



FP7-ICT-2013-10

***Xlike Pipeline &
Integration***

Outline

- **Xlike pipelines and API's**
- **Services**
- **Tutorial**
 - **Data format**
 - **Calling services**
 - **Java client**
 - **Curl example**
 - **Data format**
- **APIs**

Some examples to access to the data are at:

<https://github.com/xLime-project/Tutorial>

Send an email to:

ngarcia@isoco.com

or

egarcia@isoco.com

For gaining accessibility, though it is public so far....

Multilingual Pipelines and API's

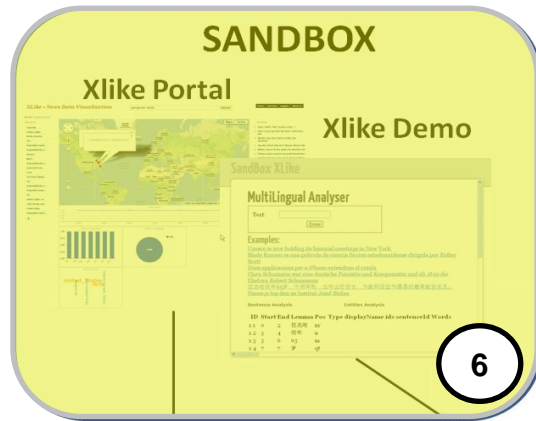
| Service | Description | URL |
|--------------------------------|---------------------------------|---|
| Language Identification | Language Identification. | http://95.87.154.251:9090/services/language_code/ident |
| English Service | English pipeline. | http://95.87.154.251:9090/services/analysis_en/analyze |
| Spanish Service | Spanish pipeline. | http://95.87.154.251:9090/services/analysis_es/analyze |
| Catalan Service | Catalan pipeline. | http://95.87.154.251:9090/services/analysis_ca/analyze |
| German Service | German pipeline. | http://lt.ffzg.hr:9090/xlike/analysis_de/analyze |
| Chinese Service | Chinese pipeline. | http://keg.cs.tsinghua.edu.cn:8080/analysis_zh/analyze |
| Slovenian Service | Slovene pipeline. | http://enrycher.ijs.si/xlike/analysis_sl/analyze |
| English annotation | Annotation service for English. | http://km.aifb.kit.edu/services/annotation-en/ |
| Spanish annotation | Annotation service for Spanish. | http://km.aifb.kit.edu/services/annotation-es/ |
| German annotation | Annotation service for German. | http://km.aifb.kit.edu/services/annotation-de/ |

- Language code following the ISO 639-2 specification (en, es, ca)
 - POST type calls

Integration (Service Level)

| Language | Description | URL SandBox | Parameters | Example of use |
|--|--|--|---|---|
| Language Identification Service | base_url/language_code/ident | http://sandbox.xike.isocp.com/services/language_code/ident | <analyze> <text>Article</text> </analyze> | <analyze><text>Bruce Springsteen is an American singer-songwriter and multi-instrumentalist.</text></analyze> |
| English Service | base_url/ analysis_en /analyze | http://sandbox.xike.isocp.com/services/analysis_en/analyze | <analyze> <text>Article</text> <target>relations</target> <conll>>true</conll> </analyze> | <analyze><text>Bruce Springsteen is an American singer-songwriter and multi-instrumentalist.</text><target>relations</target><conll>>true</conll></analyze> |
| Spanish Service | base_url/ analysis_es /analyze | http://sandbox.xike.isocp.com/services/analysis_es/analyze | <analyze> <text>Article</text> <target>relations</target> <conll>>true</conll> </analyze> | <analyze><text>Blade Runner es una película de ciencia ficción estadounidense dirigida por Ridley Scott.</text><target>relations</target><conll>>true</conll></analyze> |
| Catalan Service | base_url/ analysis_ca /analyze | http://sandbox.xike.isocp.com/services/analysis_ca/analyze | <analyze> <text>Article</text> <target>relations</target> <conll>>true</conll> </analyze> | <analyze><text>L'iPhone es un dispositiu electrònic multimèdia presentat per Apple Computer el 9 de gener de 2007.</text><target>relations</target><conll>>true</conll></analyze> |
| German Service | base_url/ analysis_de /analyze | http://sandbox.xike.isocp.com/services/analysis_de/analyze | <analyze> <text>Article</text> <conll>>true</conll> </analyze> | <analyze><text>Clara Schumann war eine deutsche Pianistin und Komponistin und ab 1840 die Ehefrau Robert Schumanns.</text><conll>>true</conll></analyze> |
| Chinese Service | base_url/ analysis_zh /analyze | http://sandbox.xike.isocp.com/services/analysis_zh/analyze | <analyze> <text>Article</text> <conll>>true</conll> </analyze> | <analyze><text>正面的观点认为,由于元朝从忽必烈即位后就开始“行汉法”,</text><conll>>true</conll></analyze> |
| Slovenian Service | base_url/ analysis_sl /analyze?text=text to identify | http://aidemo.ijs.si/xike/analysis_sl/analyze | <analyze> <text>Article</text> <conll>>true</conll> </analyze> | <analyze><text>Clara je bila žena skladatelja Roberta Schumann in ena vodilnih pianistov in skladateljev romantike.</text><conll>>true</conll></analyze> |

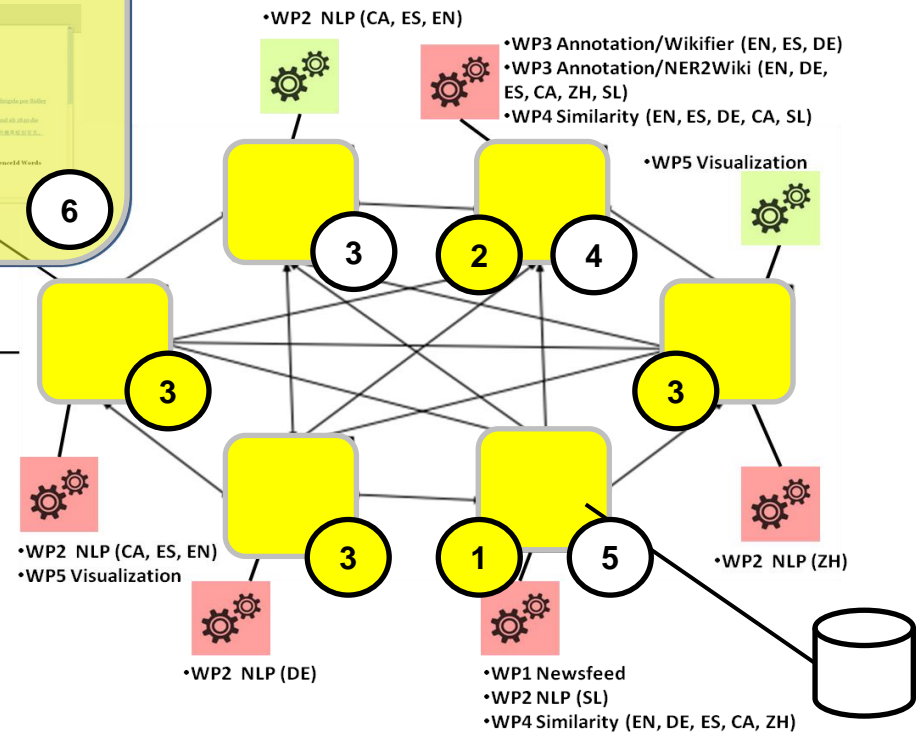
> 30 working services included at any of the demos/prototypes of the project



Initial deployment of services



Current deployed services



Linguistic analysis: token, pos, lemma, id.

```
<item>
  <services>
    <service name="UPC-analysis" date="2013-07-18"/>
  </services>
  <sentences>
    <sentence id="1">
      <text>Yesterday a house was consumed by fire between Henderson and Belle Plaine in Sibley County. </text>
      <tokens>
        <token pos="NN" end="9" lemma="yesterday" id="1.1" start="0">Yesterday</token>
        <token pos="2" end="11" lemma="1" id="1.2" start="10">a</token>
        <token pos="NN" end="17" lemma="house" id="1.3" start="12">house</token>
        <token pos="VBD" end="21" lemma="be" id="1.4" start="18">was</token>
        <token pos="VBN" end="30" lemma="consume" id="1.5" start="22">consumed</token>
        <token pos="IN" end="33" lemma="by" id="1.6" start="31">by</token>
        <token pos="NN" end="38" lemma="fire" id="1.7" start="34">fire</token>
        <token pos="IN" end="46" lemma="between" id="1.8" start="39">between</token>
        <token pos="NP00SPO" end="56" lemma="henderson" id="1.9" start="47">Henderson</token>
        <token pos="CC" end="60" lemma="and" id="1.10" start="57">and</token>
        <token pos="NP00G00" end="73" lemma="belle_plaine" id="1.11" start="61">Belle_Plaine</token>
        <token pos="IN" end="76" lemma="in" id="1.12" start="74">in</token>
        <token pos="NP00G00" end="90" lemma="sibley_county" id="1.13" start="77">Sibley_County</token>
        <token pos="Fp" end="91" lemma="." id="1.14" start="90">.</token>
      </tokens>
    </sentence>
  </sentences>
</item>
```

Node: can be a word, an entity (location, organization, person)

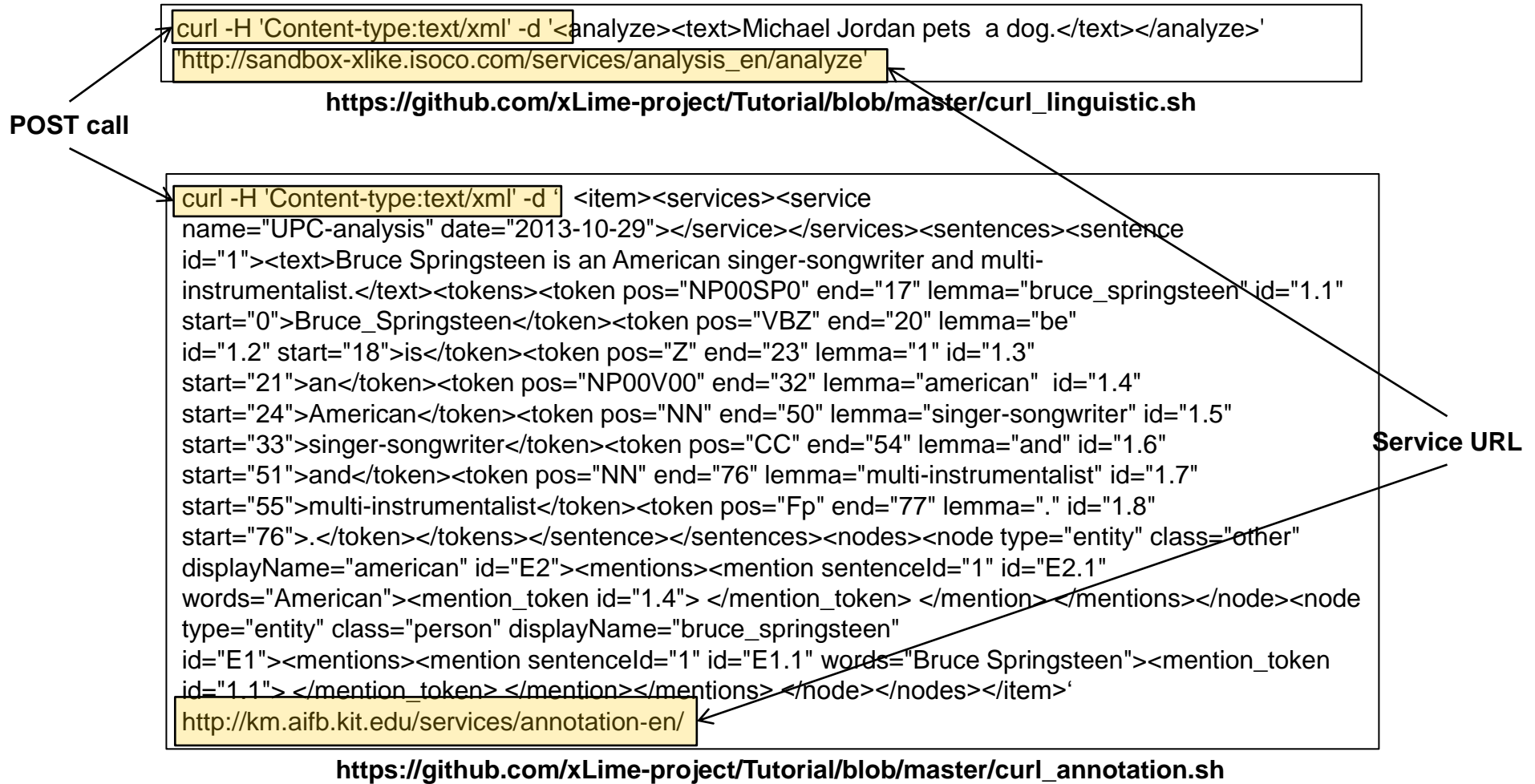
```
<node>
  <node type="entity" class="organization" displayName="assistant_fire_chief" id="E10">
    <mentions>
      <mention sentenceId="6" id="E10.1" words="Assistant Fire Chief">
        <mention_token id="6.3"/>
      </mention>
    </mentions>
  </node>
  <node type="entity" class="location" displayName="belle_plaine_fire_departments" id="E2">
    <mentions>
      <mention sentenceId="1" id="E2.1" words="Belle Plaine">
        <mention_token id="1.11"/>
      </mention>
      <mention sentenceId="5" id="E2.2" words="Belle Plaine">
        <mention_token id="5.1"/>
      </mention>
      <mention sentenceId="6" id="E2.3" words="Belle Plaine Fire Departments">
        <mention_token id="6.6"/>
      </mention>
    </mentions>
  </node>
```


Frame: includes actions as a set of composes elements

```
<frame sentenceId="1" displayName="consume.01" id="F13" tokenId="1.5">
  <argument role="A1" displayName="house.01" frame="true" id="F12"/>
  <argument role="A0" displayName="fire.01" frame="true" id="F14"/>
  <argument role="AM-LOC" displayName="sibley_county" id="E3"/>
  <descriptions>
    <description URI="01157517-v" displayName="consume,deplete,eat,eat_up,exhaust,run_through,use_up,wipe_out" knowledgeBase="WordNet-3.0"/>
  </descriptions>
</frame>
```

Frame: includes annotations as descriptors

```
<frame sentenceId="1" displayName="consume.01" id="F13" tokenId="1.5">
  <argument role="A1" displayName="house.01" frame="true" id="F12"/>
  <argument role="A0" displayName="fire.01" frame="true" id="F14"/>
  <argument role="AM-LOC" displayName="sibley_county" id="E3"/>
  <descriptions>
    <description URI="01157517-v" displayName="consume,deplete,eat,eat_up,exhaust,run_through,use_up,wipe_out" knowledgeBase="WordNet-3.0"/>
  </descriptions>
</frame>
```



Cross-lingual Annotations Output

```
</descriptions>
<mentions>
<mention sentenceId="1" words="singer-songwriter"/>
</mentions>
</annotation>
<annotation displayName="Bruce Springsteen" entityId="2" weight="0.962">
<descriptions>
<description URL="http://en.wikipedia.org/wiki/Bruce_Springsteen" lang="en"/>
<description URL="http://de.wikipedia.org/wiki/Bruce_Springsteen" lang="de"/>
<description URL="http://es.wikipedia.org/wiki/Bruce_Springsteen" lang="es"/>
<description URL="http://zh.wikipedia.org/wiki/布鲁斯·斯普林斯汀" lang="zh"/>
<description URL="http://ca.wikipedia.org/wiki/Bruce_Springsteen" lang="ca"/>
</descriptions>
<mentions>
<mention sentenceId="1" words="Bruce Springsteen"/>
</mentions>
</annotation>
<annotation displayName="Multi-instrumentalist" entityId="3" weight="0.888">
<descriptions>
<description URL="http://en.wikipedia.org/wiki/Multi-instrumentalist" lang="en"/>
<description URL="http://de.wikipedia.org/wiki/Multiinstrumentalist" lang="de"/>
<description URL="http://es.wikipedia.org/wiki/Multiinstrumentista" lang="es"/>
</descriptions>
<mentions>
<mention sentenceId="1" words="multi-instrumentalist"/>
</mentions>
</annotation>
<annotation displayName="United States" entityId="4" weight="0.566">
<descriptions>
<description URL="http://en.wikipedia.org/wiki/United_States" lang="en"/>
<description URL="http://de.wikipedia.org/wiki/Vereinigte_Staaten" lang="de"/>
<description URL="http://es.wikipedia.org/wiki/Estados_Unidos" lang="es"/>
<description URL="http://sl.wikipedia.org/wiki/Združene_države_Amerike" lang="sl"/>
<description URL="http://zh.wikipedia.org/wiki/ç%â" lang="zh"/>
<description URL="http://ca.wikipedia.org/wiki/Estats_Units_d'Amãrica" lang="ca"/>
</descriptions>
```



https://github.com/xLime-project/Tutorial/blob/master/Linguistic_Example.java

https://github.com/xLime-project/Tutorial/blob/master/Annotation_example.java.txt

```

<article id="internal article ID; consistent across streams">
  <source>
    <hostname> Publisher hostname </hostname>
    <title> Name of the publisher; failing that, title of the RSS feed </title>
    <location*?>
      <longitude*?> publisher longitude in degrees </longitude>
      <latitude*?> publisher latitude in degrees </latitude>
      <city*?> publisher city </city>
      <country*?> publisher country </country>
    </location>
    <tags*?>
      <tag*?> a tag for the publisher; the vocabulary is not controlled </tag>
    </tags>
  </source>
  <feed*?>
    <uri> URL from which the article was discovered; typically the RSS feed </uri>
  </feed>
  <uri> URL from which the article was downloaded </uri>
  <publish-date*?> The publication time and date. </publish-date>
  <retrieve-date*?> The retrieval time and date. </retrieve-date>
  <lang> 3-letter ISO 639-2 language code </lang>
  <location*?>
    <longitude*?> story content longitude in degrees </longitude>
    <latitude*?> story content latitude in degrees </latitude>
    <city*?> story city </city>
    <country*?> story country </country>
  </location>
  <tags*?>
    <tag*?> a tag for the article; the vocabulary is not controlled </tag>
  </tags>
  <img*?> The URL of a related image, usually a thumbnail. </img>
  <title> Title. Can be empty if we fail to identify it. </title>
  <body-cleartext>
    Clear text body of the article, formatted only with <p> tags
  </body-cleartext>
  <body-rych*? only English, Slovene >
    Enriched article body; an XML subtree as returned by Enrycher.
  </body-rych>
  <body-xlike*? only English, Spanish, Catalan >
    Enriched article body; an XML subtree as returned by iSOCO; experimental.
  </body-xlike>
</article>

```

Article information

Enrycher
information
XLike
information

IJS newsfeed

a clean, continuous, real-time aggregated stream of
semantically enriched news articles from RSS-enabled sites across the world.

<https://github.com/xLime-project/Tutorial/blob/master/http2fs.py>

<http://newsfeed.ijs.si/http2fs.py>

- Design of domain ontology

Ontology header:

Ontology IRI

Ontology Version IRI

Annotations +

label [language: es]
Ontología de dominio sobre X-Like (Cross-lingual Knowledge Extraction)

label [language: en]
Domain Ontology X-Like (Cross-lingual Knowledge Extraction)

homepage
<http://www.xlike.org/>

versionInfo [type: string]
v0.1

Object property hierarchy:

- topObjectProperty
 - appearsIn
 - hasArgument
 - hasDescription
 - hasFrame
 - hasMention
 - hasNode
 - hasSentence
 - hasService
 - hasToken

Annotations +

label [language: es]
Argumento

label [language: en]
Argument

comment [language: es]
Argumento de un evento. Se puede referir a un nodo, o a otro evento

Description: Argumento

- Thing
 - Argumento
 - Description
 - Frame
 - Item
 - Mention
 - Nodo
 - Sentencia
 - Servicio
 - Token

More Information

- www.xlike.org
- <http://www.xlike.org/home/deliverables/>
 - D1.3.2
 - D6.1.2
 - D7.3.1
- <http://www.xlike.org/sandbox/>
- <http://sandbox-xlike.isoco.com>
- <http://meta-share.ffzg.hr/repository/search/?q=xlike>



¿Quieres innovar?

Esteban García-Cuesta

egarcia@isoco.com



Karlsruhe 25/11/2013

iSOCO Barcelona

Av. Torre Blanca, 57
Edificio ESADE CREAPOLIS
Oficina 3C 15
08172 Sant Cugat del Vallès
Barcelona
España
(t) +34 935 677 200

iSOCO Madrid

Av. del Partenón, 16-18, 1ªª
Campo de las Naciones
28042 Madrid
España
(t) +34 913 349 797

iSOCO Pamplona

Parque Tomás
Caballero, 2, 6ªª
31006 Pamplona
España
(t) +34 948 102 408

iSOCO Valencia

C/ Prof. Beltrán Báuena, 4
Oficina 107
46009 Valencia
España
(t) +34 963 467 143

iSOCO Colombia

Complejo Ruta N
Calle 67, 52-20
Piso 3, Torre A
Medellín
Colombia
(t) +57 516 7770 ext. 1132