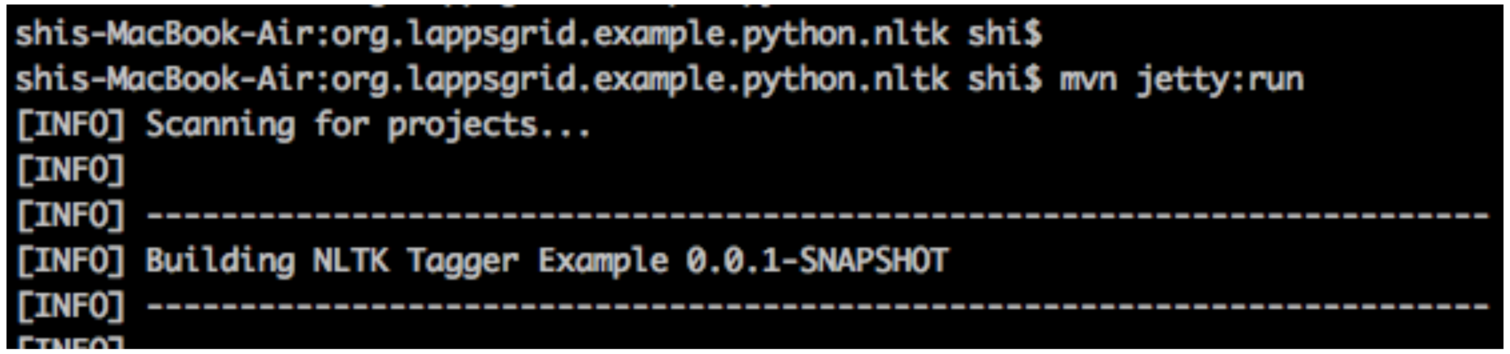
```
shis-MacBook-Air:resources shi$ python nltk_tagger.py "Hi, how are you today?"
[('Hi', 'NNP'), (',', ','), ('how', 'WRB'), ('are', 'VBP'), ('you', 'PRP'), ('today', 'NN'), ('?', '.')]
shis-MacBook-Air:resources shi$
```

- Java Wrapping



```java
List words = null;
try {
    words = (List)PyCaller.call(pythonFile, "tagger", json.getTextValue());
} catch (PyCallerException e) {
    e.printStackTrace();
    String message = "Python call error: " + e;
    return DataFactory.error(message);
}
```

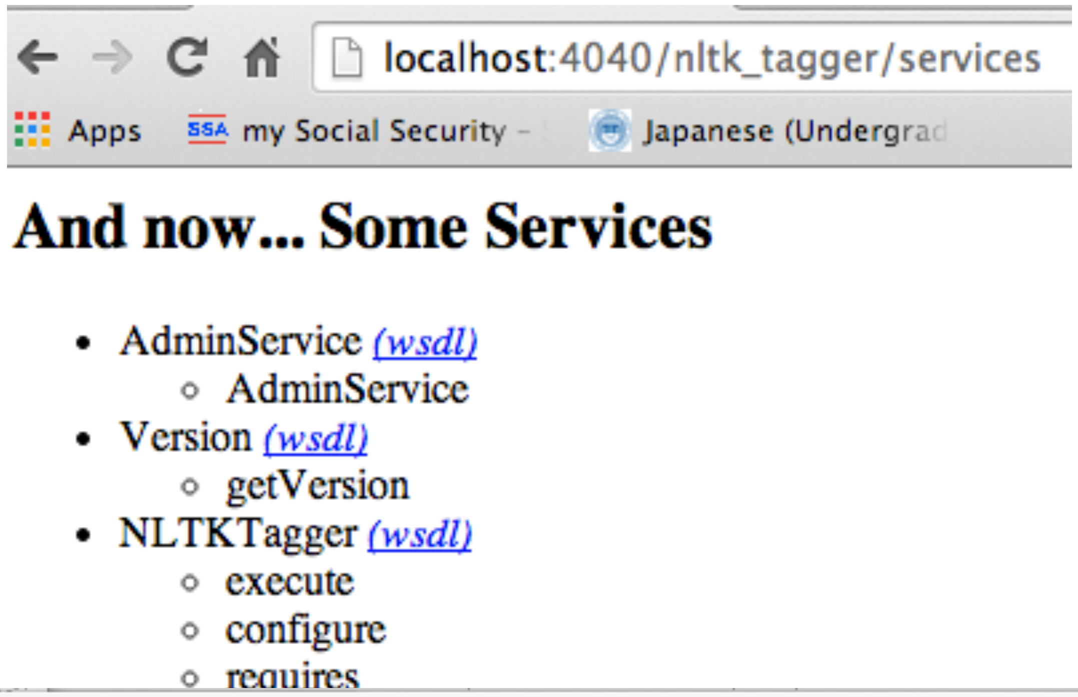
- Jetty Running



```
shis-MacBook-Air:org.lappsgrid.example.python.nltk shi$
shis-MacBook-Air:org.lappsgrid.example.python.nltk shi$ mvn jetty:run
[INFO] Scanning for projects...
[INFO]
[INFO] ------------------------------------------------------------------------
[INFO] Building NLTK Tagger Example 0.0.1-SNAPSHOT
[INFO] ------------------------------------------------------------------------
```

NLTK Tagger Testing

- Local Service



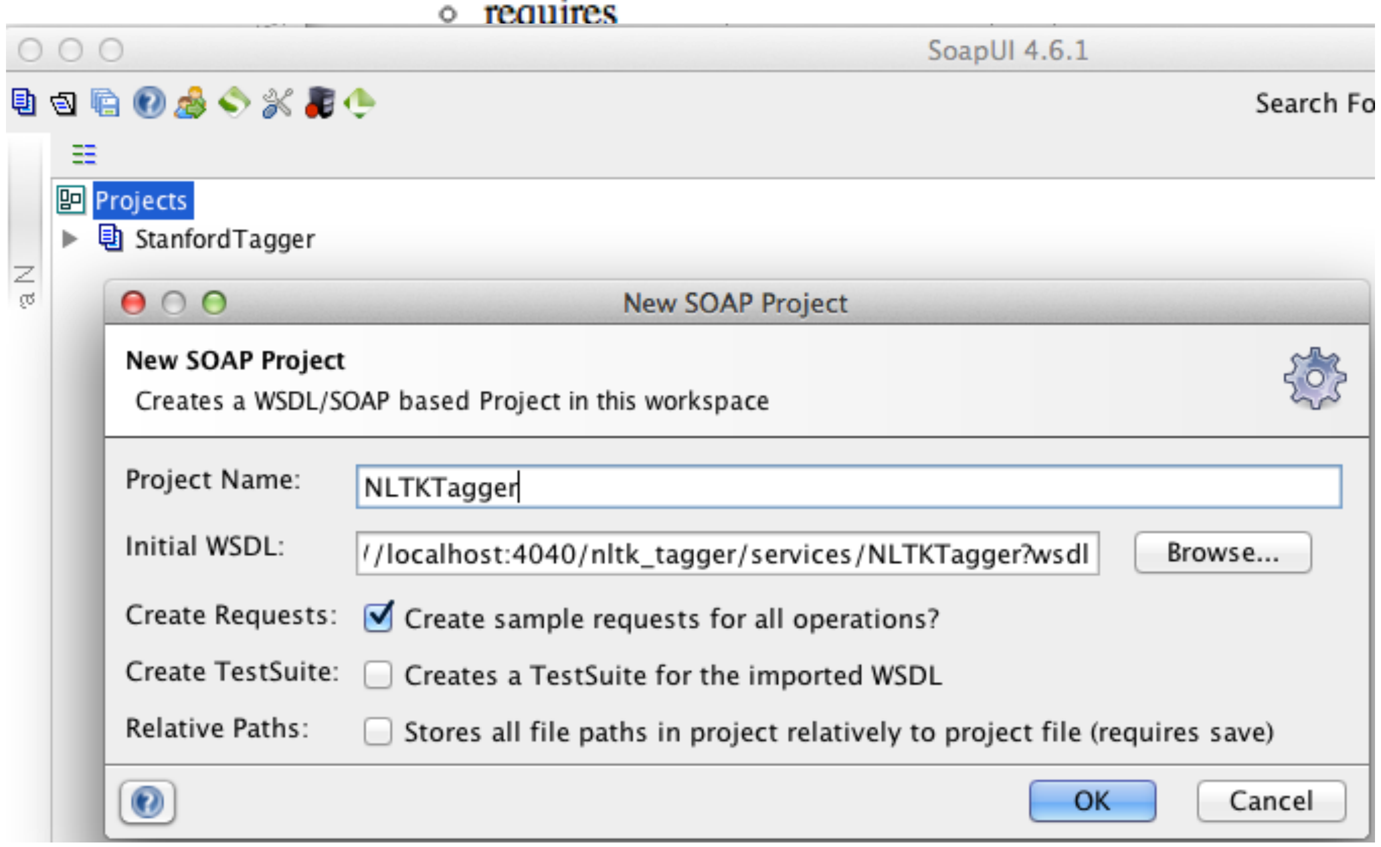
localhost:4040/nltk_tagger/services

Apps · my Social Security – · Japanese (Undergrad

And now... Some Services

- AdminService *(wsdl)*
  - AdminService
- Version *(wsdl)*
  - getVersion
- NLTKTagger *(wsdl)*
  - execute
  - configure
  - requires

- SoapUI Testing



SoapUI 4.6.1

Projects
- StanfordTagger

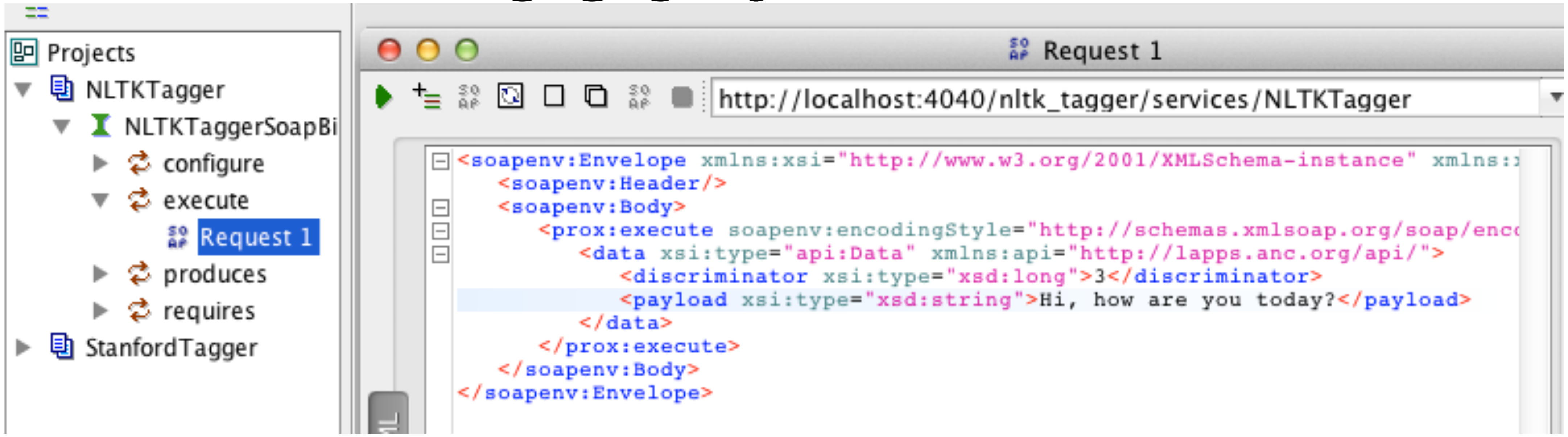
New SOAP Project

New SOAP Project
Creates a WSDL/SOAP based Project in this workspace

Project Name: NLTKTagger
Initial WSDL: //localhost:4040/nltk_tagger/services/NLTKTagger?wsdl
Create Requests: [✓] Create sample requests for all operations?
Create TestSuite: [ ] Creates a TestSuite for the imported WSDL
Relative Paths: [ ] Stores all file paths in project relatively to project file (requires save)

NLTK Tagger Testing Result

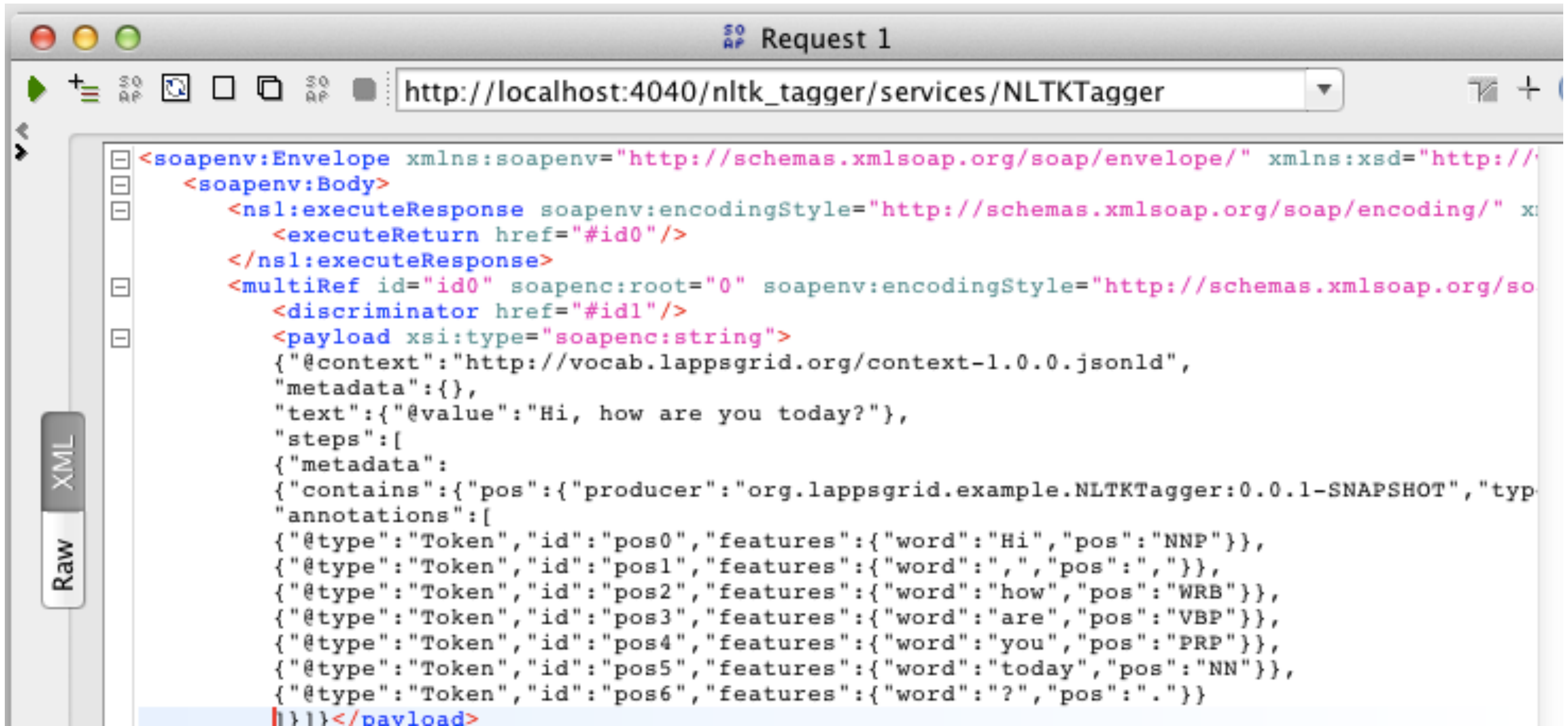
- Request



http://localhost:4040/nltk_tagger/services/NLTKTagger

```
<soapenv:Envelope xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:
   <soapenv:Header/>
   <soapenv:Body>
      <prox:execute soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/enc
         <data xsi:type="api:Data" xmlns:api="http://lapps.anc.org/api/">
            <discriminator xsi:type="xsd:long">3</discriminator>
            <payload xsi:type="xsd:string">Hi, how are you today?</payload>
         </data>
      </prox:execute>
   </soapenv:Body>
</soapenv:Envelope>
```

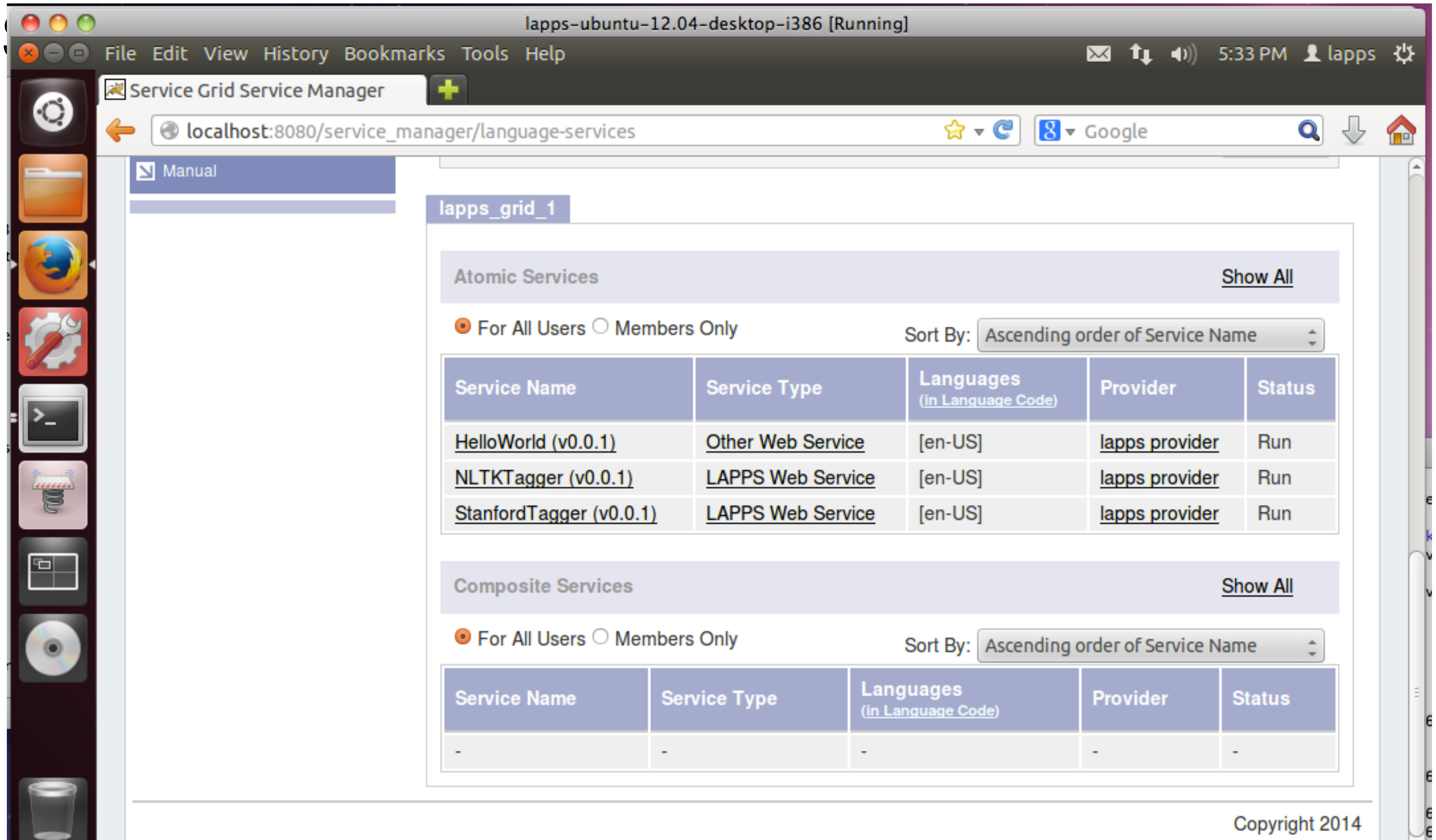
- Response



http://localhost:4040/nltk_tagger/services/NLTKTagger

```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/" xmlns:xsd="http://
   <soapenv:Body>
      <ns1:executeResponse soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/" x
         <executeReturn href="#id0"/>
      </ns1:executeResponse>
      <multiRef id="id0" soapenc:root="0" soapenv:encodingStyle="http://schemas.xmlsoap.org/so
         <discriminator href="#id1"/>
         <payload xsi:type="soapenc:string">
         {"@context":"http://vocab.lappsgrid.org/context-1.0.0.jsonld",
         "metadata":{},
         "text":{"@value":"Hi, how are you today?"},
         "steps":[
         {"metadata":
         {"contains":{"pos":{"producer":"org.lappsgrid.example.NLTKTagger:0.0.1-SNAPSHOT","typ
         "annotations":[
         {"@type":"Token","id":"pos0","features":{"word":"Hi","pos":"NNP"}},
         {"@type":"Token","id":"pos1","features":{"word":",","pos":","}},
         {"@type":"Token","id":"pos2","features":{"word":"how","pos":"WRB"}},
         {"@type":"Token","id":"pos3","features":{"word":"are","pos":"VBP"}},
         {"@type":"Token","id":"pos4","features":{"word":"you","pos":"PRP"}},
         {"@type":"Token","id":"pos5","features":{"word":"today","pos":"NN"}},
         {"@type":"Token","id":"pos6","features":{"word":"?","pos":"."}}
         ]}]}</payload>
```

Service Register


lapps-ubuntu-12.04-desktop-i386 [Running]
File Edit View History Bookmarks Tools Help
5:33 PM lapps
Service Grid Service Manager
localhost:8080/service_manager/language-services
Google
Manual

lapps_grid_1

Atomic Services — Show All

For All Users [✓] Members Only — Sort By: Ascending order of Service Name

Service Name	Service Type	Languages (in Language Code)	Provider	Status
HelloWorld (v0.0.1)	Other Web Service	[en-US]	lapps provider	Run
NLTKTagger (v0.0.1)	LAPPS Web Service	[en-US]	lapps provider	Run
StanfordTagger (v0.0.1)	LAPPS Web Service	[en-US]	lapps provider	Run

Composite Services — Show All

For All Users [✓] Members Only — Sort By: Ascending order of Service Name

Service Name	Service Type	Languages (in Language Code)	Provider	Status
-	-	-	-	-

Copyright 2014

Conclusion

- Contributing to the Lapps Grid
  - Wrapping Lapps Service
    - Java / Python Wrapping
    - Templates from Github Repository
  - Registering into Service Manager
    - Service Manager Installation Script
  - Developing Environment
    - VirtualBox Image: Ubuntu

Reference

- API Docs: http://www.anc.org/projects/lapps/api/project-info.html
- Service Templates:
  - https://github.com/chunqishi/org.lappsgrid.example.java.helloworld
  - https://github.com/chunqishi/org.lappsgrid.example.java.stanfordnlp
  - https://github.com/chunqishi/org.lappsgrid.example.python.nltk
- Service Managers:
  - http://eldrad.cs-i.brandeis.edu/service_manager/language-services
  - http://grid.anc.org:8080/service_manager/language-services
- VirtualBox Image:
  - http://eldrad.cs-i.brandeis.edu/download/lapps-ubuntu-12.04-desktop-i386.tar.gz