

XML e Web Services (e workflow management systems)

Paolo Romano
Istituto Nazionale per la Ricerca sul Cancro
(paolo.romano@istge.it)

Sommario

- Caratteristiche dei dati e della loro integrazione in biologia
- eXtensible Markup Language (XML)
- Web Services
- Soaplab ed esempi di implementazione
- Workflow management systems
- Oncology over Internet

Gestione delle informazioni

- Archiviazione locale, isolata, mainframe
- Archiviazione locale, condivisa, LAN
- Archiviazione decentrata, non condivisa
- Archiviazione decentrata, condivisa
→ integrazione
- La rete come archivio locale, anche se decentrato
→ data GRID
- Quale integrazione: dati, applicazioni, processi

L'informazione biologica

- La ricerca biomedica produce una quantità sempre crescente di dati
- Alcuni settori, quali la genomica e la proteomica, contribuiscono alla realizzazione di banche dati di rilevanti dimensioni
- Altri settori emergenti, legati all'analisi delle mutazioni, ai polimorfismi, al metabolismo, e derivati da nuovi strumenti quali i microarray, contribuiranno anch'essi con quantità di dati ancora superiori

Dimensione informazioni biologia

- EMBL Data Library 81 (Dic 2004):
 - Sequenze: 46,105,397 (Basi: 79,271,300,840)
 - EMBL Data Library 82 (Mar 2005):
 - Sequenze: 49,474,402 (Basi: 85,134,714,382)
 - Aumento: +7,3% (+7,4%)
 - EMBL Data Library 83 (Giu 2005):
 - Sequenze: 54,507,330 (Basi: 94,994,772,342)
 - Aumento: +10,2% (+11,6%)
 - EMBL Data Library 84 (Set 2005):
 - Sequenze: 58,758,902 (Basi: 107,562,580,723)
 - Aumento: +7,8% (+13,2%)
- Questa enorme quantità di dati può essere analizzata solo tramite software specialistici

Banche dati eterogenee

- Alcune banche dati sono gestite da pochi Centri (EBI, NCBI, DDBJ) in modo disomogeneo
- Le banche dati secondarie sono di ottima qualità (annotazione estesa, controllo accurato)
- Molte banche dati sono specializzate: gene/genoma, organismo, malattia, mutazione, etc...
- Molte banche dati sono sviluppate da singoli ricercatori o piccolo gruppi
- Le banche dati sono distribuite:
 - strutture dati, DBMS e metodi di distribuzione sono differenti
 - informazioni e significati (semantica) sono differenti

Software per analisi biologiche

- software specialistici sono essenziali per quasi tutte le analisi in biologia molecolare:
 - analisi di sequenza, predizione di strutture proteiche secondarie e terziarie, predizione di siti di splicing, analisi dell'evoluzione molecolare, ricostruzione sequenze, etc...
- è necessario integrare questi software con le banche dati
 - Banche dati come input dei software
 - Risultati elaborazione/analisi come nuovi dati da memorizzare e analizzare

Obiettivi dell'integrazione

- L'integrazione dei dati e l'automazione dei processi sono necessari per:
 - Eseguire automaticamente interrogazioni e/o analisi che coinvolgono più database e software
 - Eseguire con efficienza analisi che coinvolgono grosse quantità di dati
 - Ottenere una visione complessiva e più precisa di tutte le informazioni disponibili
 - Realizzare un effettivo data mining

Integrazione: grado

	Sistemi strettamente integrati	Sistemi a integrazione dinamica
Database	data warehouse	decentrati
Applicazioni	centralizzate, CORBA	Web Services, Data GRID
Processi	servizi statici, ripetitivi	servizi dinamici, adattati su richiesta
Integrazione	statica, predefinita	dinamica, a richiesta o su scoperta
Link	predefiniti, sintattici	“on-the-fly”, semantici
Trasparenza	nulla o scarsa	possibile

Integrazione: longevità

- L'integrazione necessita di stabilità:
 - Buona conoscenza e comprensione del dominio
 - Buona definizione dei dati
 - Standardizzazione
 - Obiettivi ben definiti
- L'integrazione teme:
 - Incertezza nella comprensione del dominio
 - Eterogeneità dei dati e dei sistemi
 - Specializzazione dei dati
 - Rapida evoluzione dei dati
 - Spontaneità, sperimentalismo
 - Mancanza di obiettivi predefiniti

Specificità dell'integrazione

- In ambito biologico:
 - Una pre-analisi e riagggregazione delle informazioni è impossibile, perché dati e conoscenze teoriche cambiano frequentemente e rapidamente
 - La complessità delle informazioni rende molto difficile creare modelli validi in diversi ambiti e che possano rimanere tali nel tempo
 - La disponibilità di strumenti assestati riduce le possibilità di implementare standard comuni
 - Le esigenze e gli obiettivi di ricerca evolvono rapidamente, seguendo le nuove acquisizioni e teorie
- L'integrazione deve quindi essere sviluppata utilizzando sistemi flessibili che siano facilmente adattabili ed espandibili

Integrazione delle banche dati

- L'integrazione delle banche dati comporta
 - L'analisi dei dati disponibili
 - L'identificazione dei collegamenti tra informazioni presenti in banche dati diverse
 - La definizione e l'implementazione di formati comuni per l'interscambio delle informazioni
 - L'analisi e la definizione accurata e univoca degli “oggetti biologici” coinvolti

I metodi dell'integrazione

- **Sintattici**

- Riferimenti diretti espliciti (xrefs) o impliciti (nomi)
- Descrizioni condivise (vocabolari)

- **Semantici**

- Modelli dei dati e schemi condivisi
- Ontologie condivise e associate a dati/strutture

Riferimenti reciproci

- Da un record, a un record collegato di un'altra banca dati:
 - Link diretto, univoco, non reciproco
 - ID database remoto
 - Formati standardizzati
 - Life Science ID
 - Standard I3C
- Limitazioni:
 - Annotazione manuale
 - Significato del collegamento
 - Predefiniti

Descrizioni condivise

- Da un record ai record di un'altra banca dati tramite ricerca testuale:
 - Link implicito, reciproco
 - Determinabile automaticamente
 - Termine di vocabolario
 - Vocabolari standardizzati
- Limitazioni:
 - Diffusione di vocabolari condivisi nell'annotazione
 - Significato del collegamento
 - Necessità di definire l'ambito
 - Text mining

Modelli, schemi, ontologie

- Da un record ai record di un'altra banca dati tramite ricerca mediata da interfaccia:
 - Semantica e ambito ben definiti
 - Determinabile automaticamente
 - Ricerca tramite interfaccia standard astratta
 - Esecuzione della richiesta sulla banca dati
 - Restituzione del risultato in formato standardizzato
 - Modelli, schemi, ontologie condivise
- Limitazioni:
 - Diffusione di modelli e strumenti condivisi
 - Competenze informatiche di alto livello

XML (eXtensible Markup Language)

- Linguaggio Markup per inserire semantica del documento:
 - Supera limiti HTML
 - Semplice definizione e implementazione tramite Document Type Definitions (DTDs)
 - Modulare, nuovi DTD utilizzano precedenti
- Utilizzabile da applicazioni software:
 - Corretto (Well formed, conforme allo standard XML)
 - Valido (conforme al DTD)
 - Definizioni standardizzate (namespaces)
 - Analisi sintattica ed estrazione dati automatiche

ID AA415057 standard; RNA; EST; 337 BP.
AC AA415057;
SV AA415057.1
DT 27-OCT-1997 (Rel. 53, Created)
DT 14-DEC-1999 (Rel. 62, Last updated, Version 2)
DE Mg0001 RCW Lambda Zap Express Library Pyricularia grisea cDNA clone RCW1
DE 5', mRNA sequence.
KW EST.
OS Magnaporthe grisea
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
RN [1]
RP 1-337
RA Wu S.-C., Bernstein B.D., Darvill A.G., Albersheim P.;
RT "Expressed sequence tags of the rice blast fungus grown on rice cell
RT walls";
RL Unpublished.
DR UNILIB; 863; 863.
CC Contact: Sheng-Cheng Wu
CC CCRC
CC University of Georgia
CC 220 Riverbend Road, Athens, GA 30602-4712, USA
CC Tel: 706 542 4446
CC Fax: 706 542 4412
CC Email: wusc@bscr.uga.edu
CC Seq primer: T3.

```
FH      Key      Location/Qualifiers
FH
FT      source    1..337
FT              /db_xref="taxon:148305"
FT              /db_xref="UNILIB:863"
FT              /note="Vector: Lambda Zap; Messenger RNAs prepared from
FT              Magnaporthe grisea grown at 23C in the dark with constant
FT              gyratory shaking (100 rpm) in Vogel's medium containing
FT              0.5% isolated rice cell walls as the sole carbon source"
FT              /organism="Magnaporthe grisea"
FT              /strain="CP987"
FT              /clone="RCW1"
FT              /clone_lib="RCW Lambda Zap Express Library"
FT              /tissue_type="Mycelium"
FT              /dev_stage="Day 5 post-inoculation"
XX
SQ      Sequence 337 BP; 56 A; 111 C; 74 G; 96 T; 0 other;
      ctttttcaat cagcccgaga actcctgggtt ggggttttctg cctgttctga cagctacttg      60
      tcatcgcata gcccgttctt tggttccaga taccacaagc ctgggacatt gatttcccag      120
      caactctttc aaaatggtat tattagcctc ctcacgatcc ctcgcgcggt cgcttggtcc      180
      ccttgcttca cgctctcgag cttttcagag cagtgctgct tcccgttccc tctcgactgc      240
      caccgctcgc ggccagggca aatccacaac tctcctgagg cccgcggccg ccacgaggac      300
      aagcagggttg ttgtcgactg gttcgcctt tcgtatt      337
//
```

```

<interpro id="IPR000002">
  <name>FIZZY/CDC20 domain</name>
  <type>Domain</type>
  <abstract> This domain is found in proteins ...</abstract>
  <examplelist>
    <example>
      <protein sptr_ac="Q12834" />Mammalian protein, p55CDC
    </example>
    <example>
      <protein sptr_ac="Q09649" />
    </example>
  </examplelist>
  <publist>
    <publication pub_id="PUB00006167">
      <authorlist>Shirayama M., Toth A., Galova M., Nasmyth K.</authorlist>
      <title>APC(Cdc20) promotes exit from mitosis by .....</title>
      <dbxref db="MEDLINE" dbkey="20110935" />
      <journal>Nature</journal>
      <location firstpage="203" lastpage="207" volume="402" />
      <year>1999</year>
    </publication>
  </publist>
  <memberlist>
    <dbxref db="PREFILE" dbkey="PS50218" name="FIZZY_DOMAIN" />
    <dbxref db="PRODOM" dbkey="PD004563" name="PD004563" />
  </memberlist>
</interpro>

```

```

<!ELEMENT interpro (name, type, examplelist, memberlist, publist, parlist*, chlist*,
seclist*, abstract)>

<!ELEMENT name (#PCDATA)>
<!ELEMENT type (#PCDATA)>
<!ELEMENT abstract (#PCDATA|cite|dbxref|sub|sup|p|li|i|ol|reaction|pre)*>
<!ELEMENT examplelist (example*)>
<!ELEMENT example (#PCDATA|protein|dbxref|cite)*>
<!ELEMENT publist (publication*)>
<!ELEMENT memberlist (dbxref*)>
<!ELEMENT protein (#PCDATA|protein)*>

<!ATTLIST interpro
  id          ID          #REQUIRED>
<!ATTLIST dbxref
  db          CDATA #IMPLIED
  dbkey       CDATA #IMPLIED
  name        CDATA #IMPLIED>
<!ATTLIST protein
  sptr_ac     CDATA #REQUIRED
  status      (?|T|P|F|N) #IMPLIED
  start       CDATA #IMPLIED
  end         CDATA #IMPLIED>

```

Linguaggi XML in biologia

- Sequenze
 - Bioinformatic Sequence Markup Language (BSML)
 - Agave
- Proteine (SPML)
- NCBI outputs (BlastXML)
- Microarray (MAGE-ML)
- Systems Biology Markup Language (SBML)
- Variabilità individuale
 - Biological Variation Markup Language (BVML)

Web Services

"A Web service is a software system designed to support **interoperable machine-to-machine interaction over a network**. It has an **interface described in a machine-processable format** (specifically WSDL). Other systems interact with the Web service in a manner prescribed by its description using SOAP-messages, typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards. “

W3C (working group) definition

Web Services

- Interfacce programmatiche (API) per l'accesso a servizi telematici basate su XML e protocolli standard di trasporto (HTTP, SOAP)
- Consentono alle applicazioni di accedere ai dati in maniera intelligente: individuazione dei contenuti e comprensione semantica
- Sono disponibili standard per la loro descrizione (WSDL), identificazione (UDDI) e composizione (WSFL)
- Utilizzano dati identificativi e descrittivi delle informazioni e dei servizi offerti (metadata)
- Già disponibili per molti servizi e banche date

Web Services in bioinformatics

- EMBOSS, XEMBL, Interpro (EBI)
- eUtils (NCBI)
- caBIO (NCICB)
- KEGG API

- GeneCruiser, Biosphere
- SIMAP, ...

- CABRI, TP53 (IST, Genova)

WSDL: chi fa cosa

Web Services Description Language (WSDL)

- Standard per la descrizione dei Web Services
- Comprende localizzazione, modalità di accesso e descrizione dettagliata
- Funzionalità astratte e dettagli per l'accesso concreti
- WSDL Binding: implementazione di WSDL per SOAP, HTTP, MIME

```
<definitions name="XEMBL" targetNamespace=http://www.ebi.ac.uk/XEMBL
  xmlns:tns=http://www.ebi.ac.uk/XEMBL
  xmlns:xsd=http://www.w3.org/2001/XMLSchema
  xmlns:soap=http://schemas.xmlsoap.org/wsdl/soap/
  xmlns:soapenc=http://schemas.xmlsoap.org/soap/encoding/
  xmlns="http://schemas.xmlsoap.org/wsdl/"
  <documentation>Documentation of this Web Service, together with a
    sample client and links to Bsm1 (Labbook, Inc.) and AGAVE
    (DoubleTwist, Inc.) can be found at the European Bioinformatics
    Institute http://www.ebi.ac.uk/xembl/</documentation>
  + <message name="getNucSeqRequest" xmlns:tns="http://www.ebi.ac.uk/XEMBL">
  + <message name="getNucSeqResponse">
  + <portType name="XEMBLPortType">
  + <binding name="XEMBLServiceBinding" type="tns:XEMBLPortType">
  + <service name="XEMBLService">
</definitions>
```

WSDL: XEMBL.wsdl <service> </service>

```
<service name="XEMBLService">
  <documentation>Returns full information on EMBL Nucleotide Sequences
  formatted as Bsm1 XML or Agave XML. I.e. returns sequence itself,
  cross-references, taxonomy, literature, full feature information,
  etc.</documentation>
  <port name="XEMBLPort" binding="tns:XEMBLServiceBinding">
    <soap:address
      location="http://www.ebi.ac.uk:80/cgi-bin/xembl/XEMBL-SOAP.pl" />
  </port>
</service>
```

WSDL: XEMBL.wsdl <binding> </binding>

```
<binding name="XEMBLServiceBinding" type="tns:XEMBLPortType">
  <soap:binding style="rpc"
    transport="http://schemas.xmlsoap.org/soap/http" />
  <operation name="getNucSeq">
    <soap:operation soapAction="http://www.ebi.ac.uk/XEMBL#getNucSeq" />
    <input>
      <soap:body use="encoded" namespace=http://www.ebi.ac.uk/XEMBL
        encodingStyle="http://schemas.xmlsoap.org/soap/encoding/" />
    </input>
    <output>
      <soap:body use="encoded" namespace=http://www.ebi.ac.uk/XEMBL
        encodingStyle="http://schemas.xmlsoap.org/soap/encoding/" />
    </output>
  </operation>
</binding>
```

WSDL: XEMBL.wsdl

<portType>

</portType>

```
<portType name="XEMBLPortType">
  <operation name="getNucSeq">
    <input message="tns:getNucSeqRequest" name="getNucSeq" />
    <output message="tns:getNucSeqResponse" name="getNucSeqResponse" />
  </operation>
</portType>
```

WSDL: XEMBL.wsdl

<message></message>

```
<message name="getNucSeqRequest" xmlns:tns="http://www.ebi.ac.uk/XEMBL">
  <part name="format" type="xsd:string">
    <documentation>Input parameter that indicates the result format that
      should be returned. Legit values: Bsml or sciobj.
      Defaults to Bsml if format not recognised.
    </documentation>
  </part>
  <part name="ids" type="xsd:string">
    <documentation>A space delimited list of international Nucleotide
      Sequence accession numbers (IDs).
      For example: "HSERPG U83300 AC000057".
      Minimum number of IDs is 1.
    </documentation>
  </part>
</message>

<message name="getNucSeqResponse">
  <part name="result" type="xsd:string">
    <documentation>An XML formatted result in either Bsml or AGAVE
      format.</documentation>
  </part>
</message>
```

Comporre servizi complessi

Web Services Flow Language (WSFL)

- Consente di descrivere insiemi di web services
- Flow models: specifica come utilizzare un insieme di web services per raggiungere un certo obiettivo
- Global models: descrive le interazioni tra più web services e il comportamento globale di un insieme
- Recursive composition: ogni flow model o global model viene considerato come un nuovo web service e rientrare in altri modelli
- Esistono alternative:
 - WSCDL (Web Services Choreography Description Language)
 - Scufi (Simple Conceptual Unified Flow Language)

Registrazione dei servizi

Universal Description, Discovery and Integration (UDDI)

- Realizzazione di un framework per la descrizione dei Web Services, indipendente da HW e SW
- Consente la creazione di registri di Web Services
- Basato su standard World Wide Web Consortium (W3C) and Internet Engineering Task Force (IETF)
- Esistono alternative: bioMOBY

“Middleware is not magic” (C. Goble)

- Qualità dei dati
- Gestione del contenuto dei db (vocabolari controllati)
- Creazione e aggiornamento delle ontologie
- Gestione delle informazioni sull'origine dei dati e la release
- Strumenti appropriati, da usare in maniera appropriata
- Difficoltà nell'accedere all'annotazione a testo libero

CABRI: obiettivi

Common Access to Biological Resources and Information

- Distribuzione di materiali biologici di qualità
- Linee Guida per la conservazione del materiale
- Centro Risorse Biologiche virtuale
- Cataloghi integrati tramite SRS
- Integrazione con db esterni
- Shopping cart

CABRI: partners e materiali

Partners:

- BCCM, CABI, CBS, CIP, DSMZ, ICLC, NCCB, NCIMB (collezioni)
- IST, CERDIC (ITC)

Materiali:

- Microrganismi (Batteri, lieviti, funghi filiformi)
- Linee cellulari animali e umane, ibridomi, linee B tip. HLA
- Plasmidi, fagi, virus, sonde DNA
- Complessivamente più di 100.000 risorse

CABRI: ricerca

CABRI Simple Search:

- Ricerca per ID, nome, tutti i campi restanti
- Ricerca per sinonimi

SRS standard query form:

- Utilizzo appieno link SRS
- Gestione viste personalizzate
- Selezione singoli field
- Accesso indici SRS

CABRI HyperCatalogue:

- Indici statici per materiale e per catalogo
- Accesso finale alla descrizione dettagliata tramite SRS

IARC TP53 database

IARC TP53 Mutation Database

<http://www-p53.iarc.fr/>

- Release 9: 19,809 mutazioni somatiche, 1,769 articoli
- Informazioni: mutazione, materiale, stile di vita
- Vocabolari e annotazioni standard
- Ricerche on-line richiedono interazione

Implementazione SRS del database IARC TP53

<http://srs.o2i.it/srs71/>

- Basato su SRS
- Definizione di un DTD ad-hoc
- Trasferimento dati basato su XML
- Accesso programmatico semplificato

Web Services per CABRI e TP53

Implementare Web Services che consentano:

- L'accesso ai database CABRI e TP53 tramite i siti SRS
- La possibilità di includere questi task in workflow complessi

Riproducendo il comportamento attuale:

- Ricerca per nome, identificatore e a testo libero (CABRI)
- Ricerca per funzioni e proprietà (TP53)
- Combinare risultati
- Integrare i dati con altri sorgenti tramite IDs o termini comuni

Due tipologie di Service:

- Ricerca per una specifica proprietà e restituzione degli ID
- Ricerca per ID e restituzione del record completo

Soaplab: SOAP-based Analysis Web Service

“Soaplab is a set of Web Services providing a programatic access to some applications on remote computers. It is often referred to as an **Analysis (Web) Service**” (Martin Senger, EBI).

Consente di implementare Web Services in grado di accedere a:

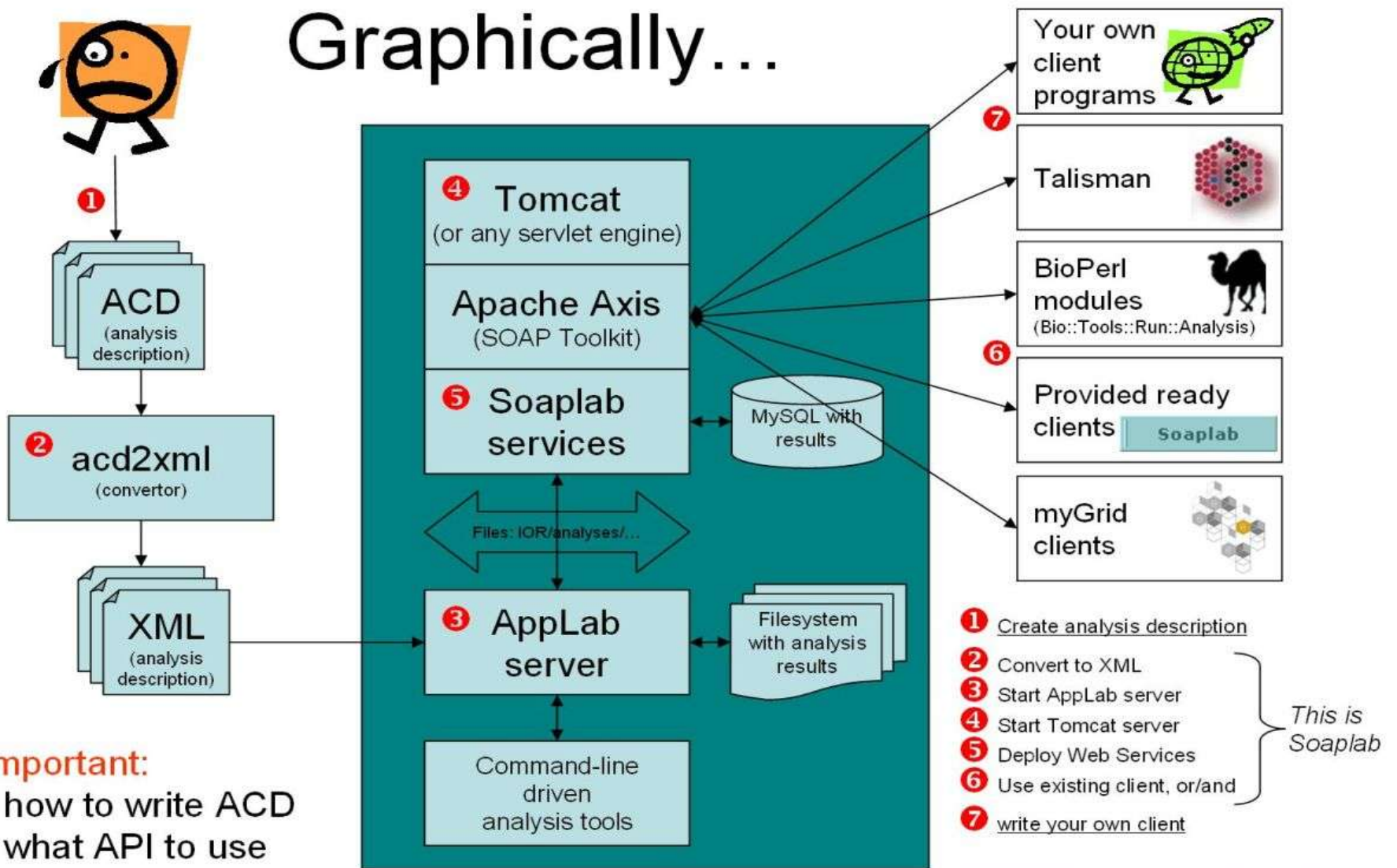
- Applicazioni locali eseguibile con “command-line”
- Applicazioni presenti in EMBOSS
- Contenuto di qualunque pagina Web (GowLab)

Requisiti

- Apache Tomcat servlet engine, Axis SOAP toolkit, Java
- perl, MySQL



Graphically...



Important:

- how to write ACD
- what API to use by the clients

Download: [soaplab-clients-version](#)

Download: [applab-cembalo-soaplab-version](#) (for EMBOSS)

Download: [analysis-interfaces-version](#)

<http://industry.ebi.ac.uk/soaplab/dist/>

Binary
distribution

Soaplab: file ACD

```
appl: getCellLineIdsByName [  
  documentation: "Get cell lines by name from CABRI human and  
    animal cell lines catalogues (see www.cabri.org)"  
  groups: "CABRI"  
  nonemboss: "Y"  
  comment: "launcher get"  
  supplier: "http://www.cabri.org/CABRI/srs-bin/wgetz"  
  comment: "method [{libs}-nam:'$name'] -ascii"  
]  
  
string: libs [ parameter: "Y" ]  
string: name [ parameter: "Y" ]  
  
outfile: result [ ]
```

Soaplab: file ACD

```
appl: getCellLineIdsByProperty [  
  documentation: "Get cell lines by properties (all text) from CABRI  
    human and animal cell lines catalogues (see www.cabri.org)"  
  groups: "CABRI"  
  nonemboss: "Y"  
  comment: "launcher get"  
  supplier: "http://www.cabri.org/CABRI/srs-bin/wgetz"  
  comment: "method [{libs}-all:'$text'] -ascii"  
]
```

```
string: libs [ parameter: "Y" ]  
string: text [ parameter: "Y" ]
```

```
outfile: ids [ ]
```

Soaplab: file ACD

```
appl: getCellLinesById [  
  documentation: "Get cell lines by Id from CABRI human and  
    animal cell lines catalogues (see www.cabri.org)"  
  groups: "CABRI"  
  nonemboss: "Y"  
  comment: "launcher get"  
  supplier: "http://www.cabri.org/CABRI/srs-bin/wgetz"  
  comment: "method -e [{libs}: '$id'] -ascii"  
]
```

```
string: libs [ parameter: "Y" ]  
string: id [ parameter: "Y" ]
```

```
outfile: result [ ]
```

Web Services CABRI

Web Service Name	Involved catalogues	Input	Output
getBacteriaIdsByName	Bacteria strains	lib(s), name	id(s)
getBacteriaIdsByProperty	Bacteria strains	lib(s), text	id(s)
getBacteriaById	Bacteria strains	id	full record
getFungiIdsByName	Filamentous fungistrains	lib(s), name	id(s)
getFungiIdsByProperty	Filamentous fungi strains	lib(s), text	id(s)
getFungiById	Filamentous fungi strains	id	full record
getYeastIdsByName	Yeasts strains	lib(s), name	id(s)
getYeastIdsByProperty	Yeasts strains	lib(s), text	id(s)
getYeastsById	Yeasts strains	id	full record
getPlasmidIdsByName	Plasmids	lib(s), name	id(s)
getPlasmidIdsByProperty	Plasmids	lib(s), text	id(s)
getPlasmidsById	Plasmids	id	full record
getPhageIdsByName	Phages	lib(s), name	id(s)
getPhageIdsByProperty	Phages	lib(s), text	id(s)
getPhagesById	Phages	id	full record
getCellLinesIdsByName	Human and animal cell lines	lib(s), name	id(s)
getCellLinesIdsByProperty	Human and animal cell lines	lib(s), text	id(s)
getCellLinesById	Human and animal cell lines	id	full record
getResourceIdsByName	All	lib(s), name	id(s)
getResourcesById	All	id	full record

Web Services TP53

Web Service Name	Input	Output
getP53MutationsByProperty	lib, text	Full record
getP53MutationsByIds	Id	Full record
getP53MutationIdsByType	lib, mutation type	id(s)
getP53MutationIdsByEffect	lib, effect	id(s)
getP53MutationIdsByExon	lib, exon number	id(s)
getP53MutationIdsByIntron	lib, intron number	id(s)
getP53MutationIdsByCodonNumber	lib, codon number	id(s)
getP53MutationIdsByCpgSite	lib, cpg site (true/false)	id(s)
getP53MutationIdsBySpliceSite	lib splice site (true/false)	id(s)
getP53MutationIdsByMetastasisLocalization	lib, metastasis localization (organ)	id(s)
getP53MutationIdsByTumorOrigin	lib, origin (primary, secondary, ...)	id(s)

Workflows management

“A computerized facilitation or automation of a business process, in whole or part”. (Workflow Management Coalition)

Obiettivo:

- implementazione di processi di analisi dei dati in ambienti standardizzati

Vantaggi principali:

- **efficienza:** in quanto procedura automatica, libera il ricercatore dai compiti ripetitivi sul web e contribuisce a una “good practice”,
- **reproducibilità:** le analisi possono essere ripetute nel tempo,
- **riuso:** I risultati intermedi possono essere riutilizzati,
- **tracciabilità:** il workflow è eseguito in un ambiente trasparente nel quale la provenienza dei dati può essere verificata a posteriori.

Workflow management software

Gestione di workflow per applicazioni bioinformatiche:

- Biopipe, un add-on per bioperl
- GPipe, una estensione dell'interfaccia Pise
- Taverna (EBI), una componente della piattaforma myGrid
- Pegasys (University of British Columbia)
- EGene (Universidade de São Paulo)
- Wildfire (Bioinformatics Institute, Singapore)
- Pipeline Pilot (SciTegic)
- BioWBI, Bioinformatic Workflow Builder Interface, di IBM

Richiedono una notevole conoscenza dei sistemi coinvolti e competenze e tempo per lo sviluppo dei workflow.

Workflow management: Taverna

Taverna Workbench consente di

- costruire workflow per analisi complesse
- accedere a processori sia remoti che locali
- definire processori alternativi
- eseguire I workflow
- visualizzare i risultati in diversi formati
- descrivere i dati bioinformatici tramite un'ontologia

Requirements: java, Windows or Linux

Open source: <http://taverna.sourceforge.net/>

Taverna: processori disponibili

Web Services descritti tramite WSDL

- Mette a disposizione i servizi descritti

Web Services accessibili tramite server Soaplab

- Mette a disposizione i servizi pubblicizzati dal server Soaplab

Registri BioMOBY

- Interagisce con un MOBY Centrale repository per accedere ai servizi registrati

Workflow

- Incorpora interi workflow definiti con Scufi o aggiunge singoli processori

Biomart

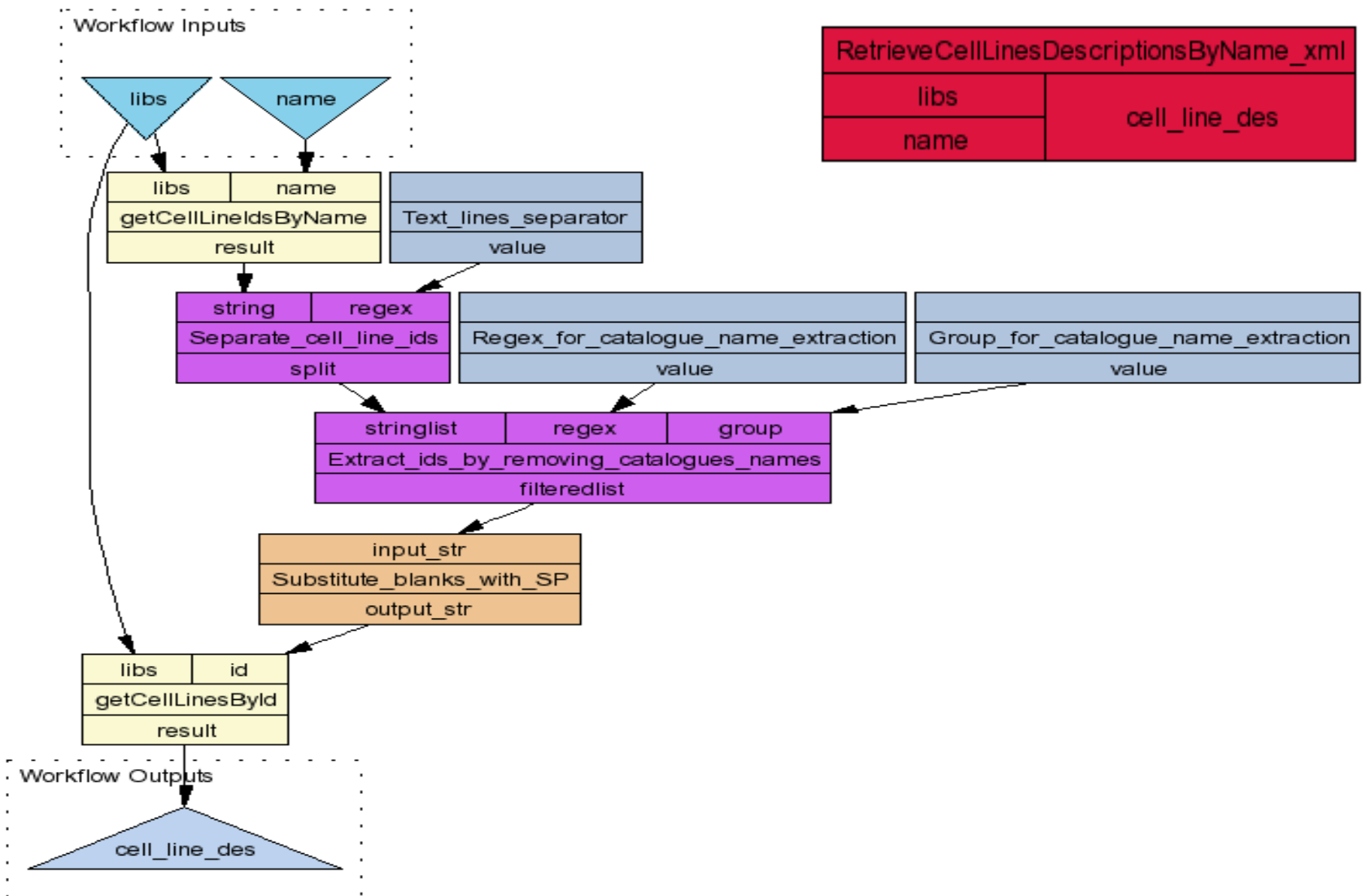
- Interagisce con database Biomart per comporre query

SeqHound

- Interroga SeqHound (Sequence and Structure Database Management System), un insieme di Web Services basati sul modello dati NCBI e accede a informazioni di sequenza e struttura

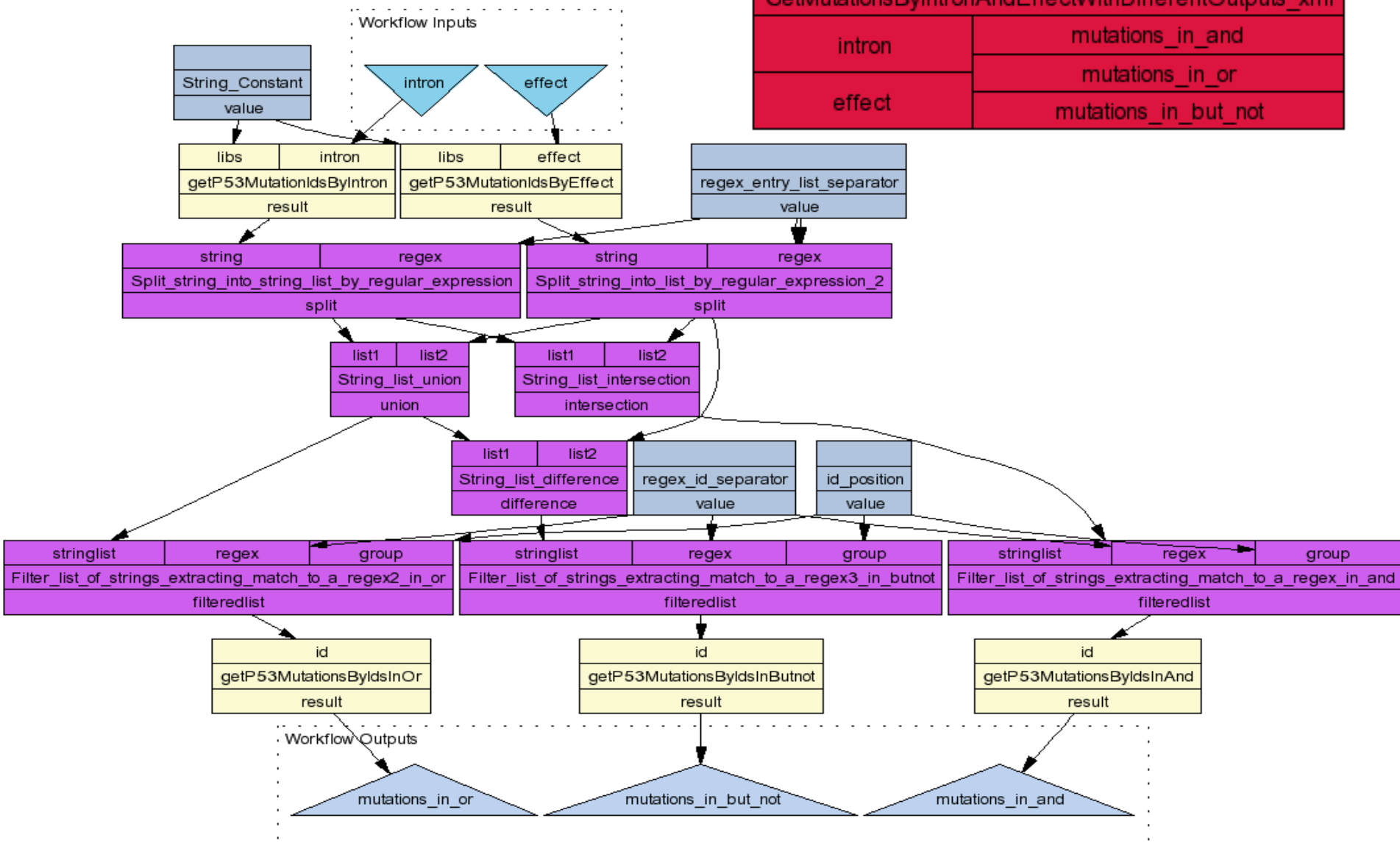
Processori locali

- Funzioni Java in grado di elaborare liste e stringhe, definire valori costanti, eseguire semplici elaborazioni. È anche possibile scrivere propri script con beanshell



Workflow per database CABRI

GetMutationsByIntronAndEffectWithDifferentOutputs_xml		
intron	mutations_in_and	
effect	mutations_in_or	
	mutations_in_but_not	



Workflow per database TP53

Oncology over Internet (O₂I)

Obiettivi del sistema:

- Consentire l'esecuzione di un insieme predefinito di workflow (di interesse oncologico)
- Consentire l'annotazione dei workflow tramite una semplice ontologia dei processori bioinformatici (dominio, task, i/o)
- Consentire la ricerca di workflow sulla base della loro annotazione
- Consentire la selezione di workflow sulla base della tipologia dell'utente (ruolo, interessi)
- Consentire la memorizzazione e il recupero dei workflow eseguiti e dei relativi risultati

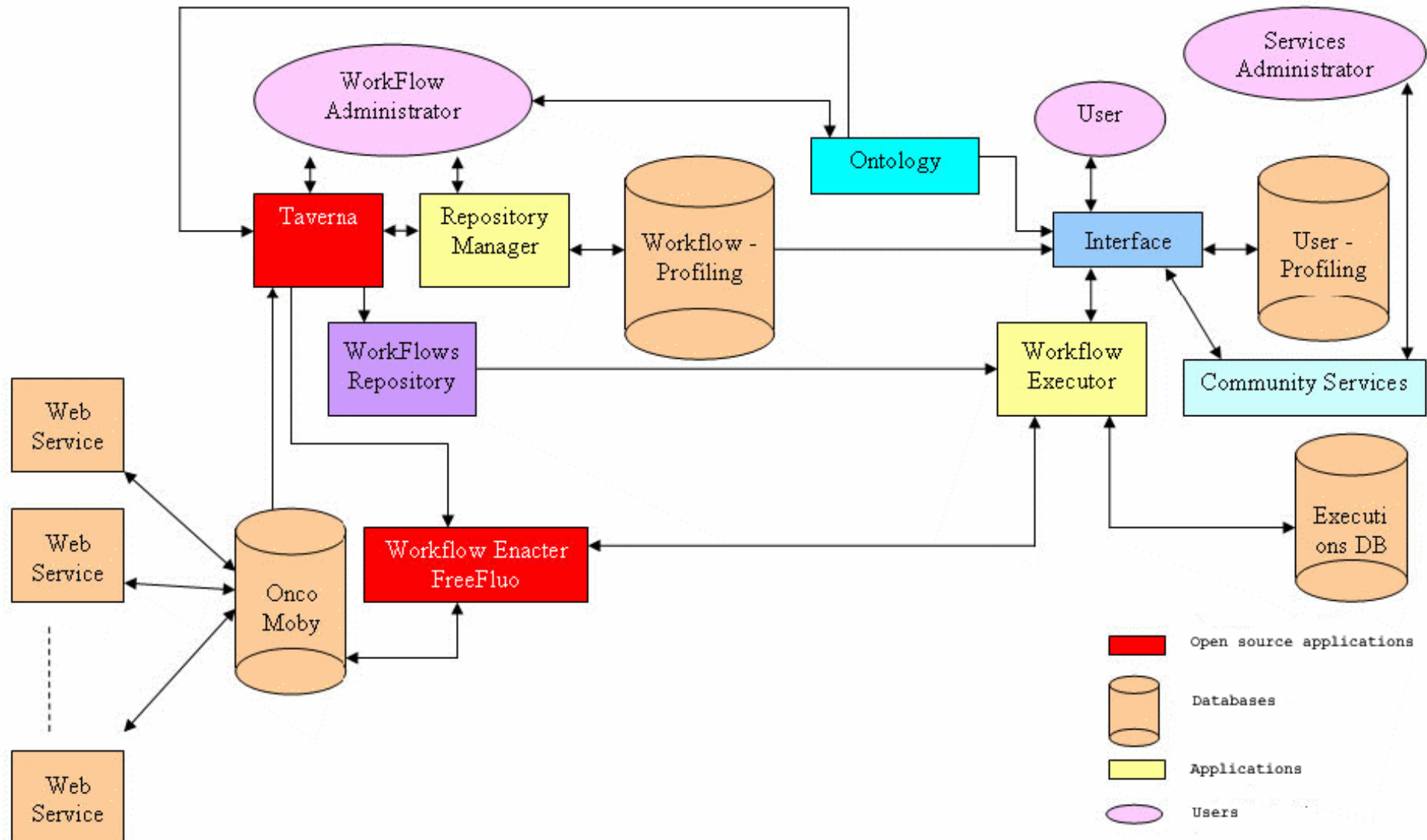
Oncology over Internet (O₂I)

Il sistema:

- Accede e interroga/utilizza Web Services, anche tramite registri bioMOBY
- Archivia i workflow in formato Scufi
- È parzialmente basato su strumenti open source (Taverna WB, FreeFluo e MySQL)

Prototipo disponibile on-line entro fine 2005:
<http://www.o2i.it:8080/portal/>

Architettura O₂I



Workflow predefiniti

I workflow sono:

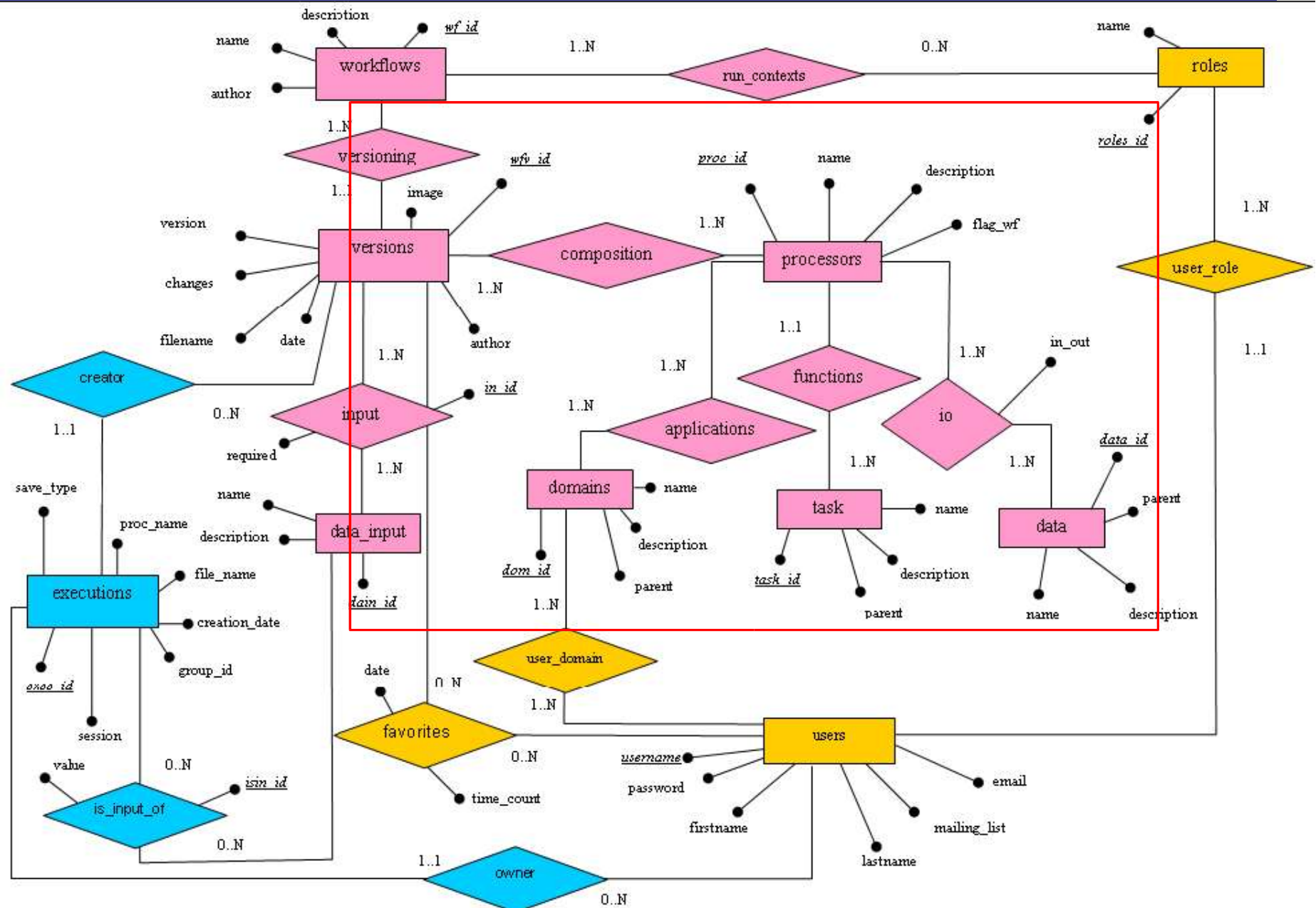
- creati dall'amministratore con Taverna
- archiviati in formato Scufi
- aggiornati (workflow vs versione)
- sottomessi da:
 - utenti
 - service providers

Annotazione dei workflow

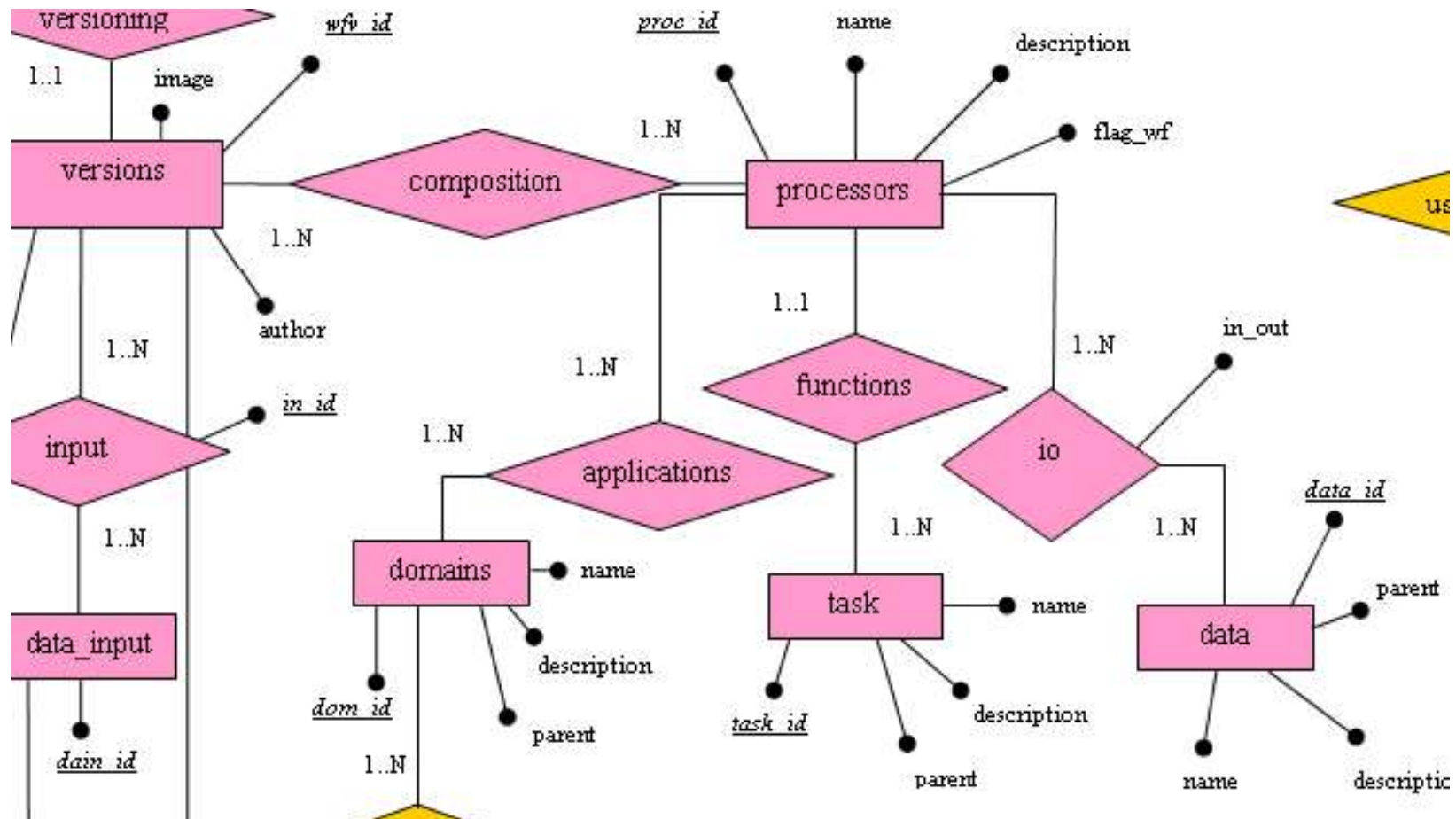
I workflow sono annotati sulla base di:

- una semplice ontologia per processori bioinformatici:
 - dominio applicativo
 - task
 - input e output
- l'ontologia deriva da quella di Taverna:
 - nuova struttura
 - alcune aggiunte (biological resources, images, ...)
 - in via di ulteriore sviluppo e adattamento

Schema ERA



Annotazione dei workflow



Registrazione e profilo utenti

Gli utenti sono raggruppati in base a:

- ruolo ricoperto
 - informatico / medico / ricercatore / paziente / giornalista / ...
- dominio di interesse

Sono conservate le precedenti esecuzioni di workflow

02I Project - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo  http://www.o2i.it:8080/o2i/  Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

Please login

username:

password:

New user? Please [register!](#)

Please install the **new** [library](#). ([Instructions](#))
Applets digital [certificate](#).

O₂I project

O2I Project - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo http://www.o2i.it:8080/o2i/change_profile.jsp Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

USER: PaoloR Clone Window logout

All workflows list

My last executed

My domains workflows

My role most popular

My role last executed

Search by ontology

All available results

Unsaved results

Temporary saved results

Persistently saved results

Edit your profile

Username: PaoloR

Password: (Must be six characters or more)

Re-type Password:

First Name:

Last Name:

Role:

computer scientist - bioinformatician
computer scientist - unspecified
journalist
patient
physician (clinical oncologist)

E-Mail Address:

Domain:

Regulatory affairs and policies
Research
Scientific divulgation
Translational research
Virology

Do you want to receive news through your e-mail address?
☒ Yes ☐ No

Choosing your password
You will use this information to access O₂I each time. Capitalization matters for your password!

Info about you
Please give us some information about you.

O2I Project - Microsoft Internet Explorer

FileModificaVisualizzaPreferitiStrumenti?

Indirizzohttp://www.o2i.it:8080/o2i/main.jsp?modo=0

Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

USER: PaoloR

Clone Windowlogout

All workflows list

My last executed

My domains workflows

My role most popular

My role last executed

Search by ontology

All available results

Unsaved results

Temporary saved results

Persistently saved results

Edit your profile

All workflows:

Workflow			Description	Version
Conditional Branch Choice	details	run	This is a demo workflow distributed with Taverna Workbench (see taverna site). If the input is true then the string 'foo' is emitted, if false then 'bar'. Just a simple example to show how the conditional branch processor works.	1.0
Retrieve Cell Lines Descriptions By Name	details	run	This workflow takes the cell line name and the catalogue(s) name(s) as input and retrieve the full cell line description(s) by first retrieving the cell lines' unique IDs associated with the input (done via a call to the <code>getCellLineIdsByName</code> web service) and then using IDs for retrieving the full cell lines descriptions (done via a call to the <code>getCellLinesByIds</code> web service). Both these web services are available at the soaplab system at http://www.o2i.it:8080/axis/services A number of string or string list local elaborations are required: - returned IDs are in a string and this must be transformed in a list (done by the 'Separate_cell_line_ids' processor, that is implemented by using a <code>Split_string_into_string_list_by_regular_expression</code> local processor) - returned IDs include catalogues' names and this must be removed before their utilization for further processing (done by the 'Extract_ids_by_removing_catalogues_names' processor, that is implemented by using a <code>Filter_list_of_strings_extracting_match_to_a_regex</code> local processor) - returned IDs include a blank character and this must be substituted by a '_SP_' characters string before submitting the data to the 'getCellLinesByIds' web service (done by a trivial beanshell script). Special requirements on input data are: - one or more of the following catalogues names can be specified: 'iclc', 'ecacc_cell', 'dsmz_mutz'. Other names may lead to errors, - when specifying more than one catalogue names, they must be in a unique input string but on distinct text lines, - cell lines names can only be made by a single word, excluding special characters as '/', '-' and '*', - cell lines names are case insensitive. Example of valid cell lines names are: - vero - hela - a172 - calu6	1.0
Retrieve decriptions of	details	run	Retrieve full descriptions of bacteria strains from CABRI catalogues (see CABRI site) by their scientific name (genus and species only). Inputs of the workflow are the name of the involved CABRI catalogues (text/plain string with one catalogue name per line) and the scientific name of the desired bacteria strain (a text/plain string including genus and species separated by a blank space).	1.0

O2I Project - Microsoft Internet Explorer

FileModificaVisualizzaPreferitiStrumenti?

Indirizzohttp://www.o2i.it:8080/o2i/index.jspVai



O₂I (Oncology over Internet) Project
Your personalized project research web site.

USER: PaoloR

Clone Windowlogout

All workflows list

My last executed

My domains workflows

My role most popular

My role last executed

Search by ontology

All available results

Unsaved results

Temporary saved results

Persistently saved results

Edit your profile

My role most popular workflows:

Workflow	Description	Version	Executed
No workflows corresponds to your selected criteria.			

O₂I project

O2I Project - Microsoft Internet Explorer
File Modifica Visualizza Preferiti Strumenti ?
Indirizzo http://www.o2i.it:8080/o2i/main.jsp?modo=4
Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

USER: PaoloR

Clone Window logout

All workflows list
My last executed
My domains workflows
My role most popular
My role last executed
Search by ontology

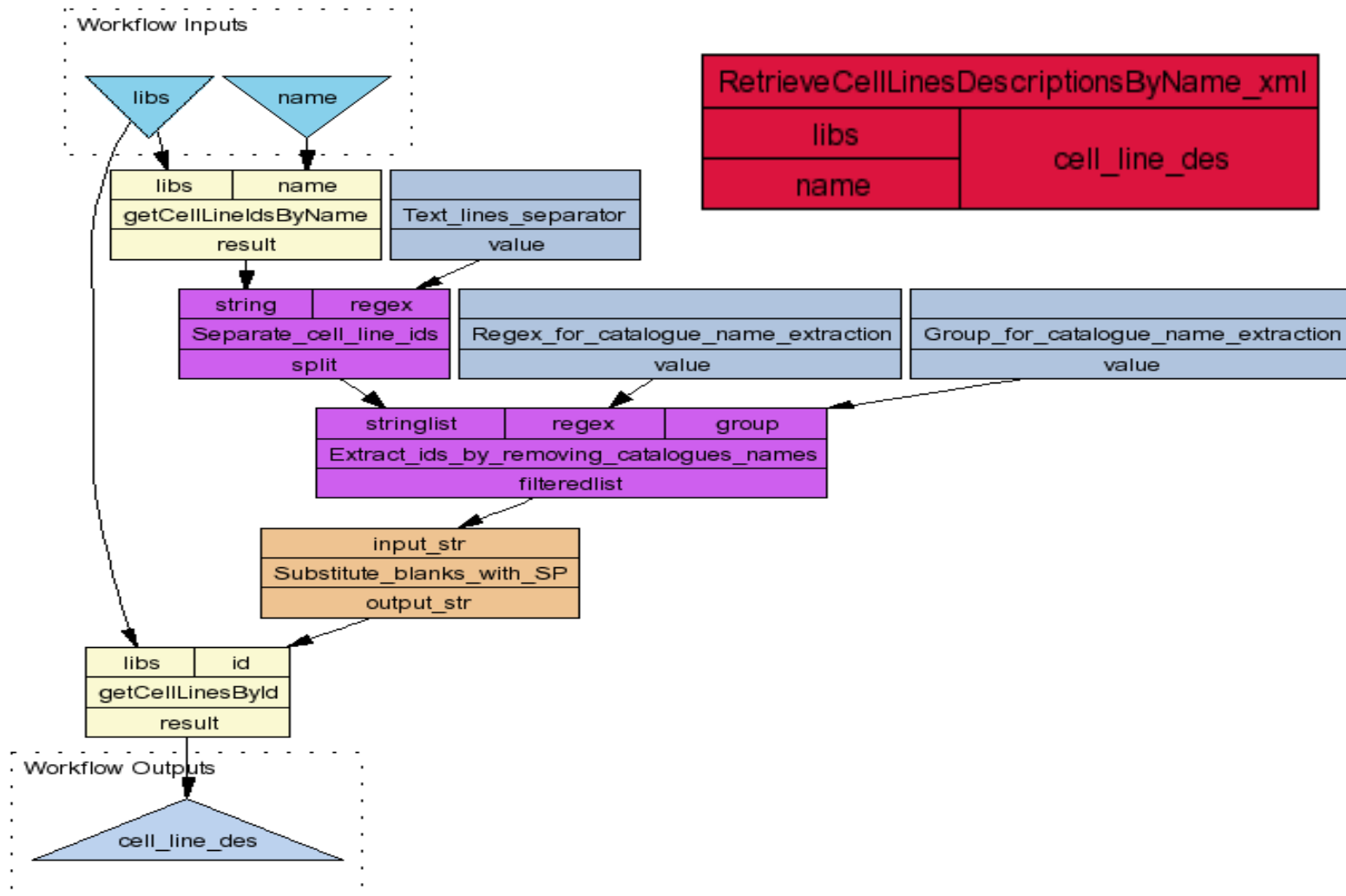
All available results
Unsaved results
Temporary saved results
Persistently saved results

Edit your profile

My application domains workflows:

Workflow	Description	Version
Retrieve Cell Lines Descriptions By Name details run	This workflow takes the cell line name and the catalogue(s) name(s) as input and retrieve the full cell line description(s) by first retrieving the cell lines' unique IDs associated with the input (done via a call to the <code>getCellLineIdsByName</code> web service) and then using IDs for retrieving the full cell lines descriptions (done via a call to the <code>getCellLinesByIds</code> web service). Both these web services are available at the soaplab system at http://www.o2i.it:8080/axis/services A number of string or string list local elaborations are required: - returned IDs are in a string and this must be transformed in a list (done by the 'Separate_cell_line_ids' processor, that is implemented by using a <code>Split_string_into_string_list_by_regular_expression</code> local processor) - returned IDs include catalogues' names and this must be removed before their utilization for further processing (done by the 'Extract_ids_by_removing_catalogues_names' processor, that is implemented by using a <code>Filter_list_of_strings_extracting_match_to_a_regex</code> local processor) - returned IDs include a blank character and this must be substituted by a 'SP' characters string before submitting the data to the 'getCellLinesByIds' web service (done by a trivial beanshell script). Special requirements on input data are: - one or more of the following catalogues names can be specified: 'iclc', 'ecacc_cell', 'dsmz_mutz'. Other names may lead to errors, - when specifying more than one catalogue names, they must be in a unique input string but on distinct text lines, - cell lines names can only be made by a single word, excluding special characters as '/', '-' and '*', - cell lines names are case insensitive. Example of valid cell lines names are: - vero - hela - a172 - calu6	1.0
Retrieve decriptions of bacteria strains details run	Retrieve full descriptions of bacteria strains from CABRI catalogues (see CABRI site) by their scientific name (genus and species only). Inputs of the workflow are the name of the involved CABRI catalogues (text/plain string with one catalogue name per line) and the scientific name of the desired bacteria strain (a text/plain string including genus and species separated by a blank space). Data is retrieved from CABRI Web Services in two steps. First, all bacteria strains IDs are retrieved by using the <code>getBacteriaIdsByName</code> method, and after descriptions are retrieved by using the <code>getBacteriaById</code> method. Some list/text elaboration is required to remove catalogue names from returned IDs	1.0

Simple demo workflow



O2I Project - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo http://www.o2i.it:8080/o2i/details.jsp?wfv_id=2 Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

USER: PaoloR Clone Window logout

[All workflows list](#)[My last executed](#)[My domains workflows](#)[My role most popular](#)[My role last executed](#)[Search by ontology](#)[All available results](#)[Unsaved results](#)[Temporary saved results](#)[Persistently saved results](#)[Edit your profile](#)

Workflows details

Name: Retrieve Cell Lines Descriptions By Name

Description: This workflow takes the cell line name and the catalogue(s) name(s) as input and retrieve the full cell line description(s) by first retrieving the cell lines' unique IDs associated with the input (done via a call to the `getCellLineIdsByName` web service) and then using IDs for retrieving the full cell lines descriptions (done via a call to the `getCellLinesByIds` web service). Both these web services are available at the soaplab system at <http://www.o2i.it:8080/axis/services>. A number of string or string list local elaborations are required: - returned IDs are in a string and this must be transformed in a list (done by the 'Separate_cell_line_ids' processor, that is implemented by using a `Split_string_into_string_list_by_regular_expression` local processor) - returned IDs include catalogues' names and this must be removed before their utilization for further processing (done by the 'Extract_ids_by_removing_catalogues_names' processor, that is implemented by using a `Filter_list_of_strings_extracting_match_to_a_regex` local processor) - returned IDs include a blank character and this must be substituted by a '_SP_' characters string before submitting the data to the 'getCellLinesByIds' web service (done by a trivial beanshell script). Special requirements on input data are: - one or more of the following catalogues names can be specified: 'idcl', 'ecacc_cell', 'dsmz_mutz'. Other names may lead to errors, - when specifying more than one catalogue names, they must be in a unique input string but on distinct text lines, - cell lines names can only be made by a single word, excluding special characters as '/', '-' and '*', - cell lines names are case insensitive. Example of valid cell lines names are: - vero - hela - a172 - calu6

Author: Paolo Romano, IST, Genova, Italy (paolo.romano@istge.it)

Roles list:

- researcher - molecular biologist
- researcher - cellular biologist
- researcher - structural biologist
- researcher - microbiologist
- researcher - immunologist
- researcher - virologist

Version details

Version: 1.0

Date: 18:05 - 27/09/2005

02i Project - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo http://www.o2i.it:8080/o2i/details.jsp?wfv_id=2 Vai

- researcher - virologist

Version details

Version: 1.0

Date: 18:05 - 27/09/2005

Author: Paolo Romano, IST, Genova, Italy (paolo.romano@istge.it)

Filename: RetrieveCellLinesDescriptionsByName.xml

Changes: Original version

Image: [show](#) in a new window

Input list:

- biological resource database (required)
- cell line name (required)

Output list:

- CABRI human and animal cell lines record

Domains list:

- Microbiology
- Cellular Biology

Processors:
Please be advised that is only a list of main components of the workflow. They are not ordered.

Name	Description	Task	Domains	Inputs	Outputs
Get cell lines id by name	Retrieve CABRI cell lines' IDs after a search in CABRI Web Services by cell lines' name	biological resource retrieval	Microbiology Cellular Biology	biological resource database cell line name	biological resource identifier
Get cell lines descriptions by id	Retrieves cell lines' descriptions by CABRI id	biological resource retrieval	Microbiology Cellular Biology	biological resource database biological resource identifier	CABRI human and animal cell lines record

• researcher - virologist

Version details

Version: 1.0

Date: 18:05 - 27/09/2005

Author: Paolo Romano, IST, Genova, Ita

Filename: RetrieveCellLinesDescriptions

Changes: Original version

Image: [show](#) in a new window

Input list:

- biological resource database (requ
- cell line name (required)

Output list:

- CABRI human and animal cell lines

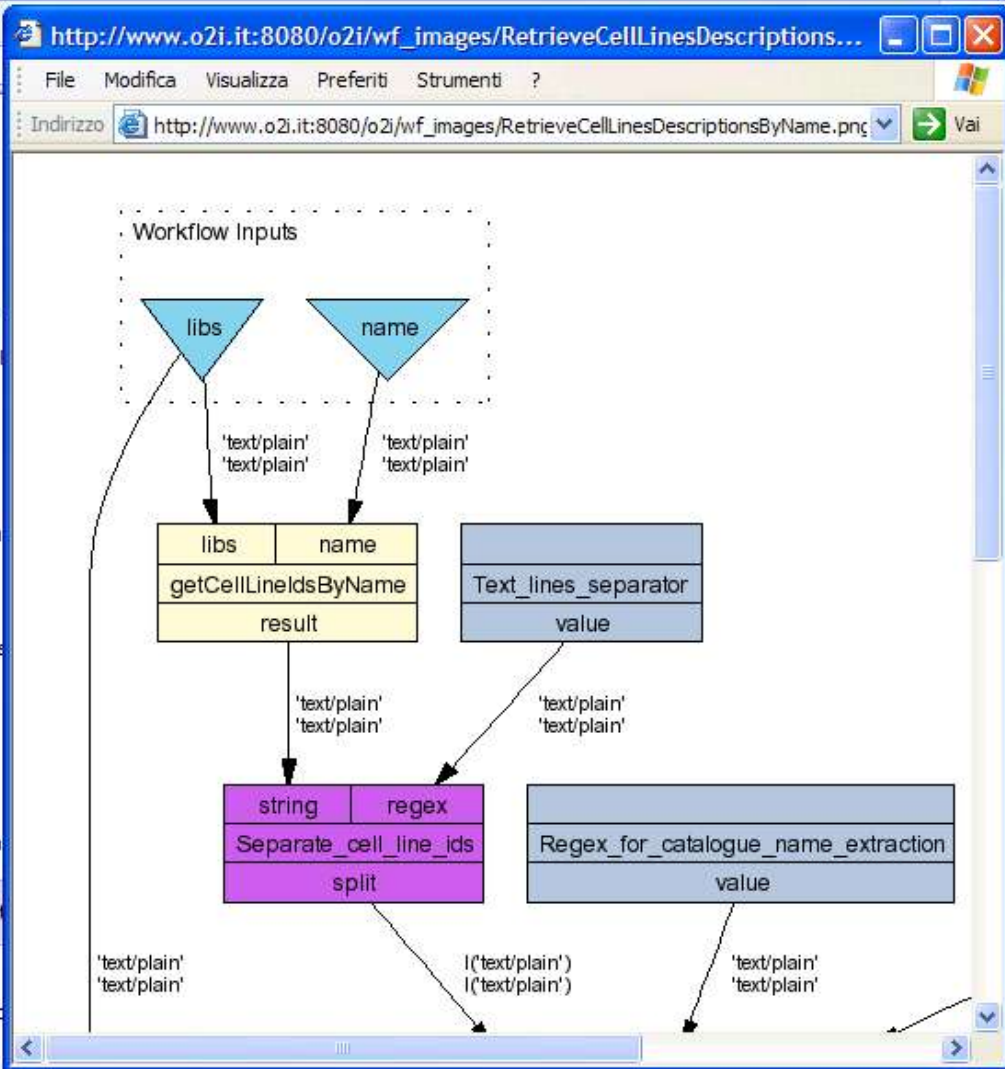
Domains list:

- Microbiology
- Cellular Biology

Processors:

Please be advised that is only a list of ma

Name	Description
Get cell lines id by name	Retrieve CABRI d after a search in Services by cell
Get cell lines descriptions by id	Retrieves cell lines' descriptions by CABRI id



O₂I Project - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo <http://www.o2i.it:8080/o2i/ontosearch.jsp> Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

USER: PaoloR Clone Window logout

[All workflows list](#)
[My last executed](#)
[My domains workflows](#)
[My role most popular](#)
[My role last executed](#)
[Search by ontology](#)

[All available results](#)
[Unsaved results](#)
[Temporary saved results](#)
[Persistently saved results](#)

[Edit your profile](#)

Search by ontology:

Domains	Tasks	Input	Output
- Scientific divulgation - Clinics - Regulatory affairs and polic - Information to patients and - Research	- aligning - retrieval - calculating - displaying - distinguishing - filtering - grouping - inserting - joining - manipulating - merging - narsing	- graph - database record - report - phylogenetic tree - sequence - diagram - metadata - database - biological image - name	- graph - database record - report - phylogenetic tree - sequence - diagram - metadata - database - biological image - name

search "Ctrl" key for multiple selections

O2I Project - Microsoft Internet Explorer

File

Modifica

Visualizza

Preferiti

Strumenti

?

Indirizzo

http://www.o2i.it:8080/o2i/ontosearch.jsp?dom_id=0&task_id=19&in_id=0&out_id=120

Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

USER: PaoloR

Clone Window

logout

All workflows list

My last executed

My domains workflows

My role most popular

My role last executed

Search by ontology

All available results

Unsaved results

Temporary saved results

Persistently saved results

Edit your profile

Search by ontology:

Domains

Scientific divulgation

Clinics

Regulatory affairs and polic

Information to patients and

Research

Tasks

aligning

retrieval

bibliografy retrieval

sequence retrieval

biological resource r

image retrieval

calculating

displaying

distinguishing

filtering

grouping

Input

graph

database record

report

phylogenetic tree

sequence

diagram

metadata

database

biological image

name

OMIM record

PDB record

SWISS-PROT record

UniGene record

CABRI record

CABRI human ar

CABRI bacteria s

CABRI filamentou

CABRI plasmid r

CABRI phage rec

CABRI yeast stra

CABRI plant cell

search

"Ctrl" key for multiple selections

Results List:

Name		Description	w/p
Workflow Retrieve decriptions of bacteria strains	details run	This process describes the overall elaboration of the workflow "Retrieve decriptions of bacteria strains"	w
Get bacteria strains descriptions by id	linked workflows	Retrieves bacteria strains descriptions by CABRI id	p

Romano, XML e Web Services

IST

71

O2I - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo http://www.o2i.it:8080/o2i/input_insert.jsp?wfv_id=2 Vai



O₂I (Oncology over Internet) Project

Your personalized project research web site.

USER: PaoloR Clone Window logout

All workflows list

My last executed

My domains workflows

My role most popular

My role last executed

Search by ontology

All available results

Unsaved results

Temporary saved results

Persistently saved results

Edit your profile

Please insert input:

CABRI Cell lines catalogues: (required input)

Description: This input includes the name(s) of the CABRI human and animal cell lines catalogues involved in the query. Multiple values can be specified, in a unique string field, each name in a distinct text line (thus, names must be divided by a '\n' character).
As of Sep 15, 2005, possible values are:

- 'iclc' (i.e., the Interlab Cell Line Collection, <http://www.iclc.it/>)
- 'ecacc_cell' (i.e., the European Collection of Cell Cultures, <http://www.ecacc.org.uk/>)
- 'dsmz_mutz' (i.e., the collection of human and animal cell cultures of the DSMZ, <http://www.dsmz.de>).

Catalogues can be added (or, rarely, removed) without notice. See the [CABRI site](#) for further information.

Cell line name: (required input)

Description: The input must specify the name of the required cell line(s). Due to the indexing rules in the CABRI network service (see the <http://www.cabri.org/>), only one word can be used in the search and no spaces are allowed in the cell line name.
Moreover:
cell lines names cannot include the following characters: '/', '-' and '*',
cell lines names are case insensitive.
Example of valid cell lines names are:
vero
hela
a172
calu6

Execute

O2I Project - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo http://www.o2i.it:8080/o2i/applet_page.jsp?file_name=fin20050523123547686668.xml Vai



O₂I (Oncology over Internet) Project

Your personalized research web site.

USER: PaoloRomano Clone Window Logout

All workflows list

My last executed

My domains workflows

My role most popular

My role last executed

Search by ontology

All available results

Unsaved results

Temporary saved results

Persistently saved results

Edit your profile

Results visualization

Execution state: Load com

Result Viewer

< > Save to disc

cell_line_des

List

text/plain

Click to view...

Accession_number ICLC ATL95005

Cell_line_name Vero

Brief_description Species: monkey, African green; Tissue: kidney

Description Species: monkey, African green adult; Tissue: kidney

Depositor Obtained from ECACC.

Bibliographic_reference Nippon Rinsho 1963;21:1209

Morphology fibroblast, grown as monolayer

Culture_conditions continuous culture; DMEM + 10% FBS + 2mM L-Glutamine; s

Viruses Search for viruses was not performed

Properties indicator line for mycoplasma testing; virology; virus titration; virus rep

Release_conditions Cell line available for distribution. For non-commercial inve

Hazard -

Species_validation Validated by isoenzymes: confirmed as monkey with AST, MD

Karyology modal number 58

Freezing_medium Culture medium + 50% FBS + 10% DMSO

Sterility mycoplasma negative, HOECHST and PCR

Further_bibliography Kitasato Arc Exp Med 1964;37:27-42 [PMID: 5833688]

Clone Window [logout](#)

[Edit your profile](#)

All the results in the same table have been produced by the same workflow execution.

Execution Details	Workflow Inputs	Results list
Date of Execution: 14:44 - 23/05/2005 Workflow name: Retrieve Cell Lines Descriptions By Name (Workflow diagram)	libs = 'iclc' name = 'vero'	☐ Workflow output ☐ Text lines separator ☐ Substitute blanks with SP ☐ Regex for catalogue name extraction ☐ Group for catalogue name extraction ☐ Separate cell line ids ☐ Extract ids by removing catalogues names ☐ getCellLineIdsByName ☐ getCellLinesById

02i Project - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

Indirizzo <http://149.132.158.111:8080/o2i/showallresults.jsp> Vai

My role last executed

Date of Execution: 13:08 - 04/10/2005

Workflow name: [Show Gene Ontology Term Context](#)

Search by ontology

<http://149.132.158.111:8080/o2i/workflow/RetrieveCellLinesDescrip...>

Indirizzo <http://149.132.158.111:8080/o2i/workflow/RetrieveCellLinesDescriptionsByName> Vai

Workflow Inputs

libs name

'text/plain' 'text/plain'

'text/plain' 'text/plain'

libs name

getCellLineIdsByName

result

Text_lines_separator

value

'text/plain' 'text/plain'

'text/plain' 'text/plain'

string regex

Separate_cell_line_ids

split

Regex_for_catalogue_name_extraction

value

'text/plain' 'text/plain'

I('text/plain')

I('text/plain')

'text/plain' 'text/plain'

'text/plain' 'text/plain'

Workflow output

ancestorColour

childColour

inputTermColour

getImmediateChildren

getParents

getAncestry

create

getChildren

getresults

Workflow Inputs

libs = 'iclc'

name = 'vero'

Results list

Workflow output

Text lines separator

Substitute blanks with SP

Regex for catalogue name extraction

Group for catalogue name extraction

Separate cell line ids

Extract ids by removing catalogues names

getCellLineIdsByName

getCellLinesById

Save temporarily Delete

Colors mean:

red

temporary saved

green

permanently saved

Collaborazioni

IST, Genova

Paolo Romano,
Ulrich Pfeffer,
Domenico Marra,
Valentina Mirisola,
M. Assunta Manniello

ISMAL, CNR, Genova

Patrizio Arrigo,
Matteo Fattore

ITB, CNR, Milano

Luciano Milanesi

DISCo, Università di Milano Bicocca

Guglielmo Bertolini,
Flavio De Paoli,
Giancarlo Mauri

DIST, Università di Genova

Ivan Porro,
Silvia Scaglione

DMI, Università di Camerino (MC)

Emanuela Merelli

Lavoro parzialmente supportato dal Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), Progetto Strategico "Oncology over Internet" (2002 – 2005).