

MIC ENRICHMENT



SITO WEB

- Realizzazione di un sito per conservare e rendere facilmente accessibili tutti i materiali di cui parliamo nella presentazione;
- <https://mic-ontology.herokuapp.com/>;
- Utilizzo dell'URL del sito come base dell'URI per le entità e proprietà create in fase di mapping.

RDF TRIPLE STORE MIC

[Home](#) [Xcel e CSV](#) [Trasformazioni](#) [Output](#)

ABSTRACT

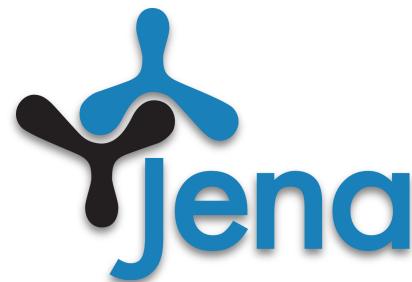
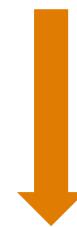
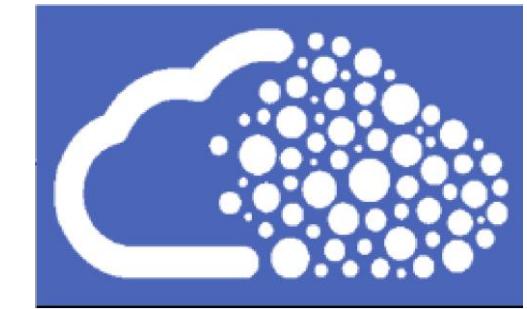
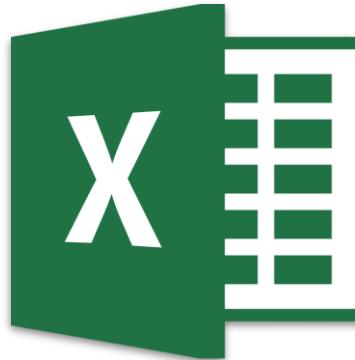
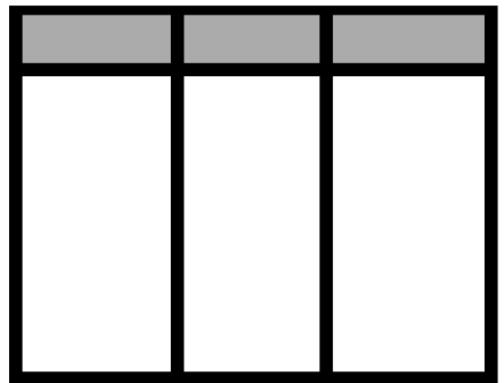
Lo scopo di questo progetto è l'arricchimento dei dati dell'archivio del MIC, in seguito al mapping RDF del suddetto su Datagraft. Per prima cosa abbiamo ripulito il file.xlsx cercando di renderlo il più possibile omogeneo e facile da manipolare. Abbiamo eseguito il mapping con Datagraft, poi abbiamo eseguito il linking con SILK tra il nostro dataset e quello di DBpedia. Infine, con Jena Fuseki abbiamo consultato il dataset attraverso delle SPARQL queries.

Tutte le fasi del progetto sono state trattate approfonditamente in [questa presentazione PowerPoint](#).



The screenshot shows a web application interface for managing an RDF triple store. At the top, there are four navigation links: Home, Xcel e CSV, Trasformazioni, and Output. Below these is a section titled 'ABSTRACT' containing a detailed description of the project's goals and methodology. At the bottom, there are three icons: a white grid icon, a green 'X' icon inside a square, and a blue cloud icon with dots. To the right of these icons are the labels 'Silk Workbench' and 'Jena' with its logo.

PIPELINE



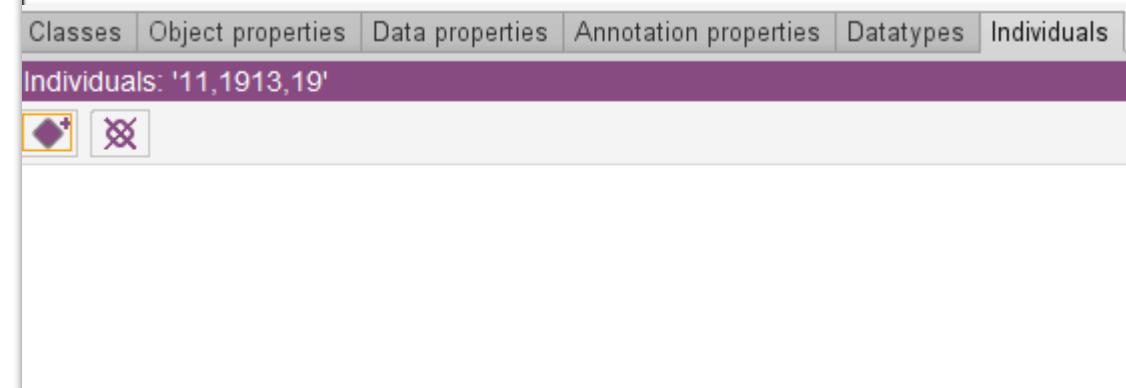
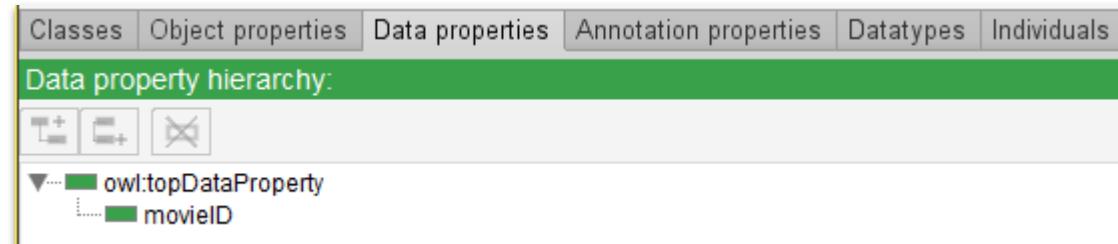
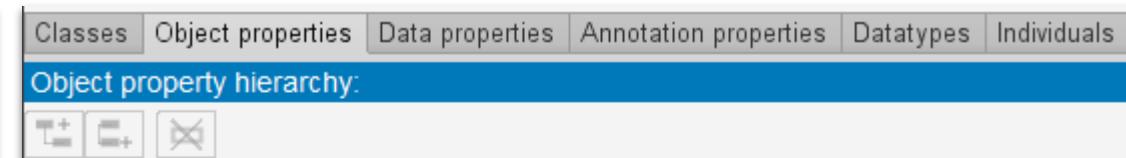
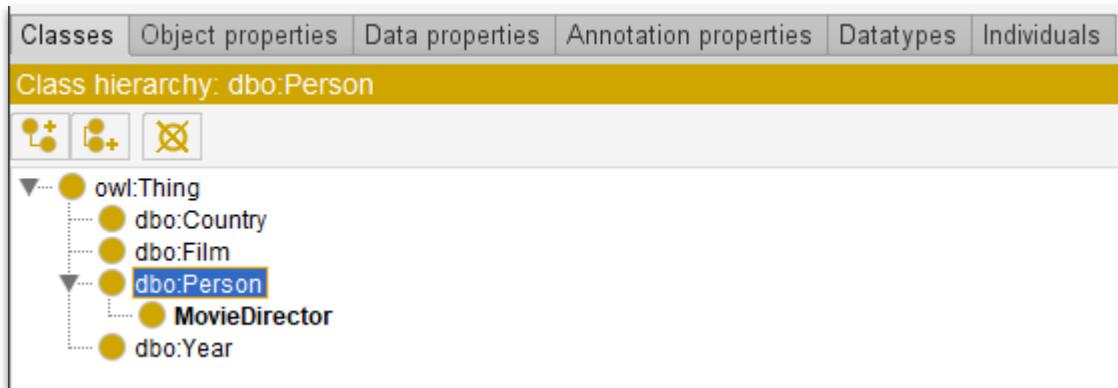
Silk Workbench

PROTÉGÉ 1

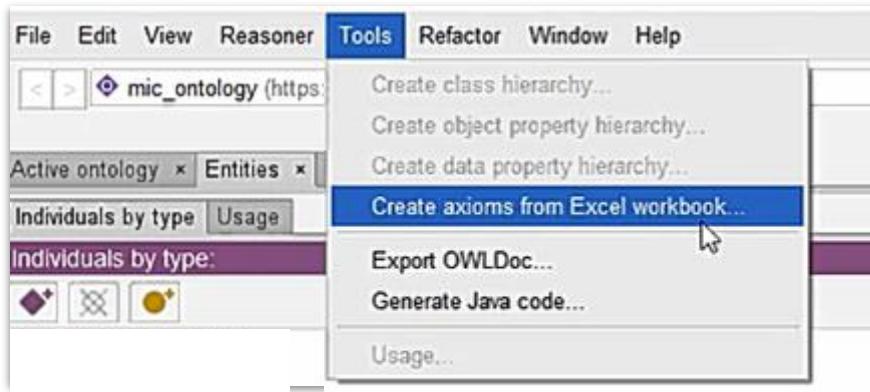
- Inizialmente è stato preso in considerazione **Protégé**;
- Protégé ha un plugin chiamato **Cellfie** che permette di mappare un documento excel, specificando delle regole di mapping che portano alla creazione degli assiomi dell'ontologia;
- Cellfie utilizza **MappingMaster**, una libreria Java open-source;
- Le regole vanno scritte in **Transformation Rule Language**.



PROTÉGÉ 2: classi e proprietà



PROTÉGÉ 3: mapper e regole



Cellfile

Target Ontology: mic_ontology (https://mic-ontology.herokuapp.com/mic_ontology.owl)

Workbook (C:\Users\User\Desktop\Cose per progetto data semantics\Dataset\regole film prova.json)

Table 1

	A	B
1	N.	Titolo
2	1	PARTE' DEGLI ANGELI, LA
3	2	SE TORNIO - ERNEST PIGNON-ERNEST E LA FIGURA DI PASOLINI
4	3	... E FUORI NEVICA
5	4	... E FUORI NEVICA
6	5	..LE L'UOMO CREO' SATANA
7	6	..LE ORA PARLIAMO DI KEVIN
8	7	¡VIVAN LAS ANTIPODAS! - VIVE LESANTIPODES! (VIVA GLI ANTIPODI!)
9	8	¡VIVAN LAS ANTIPODAS! - VIVE LESANTIPODES! (VIVA GLI ANTIPODI!)
10	9	007 - CASINO ROYALE - QUANTUMOF SOLACE - SKYFALL
11	10	007 - DALLA RUSSIA CON AMORE
12	11	007 - GOLDENEYE
13	12	007 - IL DOMANI NON MUORE MAI

Transformation Rules (C:\Users\User\Desktop\Cose per progetto data semantics\Dataset\regole film prova.json)

Add	Edit	Delete	Load Rules	Save Rules	Save As...		
<input checked="" type="checkbox"/>	Sheet Name	Start Column	End Column	Start Row	End Row	Rule	Comment
<input checked="" type="checkbox"/>	Table 1	A	E	2	2000	Individual: @D* Types: dbo:Year Facts: producedInYear @A*(rdfs:label=@B*)	
<input checked="" type="checkbox"/>	Table 1	A	E	2	2000	Individual: @C*(rdfs:label=@C*) Types: dbo:Person, MovieDirector Facts: shoted @B*	
<input checked="" type="checkbox"/>	Table 1	A	E	2	2000	Individual: @A*(rdfs:label=@B*) Types: dbo:Film Facts: dbo:director @C* Facts: productionYear @D* Facts: dbp:country @E* Facts: movieID @A*	
<input checked="" type="checkbox"/>	Table 1	A	E	2	2000	Individual: @E*(rdfs:label=@E*) Types: dbo:Country Facts: countryProduced @A*(rdfs:label=@B*)	

Generate Axioms

PROTÉGÉ 4: esempio di regole

Cellfile

Target Ontology: mic_ontology (https://mic-ontology.herokuapp.com/mic_ontology.owl)

Workbook (C:\Users\User\Desktop\Cose per progetto data semantics\CinetecaMilano_Elenco_DVD.xlsx)

Table 1

	A	B	C	D	E
1	N.	Titolo	Regia	Anno	Nazione
2	1	PARTE' DEGLI ANGELI, LA	LAVERTY Paul	2012	UK/France/Belgium/Ital
3	2	SE TORNIO' - ERNEST PIGNON-ERNEST E LA FIGURA DI PASOLINI	COLLETTIVO SIKOZEL	2016	Italy/France
4	3	... E FUORI NEVICA	SALEMME Vincenzo	2014	Italy
5	4	... E FUORI NEVICA	SALEMME Vincenzo	2014	Italy
6	5	...E L'UOMO CREO' SATANA	KRAMER, Stanley	1960	USA
7	6	...E ORA PARLIAMO DI KEVIN	RAMSAY, Lynne	2011	GB
8	7	¡VIVAN LAS ANTIPODAS! - VIVE LESANTIPODES! (VIVA GLI ANTIPODI!)	KOSSAKOVSKY Victor	2011	Germany/Argentina/Netherlands/Chile
9	8	¡VIVAN LAS ANTIPODAS! - VIVE LESANTIPODES! (VIVA GLI ANTIPODI!)	KOSSAKOVSKY Victor	2011	Germany/Argentina/Netherlands/Chile
10	9	007 - CASINO ROYALE - QUANTUM OF SOLACE - SKYFALL	CAMPBELL, Martin - FORSTER, ...	2006-2012	USA / GB / Italy
11	10	007 - DALLA RUSSIA CON AMORE	YOUNG, Terence	1963	USA
12	11	007 - GOLDENEYE	CAMPBELL Martin	1995	UK/USA
13	12	007 - IL DOMANI NON MUORE***	SPOTTISWOODE P	1997	UK/USA

Transformation Rule Editor

Transformation Rules (C:\Users\User\Desktop\...)

Add Edit Delete

Sheet Name	Start Column	End Column
Table 1	A	E

Sheet name: Table 1
Start column: A
End column: E
Start row: 2
End row: 2000
Comment:
Rule:
Individual: @A*(rdfs:label=@B*)
Types: dbo:Film
Facts: dbo:director @C*
Facts: productionYear @D*
Facts: dbp:country @E*
Facts: movieID @A*

OK Annulla

Generate Axioms

PROTÉGÉ 5: creazione degli assiomi

Cellie

Target Ontology: mic_ontology (https://mic-ontology.herokuapp.com/mic_ontology.owl)

Workbook (C:\Users\User\Desktop\Cose per progetto data semantics\CinetecaMilano_Elenco_DVD.xlsx)

Table 1

N.	A	E
1	N.	Titolo
2	1	PARTE' DEGLI ANGELI, LA
3	2	SE TORNO' - ERNEST PIGNO
4	3	... E FUORI NEVICA
5	4	... E FUORI NEVICA
6	5	...E L'UOMO CREO' SATANA
7	6	...E ORA PARLIAMO DI KEVIN
8	7	¡VIVAN LAS ANTIPODAS! - VIV
9	8	¡VIVAN LAS ANTIPODAS! - VIV
10	9	007 - CASINO ROYALE - QU
11	10	007 - DALLA RUSSIA CON AM
12	11	007 - GOLDENEYE
13	12	007 - IL DOMANI NON MUORE

Transformation Rules (C:\Users\User\Desktop\Cose per progetto data semantics\TransformationRules.xsl)

Add	Edit	Delete	
<input checked="" type="checkbox"/>	Sheet Name	Start Column	E
<input checked="" type="checkbox"/>	Table 1	A	E
<input checked="" type="checkbox"/>	Table 1	A	E
<input checked="" type="checkbox"/>	Table 1	A	E
<input checked="" type="checkbox"/>	Table 1	A	E
<input checked="" type="checkbox"/>	Table 1	A	E

Generated Axioms

Cellie generates 22872 axioms:

- Individual: A-TEAM
- Individual: A-TEAMVERSIONEESTESA
- Individual: A.INTELLIGENZAARTIFICIALE
- Individual: AA.VV.
- Individual: AAL
- Individual: AB-NORMALBEAUTY
- Individual: ABARTH-ILRITORNODELLOSCORPIONE-DVD1
- Individual: ABARTH-ILRITORNODELLOSCORPIONE-DVD2
- Individual: ABRONZATISSIMI2-UNANNODOPPO
- Individual: ABCsOFDEATHII
- Individual: ABDUCTION
- Individual: ABEAUTIFULMIND
- Individual: ABENDLAND
- Individual: ABLOOMINGBUSINESS
- Individual: ABOMINABLE
- Individual: ABOUTADAM
- Individual: ABRANDNEWLIFE
- Individual: ABRIGLIASCIOLTA

K/France/Belgium/Italy
aly/France
aly
isa
B
ermany/Argentina/Netherlands/Chile
ermany/Argentina/Netherlands/Chile
isa / GB / Italy
isa
KUSA
K/USA

Save Rules Save As...

Comment

View Log

Cancel Add to a new ontology Add to current ontology

Generate Axioms

PROTÉGÉ 6: individuals

mic_ontology (https://mic-ontology.herokuapp.com/mic_ontology.owl)

Active ontology | Entities | Individuals by class | DL Query | Search...

Individuals by type: dbo:Year

Annotations | Usage

Annotations: dbo:Year

Description: dbo:Year

Equivalent To | SubClass Of | General class axioms

SubClass Of (Anonymous Ancestor)

Instances |

Individuals: '1937,1945'

Instances |

'11,1913,19'
'1918(?)'
'1937,1945'
'2001(1984)'
'2002,2003'
'43 - '97'
-2 LIVELLO DEL TERRORE
... E FUORI NEVICA
...E L'UOMO CREO' SATANA
...E ORA PARLIAMO DI KEVIN
0.58
007 - CASINO ROYALE - QUANTUM OF SOLACE - SKYFALL
007 - DALLA RUSSIA CON AMORE
007 - GOLDENEYE
007 - IL DOMANI NON MUORE MAI

To use the reasoner click Reasoner > Start reasoner | Show Inferences

PROTÉGÉ 7: esempi di individual

Description: LUPO, Michele

Property assertions: LUPO, Michele

Types +

- dbo:Person
- MovieDirector

Object property assertions +

- shooteed 'ARIZONA COLT'
- shooteed 'CHISSA' PERCHE'... CAPITANO TUTTE A ME'
- shooteed BOMBER

Data property assertions +

Negative object property assertions +

Negative data property assertions +

Description: PARTE' DEGLI ANGELI, LA

Property assertions: PARTE' DEGLI ANGELI, LA

Types +

- dbo:Film

Object property assertions +

- dbp:country UK/France/Belgium/Ital
- dbo:director 'LAVERTY Paul'
- productionYear 2012

Data property assertions +

- movielD "1"^^xsd:string

Negative object property assertions +

Negative data property assertions +

Description: Belgium

Property assertions: Belgium

Types +

- dbo:Country

Object property assertions +

- countryProduced 'ALABAMA MONROE - UNA STORIA D'AMORE'
- countryProduced 'AVVENTURE DI SAMMY, LE'
- countryProduced 'BLUE BIRD'
- countryProduced 'AFRICAN SAFARI'

Data property assertions +

Negative object property assertions +

Negative data property assertions +

Description: 2015/2016

Property assertions: 2015/2016

Types +

- dbo:Year

Object property assertions +

- producedInYear 'ARROW - QUARTA STAGIONE'
- producedInYear 'BIG BANG THEORY, THE - NONA STAGIONE'
- producedInYear 'AMERICAN HORROR STORY- HOTEL - QUINTA STAGIONE'
- producedInYear 'C'ERA UNA VOLTA - QUINTA'

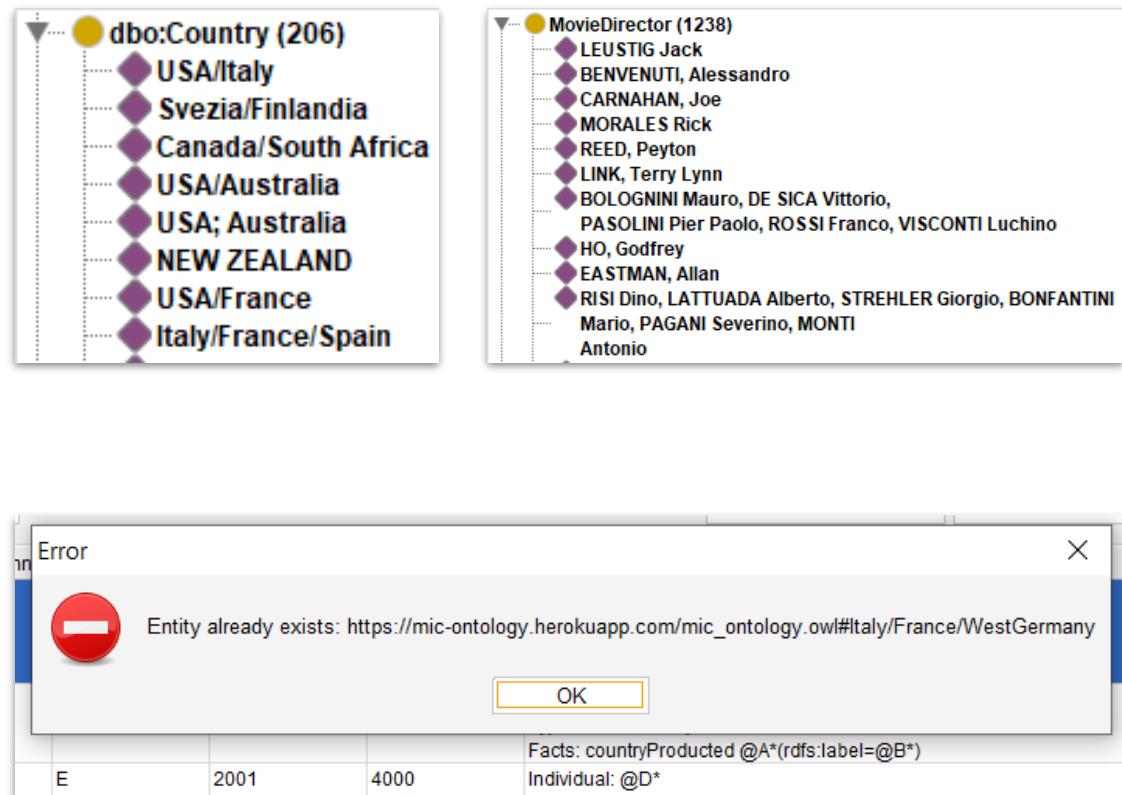
Data property assertions +

Negative object property assertions +

Negative data property assertions +

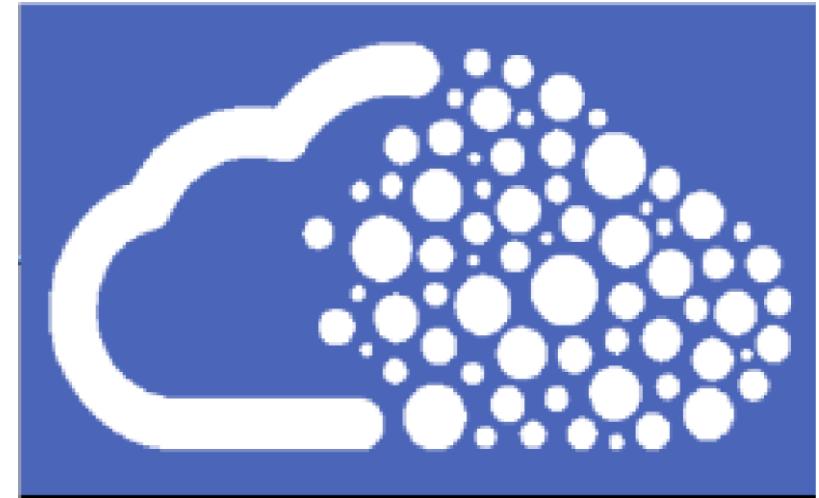
PROTÉGÉ 8: problemi e criticità

- Non si possono spartire i nomi e i paesi;
- Il mapper non è stato in grado di creare gli assiomi per tutte le righe del documento;
- Per risolvere il problema sono state applicate le regole solo alle prime 2000 righe, pensando di farne 2000 alla volta;
- Quando si passa alle righe 2001-4000, il mapper ha dà errore, segnalando entità già esistenti.



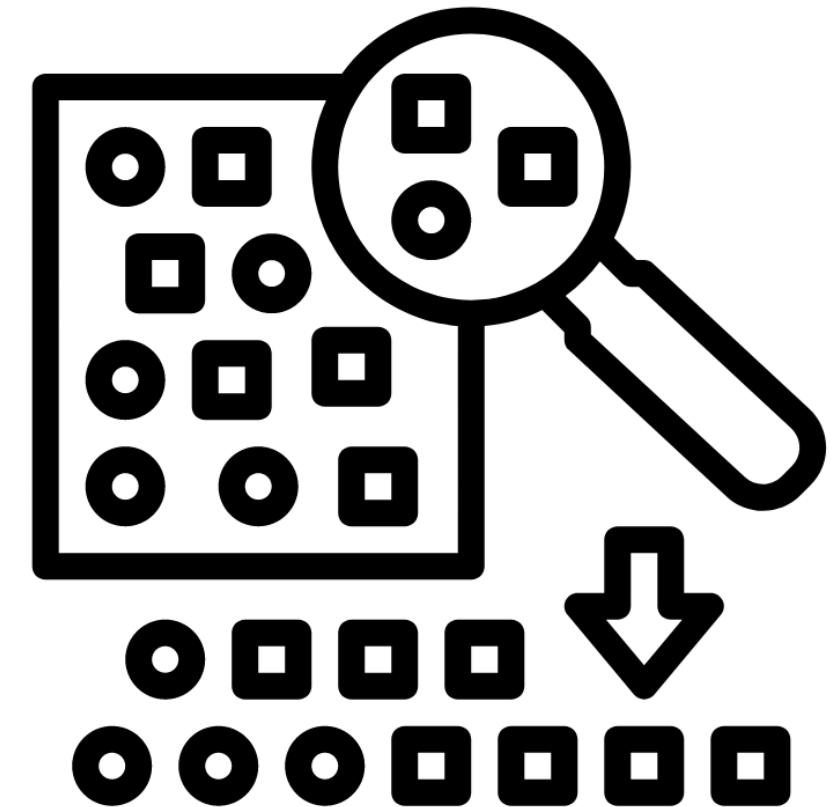
PASSAGGIO A DATAGRAFT

- A causa dei vari problemi riscontrati con il mapper di Protégé, è stato deciso di passare a DataGraft per la fase di mapping.

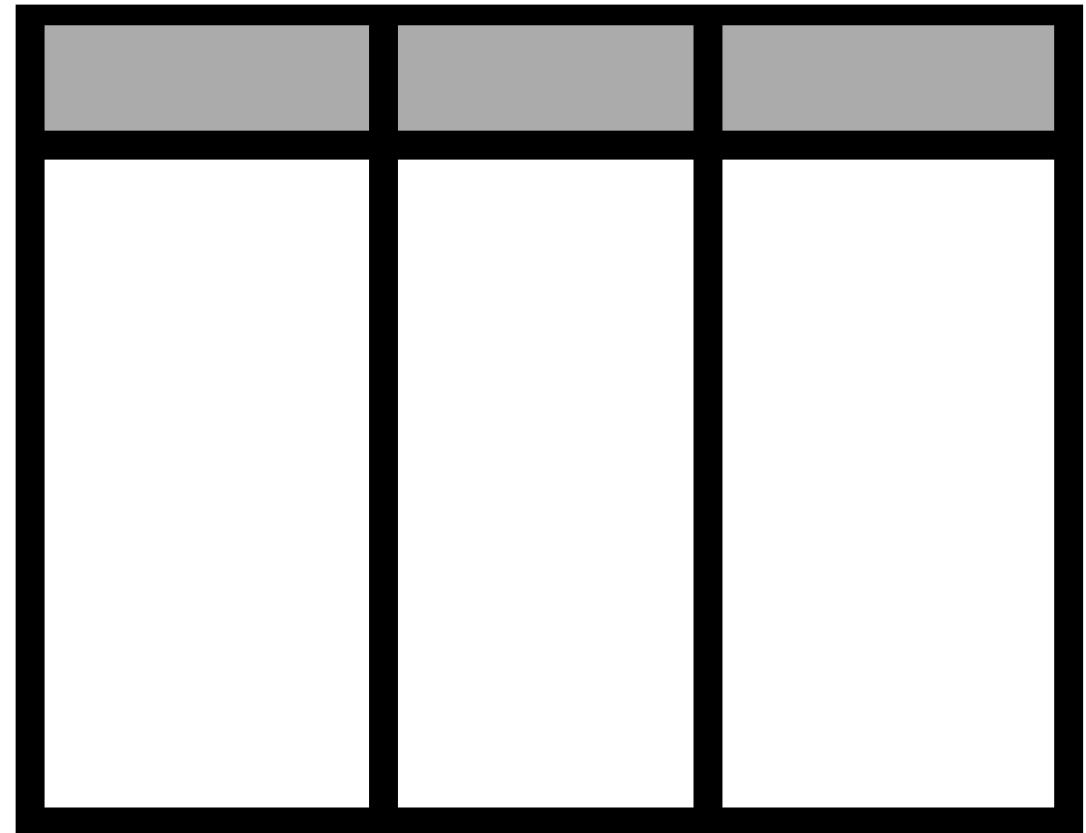


SAMPLE

- Dataset originale: 12mila righe circa.
- Sample: 200 righe circa.



1- RAW DATA



1- RAW DATA

N.	Titolo	Regia	Anno	Nazione
1	"PARTE" DEGLI ANGELI, LA	LAVERTY Paul	2012	UK/France/Belgium/Italy
2	"SE TORNO" - ERNEST PIGNON-ERNEST E LA FIGURA DI PASOLINI	COLLETTIVO SIKOZEL	2016	Italy/France
3	... E FUORI NEVICA	SALEMME Vincenzo	2014	Italy
4	... E FUORI NEVICA	SALEMME Vincenzo	2014	Italy
5	...E L'UOMO CREO' SATANA	KRAMER, Stanley	1960	USA
6	...E ORA PARLIAMO DI KEVIN	RAMSAY, Lynne	2011	GB
7	¡VIVAN LAS ANTIPODAS! - VIVE LES ANTIPODES! (VIVA GLI ANTIPODI!)	KOSSAKOVSKY Victor	2011	Germany/Argentina/Netherlands/Chile
8	¡VIVAN LAS ANTIPODAS! - VIVE LES ANTIPODES! (VIVA GLI ANTIPODI!)	KOSSAKOVSKY Victor	2011	Germany/Argentina/Netherlands/Chile
9	007 - CASINO ROYALE - QUANTUM OF SOLACE - SKYFALL	CAMPBELL, Martin - FORSTER, Marc - MENDES, Sam	2006-2012	USA / GB / Italy
10	007 - DALLA RUSSIA CON AMORE	YOUNG, Terence	1963	USA
11	007 - GOLDENEYE	CAMPBELL Martin	1995	UK/USA
12	007 - IL DOMANI NON MUORE MAI	SPOTTISWOODE Roger	1997	UK/USA
13	007 - IL MONDO NON BASTA	APTED, Michael	1999	USA
14	007 - IL MONDO NON BASTA	APTED Michael	1999	UK/USA

https://www.cinetecamilano.it/pdf/CinetecaMilano_Elenco_DVD.pdf

1- RAW DATA: titolo

Maiuscolo	AIDA DELLE MARIONETTE - ATELIER COLLA AIDA'S BROTHERS AND SISTERS AIR DOLL
Virgolette ("")	Titolo "PARTE" DEGLI ANGELI, LA "SE TORNO" - ERNEST PIGNON- ERNEST E LA FIGURA DI PASOLINI
Due punti (:)	3 SUPEREROI: DOTTOR STRANGE - IL MAGO SUPREMO 3 SUPEREROI: L'INVINCIBILE IRON MAN
Duplicati	3 SUPEREROI: DOTTOR STRANGE - IL MAGO SUPREMO 3 SUPEREROI: DOTTOR STRANGE - IL MAGO SUPREMO

1- RAW DATA: regia 1

Cognomi maiuscoli
e minuscoli

Regia
LAVERTY Paul
COLLETTIVO SIKOZEL
Malloy, Chris

Virgolette ("")

ARBUCKLE, Roscoe "Fatty"

Film con troppi autori

VIGALONGO Nacho GARCIA BOGLIANO Adrián DIAZ ESPINOZA
Ernesto SARMIENTO Marcel BETTIS Angela IGUCHI Noboru TRAUCKI
Andrew CAPPELEN MALLING
Thomas GRAU Jorge Michel YAMAGUCHI Yudai MORGENTHALER
Anders TJAHIJANTO Timo WEST Ti PISANTHANAKUN Banjong CATTET
Hélène FORZANI Bruno RUMLEY Simon WINGARD Adam SPASOJEVIC
Srdjan WEST Jake HARDCASTLE Lee WHEATLEY Ben
ANDREWS Kaare SCHNEPP Jon

Separatori disomogenei

DI FELICE, Alessandro \ DI FELICE,
Cristiano

RISTORI, Marco - BONI, Luca

1- RAW DATA: regia 2

(Alias)	<p>YOSHIDA Yoshishige (alias YOSHIDA Kiju)</p>
(Ideato da)	<p>SURNOW, Joel - COCHRAN, Robert (Ideato da)</p>
(A cura di)	<p>Comitato provinciale DC di Milano (a cura di)</p>

1- RAW DATA: regia 3

(Accreditato come)	WYNORSKI Jim (accreditato come Jay Andrews)
(As)	YOSHIDA Yoshishige (as YOSHIDA Kiju)
Caso Wachowski	TYKWER, Tom, WACHOWSKI, Andy - WACHOWSKI, Lana

1- RAW DATA: anno

Separatori disomogenei

2006-2012

2014/2015

Spazi mancanti

2006-2012

1- RAW DATA: nazione 1

Lingue diverse per le stesse nazioni	France Cina, Francia
Abbreviazioni	GB
Separatori disomogenei	SPA - CANADA / GB - USA

1- RAW DATA: nazione 2

Maiuscolo e minuscolo

HONG KONG

USA

France/Germany

Spazi casuali

Italy / France

France/Germany

1- RAW DATA: generale

Celle vuote

NASSER, Elie		
	2013	USA
	2013	USA

2- DATA PREPARATIO N



2- DATA PREPARATION: titolo

Maiuscolo

AIDA DELLE MARIONETTE - ATELIER COLLA
AIDA'S BROTHERS AND SISTERS
AIR DOLL



Aida Delle Marionette - Atelier Colla
Aida'S Brothers And Sisters
Air Doll

Virgolette ("")

Titolo
"PARTE" DEGLI ANGELI, LA
"SE TORNO" - ERNEST PIGNON- ERNEST E LA FIGURA DI PASOLINI



Parte Degli Angeli, La
Se Torno - Ernest Pignon- Ernest E La

Due punti (:)

3 SUPEREROI: DOTTOR STRANGE - IL MAGO SUPREMO
3 SUPEREROI: L'INVINCIBILE IRON MAN



3 Supereroi, Dottor Strange - Il Mago Supremo
3 Supereroi, L'Invincibile Iron Man

Duplicati

3 SUPEREROI: DOTTOR STRANGE - IL MAGO SUPREMO
3 SUPEREROI: DOTTOR STRANGE - IL MAGO SUPREMO



3 Supereroi, Dottor Strange - Il Mago Supremo

2- DATA PREPARATION: regia 1

Cognomi maiuscoli
e minuscoli

Regia
LAVERTY Paul
COLLETTIVO SIKOZEL
Malloy, Chris



Laverty Paul
Collettivo Sikozel

Virgolette (")

ARBUCKLE, Roscoe "Fatty"



Arbuckle Roscoe (Fatty)

Film con troppi autori

VIGALONGO Nacho GARCIA BOGLIANO Adrián DIAZ ESPINOZA
Ernesto SARMIENTO Marcel BETTIS Angela IGUCHI Noboru TRAUCKI
Andrew CAPPELEN MALLING
Thomas GRAU Jorge Michel YAMAGUCHI Yudai MORGENTHALER
Anders TJAHHANTO Timo WEST Ti PISANTHAKUN Banjong CATTET
Hélène FORZANI Bruno RUMLEY Simon WINGARD Adam SPASOJEVIC
Srdjan WEST Jake HARDCASTLE Lee WHEATLEY Ben
ANDREWS Kaare SCHNEPP Jon



AA.VV.

Separatori disomogenei

DI FELICE, Alessandro \ DI FELICE,
Cristiano

RISTORI, Marco - BONI, Luca



Di Felice Alessandro / Di Felice Cristiano
Ristori Marco / Boni Luca

2- DATA PREPARATION: regia

2

(Alias)

YOSHIDA Yoshishige (alias YOSHIDA
Kiju)



Yoshida Yoshishige / As Youshida Kiju

(Ideato da)

SURNOW, Joel - COCHRAN, Robert
(Ideato da)



Surnow Joel / Cochran Robert

(A cura di)

Comitato provinciale DC di Milano (a
cura di)



Comitato Provinciale Dc Di Milano

2- DATA PREPARATION: regia

3

(Accreditato come)

WYNORSKI Jim (accreditato come
Jay Andrews)



Wynorski Jim / As Jay Andrews

(As)

YOSHIDA Yoshishige (as YOSHIDA
Kiju)



Yoshida Yoshishige / As Youshida Kiju

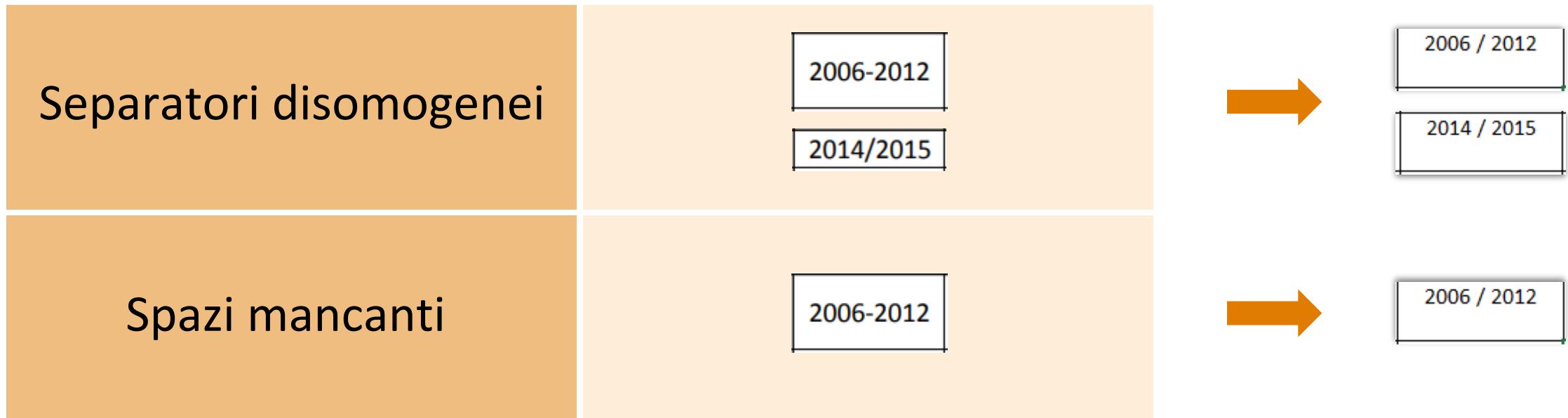
Caso Wachowski

TYKWER, Tom, WACHOWSKI, Andy -
WACHOWSKI, Lana



Tykwer Tom / Wachowski Lilly / Wachowski
Lana

2- DATA PREPARATION: anno



2- DATA PREPARATION: nazione 1

Lingue diverse per le stesse nazioni

France
Cina, Francia



France
Cina / France

Abbreviazioni

GB



United Kingdom

Separatori disomogenei

SPA - CANADA / GB - USA



Spagna / Canada / United Kingdom / United States

2- DATA PREPARATION: nazione 2

Maiuscolo e minuscolo

HONG KONG

USA

France/Germany



Hong Kong

United States

Spazi casuali

Italy / France

France/Germany



Italy / France

France / Germany

2- DATA PREPARATION: generale

Celle vuote

NASSER, Elie		
	2013	USA
	2013	USA



Nasser Elie	Nd	Nd
Nd	2013	United States
Nd	2014	United States

2- DATA PREPARATION 2

- Import del file MIC_Film_T1 in Excel;
- Duplicazione delle colonne “Registi” e “Nazioni”, creando le colonne “Registi_URI”, “Registi_Label”, “Nazioni_URI” e “Nazioni_Label”;
- Nelle colonne “Registi_URI” e “Nazioni_URI”, rimozione della punteggiatura e sostituzione il carattere spazio con l’underscore (_), in modo da non avere problemi con la forma degli URI.
- Rimozione degli a capo (\n) generati dalla trasformazione su Datagraft.

	A	B	C	D	E	F	G
1	N	Anno	Titolo	Nazioni_URI	Nazioni_Label	Registi_URI	Registi_Label
2		1	2012 Parte Degli Angeli, La	United_Kingdom	United Kingdom	Laverty_Paul	Laverty Paul
3		1	2012 Parte Degli Angeli, La	France	France	Laverty_Paul	Laverty Paul
4		1	2012 Parte Degli Angeli, La	Belgium	Belgium	Laverty_Paul	Laverty Paul
5		1	2012 Parte Degli Angeli, La	Italy	Italy	Laverty_Paul	Laverty Paul
6		2	2016 Se Torno - Ernest Pignon- Ernest E La Figura Di Pasolini	Italy	Italy	Collettivo_Sikozel	Collettivo Sikozel
7		2	2016 Se Torno - Ernest Pignon- Ernest E La Figura Di Pasolini	France	France	Collettivo_Sikozel	Collettivo Sikozel
8		3	2014 ... E Fuori Nevica	Italy	Italy	Salemme_Vincenzo	Salemme Vincenzo
9		4	1960 ...E L'Uomo Creo' Satana	United_States	United States	Kramer_Stanley	Kramer Stanley
10		5	2011 ...E Ora Parliamo Di Kevin	United_Kingdom	United Kingdom	Ramsay_Lynne	Ramsay Lynne

United Kingdom
Germany / Argentina / Net\hernlands / Chile
United States / United Kingdom / Italy

2- DATA PREPARATION 3

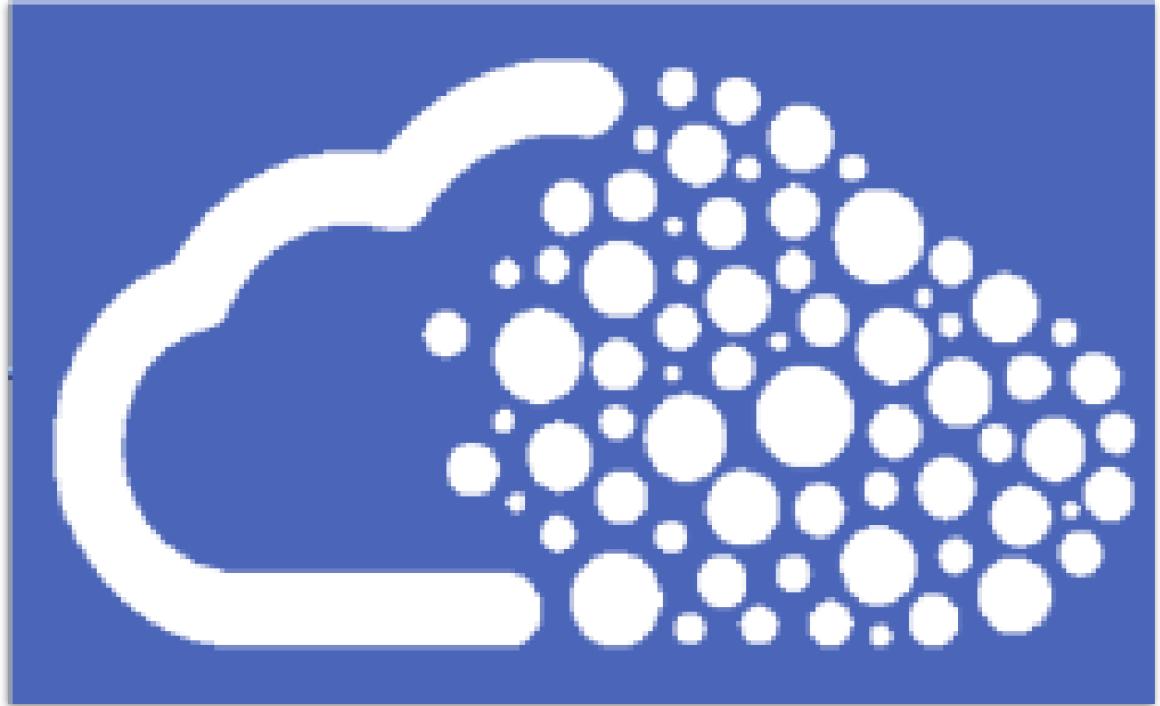
- Nelle colonne “Registi_URI” e “Registi_Label”, è stato invertito l’ordine del nome e del cognome (da cognome-nome a nome-cognome);
- La formula usata è: =DESTRA(C2;LUNGHEZZA(C2)-TROVA(" ";C2))&"&SINISTRA(C2;TROVA(" ";C2)-1).

Registi_URI	Registi_Label
Laverty_Paul	Laverty Paul
Collettivo_Sikozel	Collettivo Sikozel
Collettivo_Sikozel	Collettivo Sikozel
Salemme_Vincenzo	Salemme Vincenzo
Kramer_Stanley	Kramer Stanley
Ramsay_Lynne	Ramsay Lynne
Kossakovsky_Victor	Kossakovsky Victor

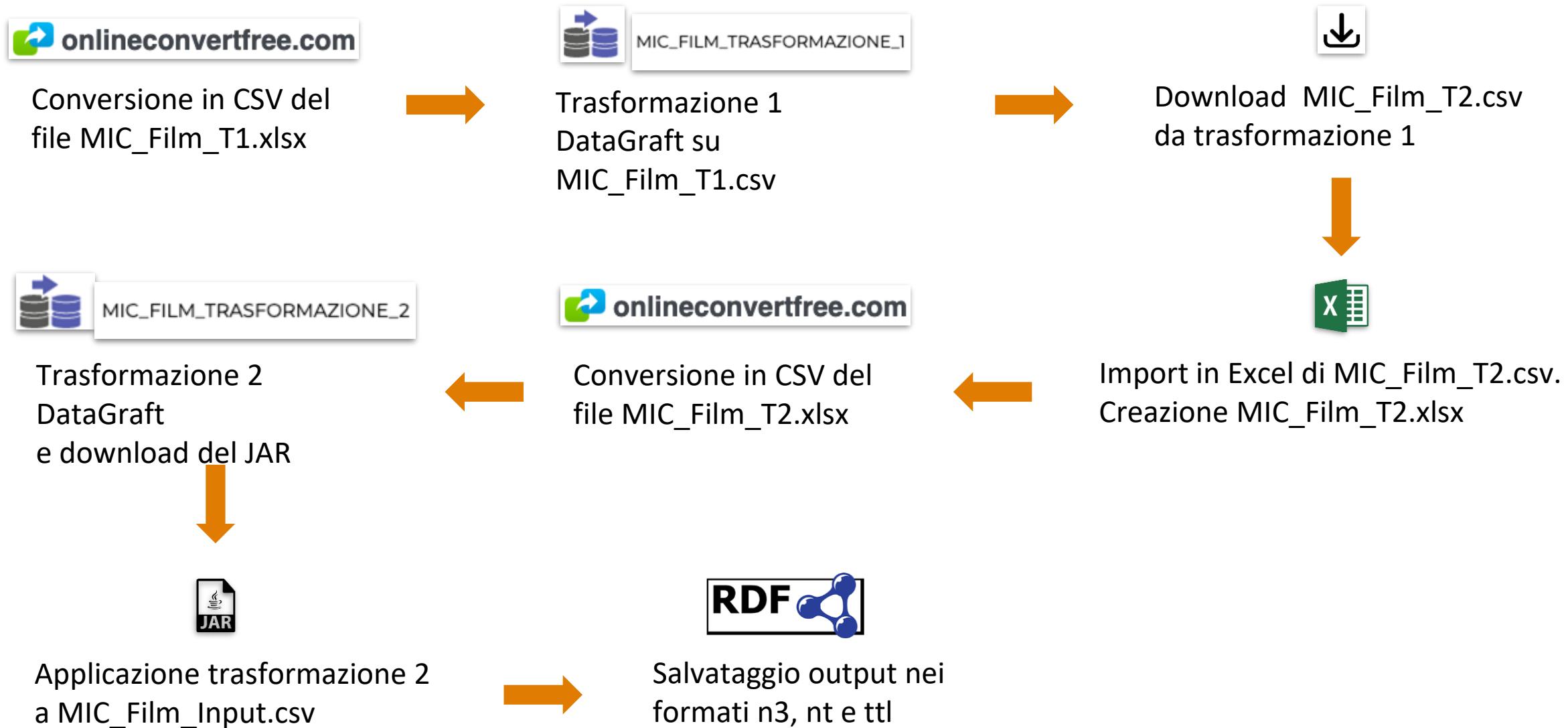


Registi_URI	Registi_Label
Paul_Laverty	Paul Laverty
Sikozel_Collettivo	Sikozel Collettivo
Sikozel_Collettivo	Sikozel Collettivo
Vincenzo_Salemme	Vincenzo Salemme
Stanley_Kramer	Stanley Kramer
Lynne_Ramsay	Lynne Ramsay
Victor_Kossakovsky	Victor Kossakovsky

3- RDF MAPPING (DataGraft)

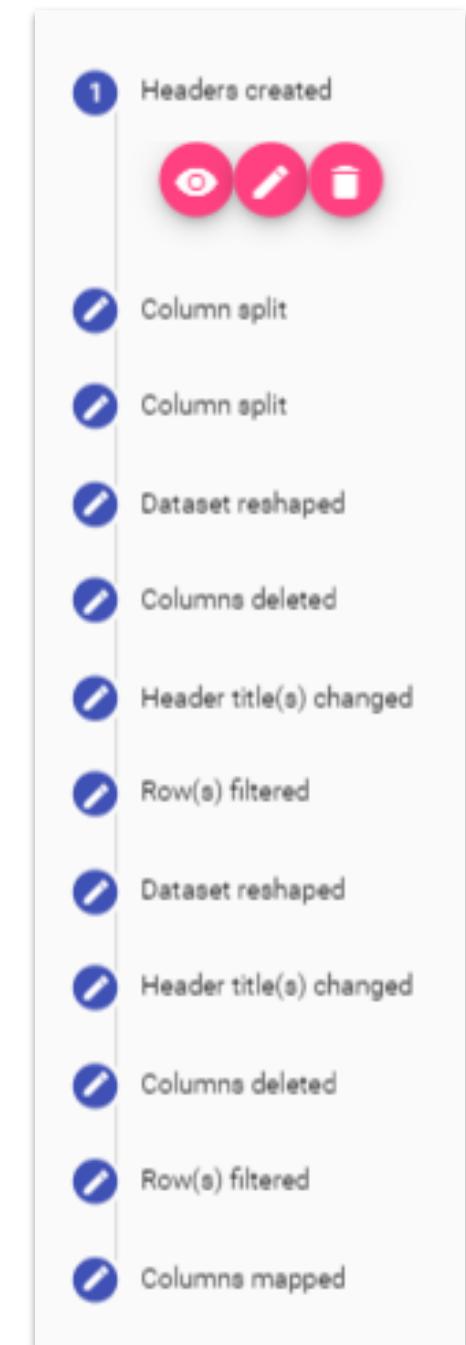


3- RDF MAPPING: pipeline



3- RDF MAPPING: trasformazione 1

- I. Creazione degli header, tramite **Make dataset > Use the first row as header**;
- II. Divisione della colonna “Regia” e della colonna “Nazione”, applicando **Split column** a ciascuna delle due;
- III. Esecuzione di **Reshape dataset**, con l’applicazione della funzione **Melt** alle colonne non spartite (eccetto “Nazione”);
- IV. **Take columns**, selezionando **value**; rimozione delle colonne listed;
- V. Cambio di nome alla colonna risultante dal passo IV (“value”) in “Nazioni”, tramite **Rename columns**;
- VI. Applicazione di **Filter rows** su “Nazioni”, tramite la funzione **not-empty**;
- VII. Applicazione di **Reshape dataset** alle colonne relative a “Regia”;
- VIII. Cambio di nome alla colonna risultante dal passo VII (“value”) in “Registi”, tramite **Rename columns**;
- IX. Applicazione di **Filter rows** su “Registi”, tramite la funzione **not-empty**;
- X. **Take columns**, selezionando **value**; rimozione delle colonne listed;
- XI. Esecuzione di **Map columns** sulle colonne “Anno”, “N”, “Titolo”, “Nazioni” e “Registi”, tramite la funzione **Trim**;
- XII. Esecuzione del **Tabular enrichment**.



3- RDF MAPPING: modifica Excel 1

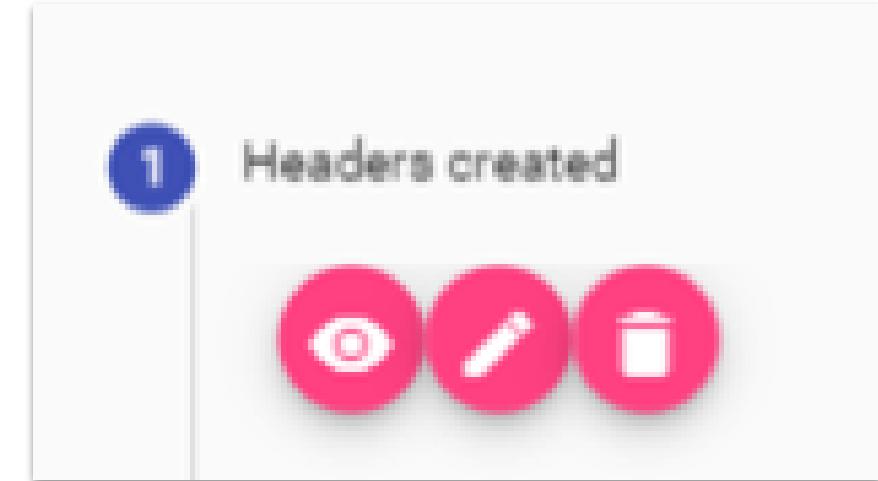
- Import del file MIC_Film_T1 in Excel;
- Duplicazione delle colonne “Registi” e “Nazioni”, creando le colonne “Registi_URI”, “Registi_Label”, “Nazioni_URI” e “Nazioni_Label”;
- Nelle colonne “Registi_URI” e “Nazioni_URI”, rimozione della punteggiatura e sostituzione il carattere spazio con l’underscore (_), in modo da non avere problemi con la forma degli URI.
- Rimozione degli a capo (\n) generati dalla trasformazione su Datagraft.

	A	B	C	D	E	F	G
1	N	Anno	Titolo	Nazioni_URI	Nazioni_Label	Registi_URI	Registi_Label
2		1	2012 Parte Degli Angeli, La	United_Kingdom	United Kingdom	Laverty_Paul	Laverty Paul
3		1	2012 Parte Degli Angeli, La	France	France	Laverty_Paul	Laverty Paul
4		1	2012 Parte Degli Angeli, La	Belgium	Belgium	Laverty_Paul	Laverty Paul
5		1	2012 Parte Degli Angeli, La	Italy	Italy	Laverty_Paul	Laverty Paul
6		2	2016 Se Torno - Ernest Pignon- Ernest E La Figura Di Pasolini	Italy	Italy	Collettivo_Sikozel	Collettivo Sikozel
7		2	2016 Se Torno - Ernest Pignon- Ernest E La Figura Di Pasolini	France	France	Collettivo_Sikozel	Collettivo Sikozel
8		3	2014 ... E Fuori Nevica	Italy	Italy	Salemme_Vincenzo	Salemme Vincenzo
9		4	1960 ...E L'Uomo Creo' Satana	United_States	United States	Kramer_Stanley	Kramer Stanley
10		5	2011 ...E Ora Parliamo Di Kevin	United_Kingdom	United Kingdom	Ramsay_Lynne	Ramsay Lynne

United Kingdom
Germany / Argentina / Net\nherlands / Chile
United States / United Kingdom / Italy

3- RDF MAPPING: trasformazione 2

- I. Esecuzione di **Make dataset > Use the first row as header**



3- RDF MAPPING: trasformazione 2 - mapping 1

DataGraft beta2

Explore Dashboard progettosemantic :

Tabular Transformation RDF Mapping Tabular Enrichment

Back Next step

UPDATE RDF MAPPINGS

ASIA Configuration

Annotation properties

Default URI prefix
<https://mic-ontology.herokuapp.com/>

Vocabulary Suggestions

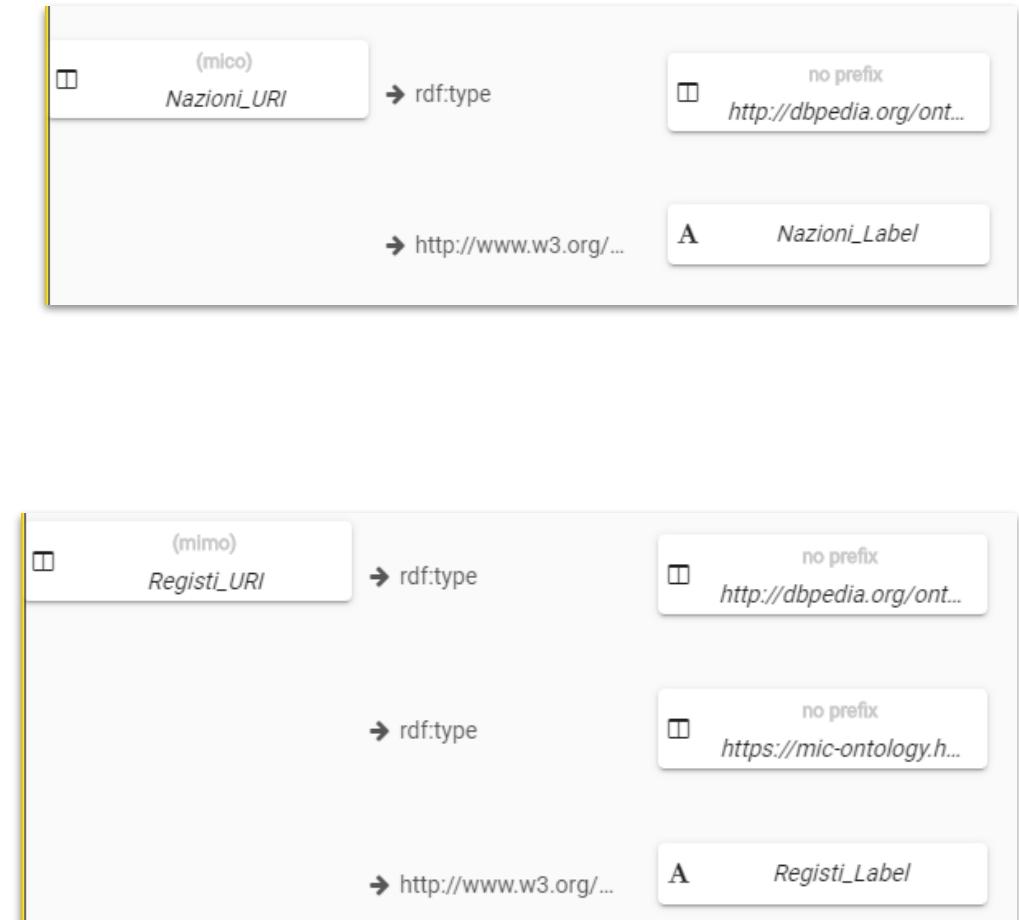
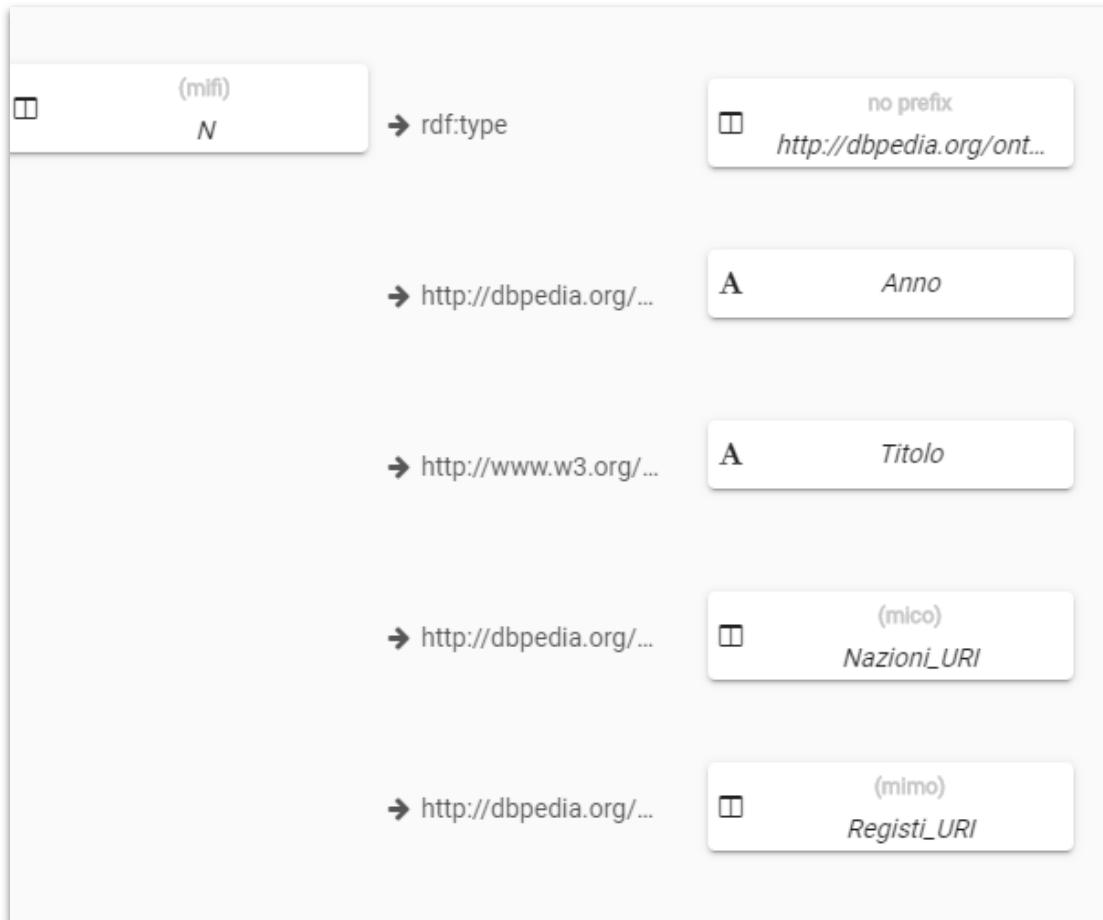
Suggerer ABSTAT Dataset language IT

Preferred summaries

DISCARD SAVE

	N	Anno	Titolo	Nazioni_URI	Nazioni_Label	Registi_URI	Registi_Label
1	1	2012	Parte Degli Angeli, La	United_Kingdom	United Kingdom	Laverty_Paul	Laverty Paul
2	1	2012	Parte Degli Angeli, La	France	France	Laverty_Paul	Laverty Paul
3	1	2012	Parte Degli Angeli, La	Belgium	Belgium	Laverty_Paul	Laverty Paul
4	1	2012	Parte Degli Angeli, La	Italy	Italy	Laverty_Paul	Laverty Paul
5	2	2016	Se Torno - Ernest Pignon- Ernest E La Figura Di Pasolini	Italy	Italy	Collettivo_Sikozel	Collettivo Sikozel
6	2	2016	Se Torno - Ernest Pignon- Ernest E La Figura Di Pasolini	France	France	Collettivo_Sikozel	Collettivo Sikozel
7	3	2014	... E Fuori Nevica	Italy	Italy	Salemme_Vincenzo	Salemme Vincenzo
8	4	1960	...E L'Uomo Creo' Sata		States	Kramer_Stanley	Kramer Stanley
9	5	2011	...E Ora Parliamo Di Ke	Kingdom	Ramsay_Lynne	Ramsay Lynne	
10	6	2011	♦♦?Vivan Las Antipodas! - Vive Les Antipo	Italy	Kossakovsky_Victor	Kossakovsky Victor	
11	6	2011	♦♦?Vivan Las Antipodas! - Vive Les Antipo	Spain	Kossakovsky_Victor	Kossakovsky Victor	
12	6	2011	♦♦?Vivan Las Antipodas! - Vive Les Antipo	Portugal	Kossakovsky_Victor	Kossakovsky Victor	
13	6	2011	♦♦?Vivan Las Antipodas! - Vive Les Antipo	Italy	Kossakovsky_Victor	Kossakovsky Victor	
14	7	2006 / 2012	007 - Casino Royale - Quantum Of	States	Campbell_Martin	Campbell Martin	
15	7	2006 / 2012	007 - Casino Royale - Quantum Of	States	Forster_Marc	Forster Marc	
16	7	2006 / 2012	007 - Casino Royale - Quantum Of	States	Mendes_Sam	Mendes Sam	
17	7	2006 / 2012	007 - Casino Royale - Quantum Of	Kingdom	Campbell_Martin	Campbell Martin	
18	7	2006 / 2012	007 - Casino Royale - Quantum Of	Kingdom	Forster_Marc	Forster Marc	
19	7	2006 / 2012	007 - Casino Royale - Quantum Of	Kingdom	Mendes_Sam	Mendes Sam	
20	7	2006 / 2012	007 - Casino Royale - Quantum Of	Italy	Campbell_Martin	Campbell Martin	
21	7	2006 / 2012	007 - Casino Royale - Quantum Of	Italy	Forster_Marc	Forster Marc	
22	7	2006 / 2012	007 - Casino Royale - Quantum Of	Italy	Mendes_Sam	Mendes Sam	

3- RDF MAPPING: trasformazione 2 - mapping 2



3- RDF MAPPING: trasformazione 2 - mapping 3

URI node Literal node Blank node

Select mapping source type
Free-defined text

Node name
<http://dbpedia.org/ontology/Film>

Specify condition

CANCEL OK

URI node Literal node Blank node

Select mapping source type
From dataset column

Select node prefix
mimo

Select column
Registi_URI

Specify condition

CANCEL OK

URI node Literal node Blank node

Select mapping source type
From dataset column

Select column Anno

Data type string

Language tag en

Assign data type

Show documentation

Specify condition

CANCEL OK

3- RDF MAPPING: trasformazione 2 - output

TTL



```
<https://mic-ontology.herokuapp.com/Film/1> a <http://dbpedia.org/ontology/Film> ;  
    <http://dbpedia.org/ontology/releaseDate> "2012"@en ;  
    <http://www.w3.org/2000/01/rdf-schema#label> "Parte Degli Angeli, La"@it ;  
    <http://dbpedia.org/property/country> <https://mic-ontology.herokuapp.com/Country/United_Kingdom> ;  
    <http://dbpedia.org/ontology/director> <https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> .  
  
<https://mic-ontology.herokuapp.com/Country/United_Kingdom> a <http://dbpedia.org/ontology/Country> ;  
    <http://www.w3.org/2000/01/rdf-schema#label> "United Kingdom"@en .  
  
<https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> a <http://dbpedia.org/ontology/Person> , <https://mic-ontology.herokuapp.com/MovieDirector> ;  
    <http://www.w3.org/2000/01/rdf-schema#label> "Paul Laverty"@en .
```

NT



```
[<https://mic-ontology.herokuapp.com/Film/1> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://dbpedia.org/ontology/Film> .  
<https://mic-ontology.herokuapp.com/Film/1> <http://dbpedia.org/ontology/releaseDate> "2012"@en .  
<https://mic-ontology.herokuapp.com/Film/1> <http://www.w3.org/2000/01/rdf-schema#label> "Parte Degli Angeli, La"@it .  
<https://mic-ontology.herokuapp.com/Film/1> <http://dbpedia.org/property/country> <https://mic-ontology.herokuapp.com/Country/United_Kingdom> .  
<https://mic-ontology.herokuapp.com/Film/1> <http://dbpedia.org/ontology/director> <https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> .  
<https://mic-ontology.herokuapp.com/Country/United_Kingdom> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://dbpedia.org/ontology/Country> .  
<https://mic-ontology.herokuapp.com/Country/United_Kingdom> <http://www.w3.org/2000/01/rdf-schema#label> "United Kingdom"@en .  
<https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://dbpedia.org/ontology/Person> .  
<https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://mic-ontology.herokuapp.com/MovieDirector> .  
<https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> <http://www.w3.org/2000/01/rdf-schema#label> "Paul Laverty"@en .  
<https://mic-ontology.herokuapp.com/Film/1> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://dbpedia.org/ontology/Film> .  
<https://mic-ontology.herokuapp.com/Film/1> <http://dbpedia.org/ontology/releaseDate> "2012"@en .  
<https://mic-ontology.herokuapp.com/Film/1> <http://www.w3.org/2000/01/rdf-schema#label> "Parte Degli Angeli, La"@it .
```

N3



```
<https://mic-ontology.herokuapp.com/Film/1> a <http://dbpedia.org/ontology/Film> ;  
    <http://dbpedia.org/ontology/releaseDate> "2012"@en ;  
    <http://www.w3.org/2000/01/rdf-schema#label> "Parte Degli Angeli, La"@it ;  
    <http://dbpedia.org/property/country> <https://mic-ontology.herokuapp.com/Country/United_Kingdom> ;  
    <http://dbpedia.org/ontology/director> <https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> .  
  
<https://mic-ontology.herokuapp.com/Country/United_Kingdom> a <http://dbpedia.org/ontology/Country> ;  
    <http://www.w3.org/2000/01/rdf-schema#label> "United Kingdom"@en .  
  
<https://mic-ontology.herokuapp.com/MovieDirector/Paul_Laverty> a <http://dbpedia.org/ontology/Person> , <https://mic-ontology.herokuapp.com/MovieDirector> ;  
    <http://www.w3.org/2000/01/rdf-schema#label> "Paul Laverty"@en .
```

4- DATA LINKING (Silk)

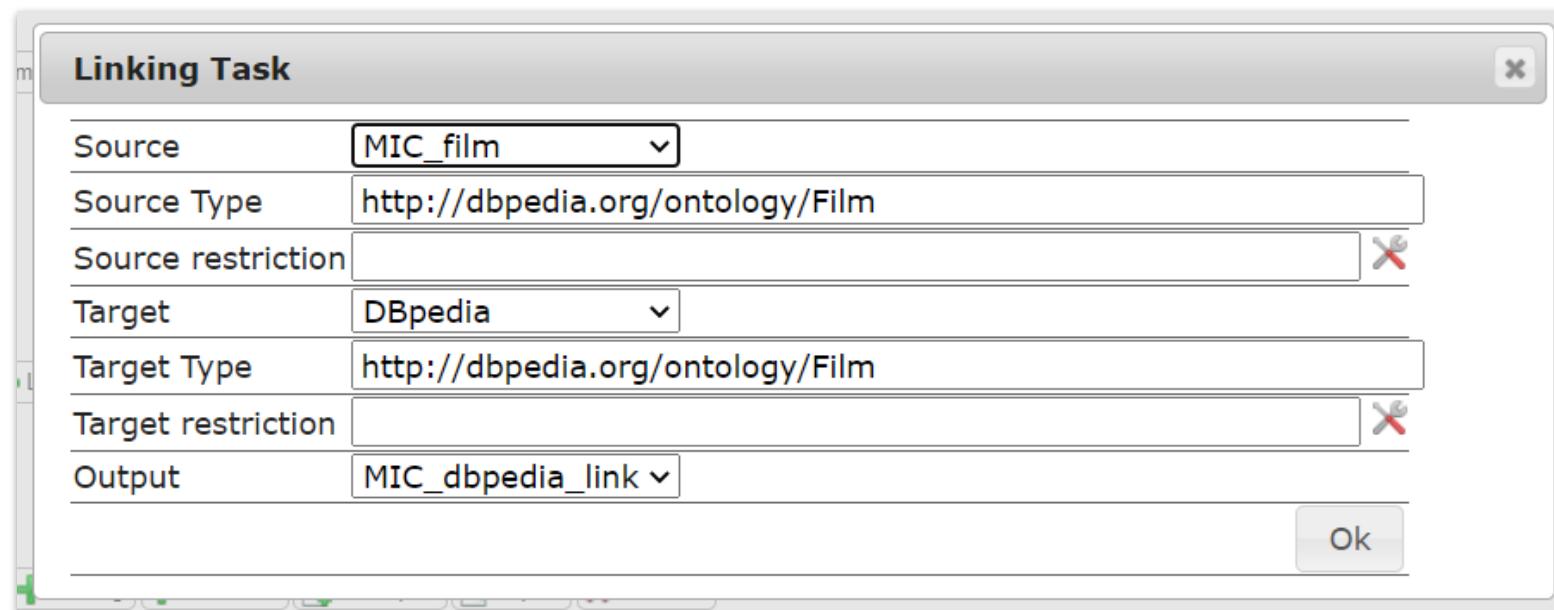
Silk Workbench

4- SILK: prefissi e risorse

The screenshot displays the SILK interface with the following components:

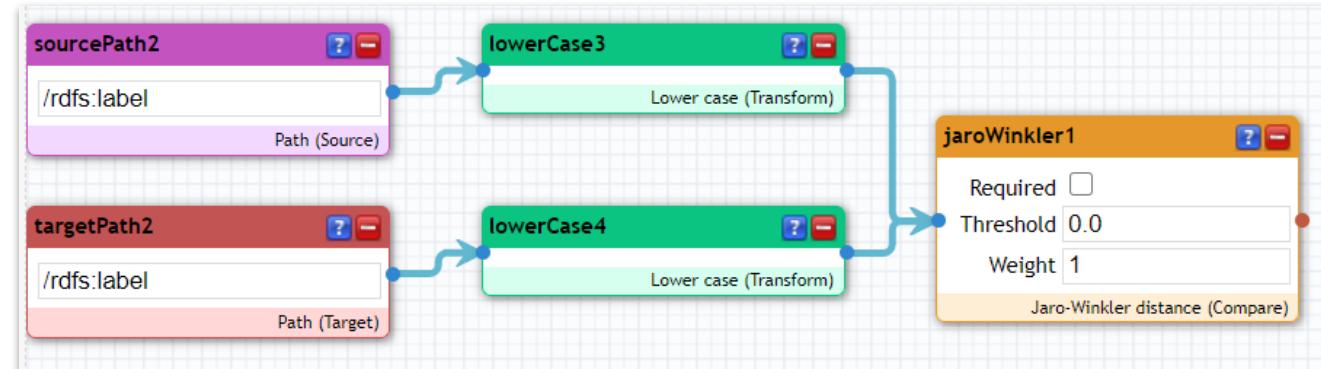
- Project Structure:** A tree view showing the project `micOntology` containing three resources: `DBpedia`, `MIC_dbpedia_link`, and `MIC_film`. Each resource has a `Properties`, `Open`, and `Remove` button.
- Edit Prefixes Dialog:** A modal window titled "Edit Prefixes" listing the following prefixes:
 - `rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#` (with a Remove button)
 - `dbpp: http://dbpedia.org/property/` (with a Remove button)
 - `dbo: http://dbpedia.org/ontology/` (with a Remove button)
 - `owl: http://www.w3.org/2002/07/owl#` (with a Remove button)
 - `rdfs: http://www.w3.org/2000/01/rdf-schema#` (with a Remove button)
 - `mic: https://mic-ontology.herokuapp.com/` (with a Remove button)A "Add" button is at the bottom right, and an "Ok" button is at the bottom left.
- Manage Resources Dialog:** A modal window titled "Manage Resources" containing sections for Available resources, Upload input resource, Import input resource from URL, and New output resource.
 - Available resources:** Lists `MIC_dbpedia_link` and `MIC_Film_Output.ttl`, each with a Remove button.
 - Upload input resource:** Fields for Name (set to "unnamed") and File (button labeled "Scegli file Nessun file selezionato"), with an "Upload" button.
 - Import input resource from URL:** Fields for Name (set to "unnamed") and URL, with an "Import" button.
 - New output resource:** Fields for Name (set to "output") and Create button.A "Close" button is at the bottom right.

4- SILK: linking task



4- SILK: linking pipeline 1

- Compariamo solo le label dei film.



Source Entity	Target Entity	Link Status
mic:Film/1122	http://dbpedia.org/resource/Blade_II	100,0%
mic:Film/1137	http://dbpedia.org/resource/Blood_Brothers_(2007_drama_film)	100,0%
mic:Film/1136	http://dbpedia.org/resource/Blood_and_Wine	100,0%
mic:Film/9341	http://dbpedia.org/resource/Blonde_Venus	100,0%
mic:Film/1138	http://dbpedia.org/resource/Blood_Trails	100,0%
mic:Film/1146	http://dbpedia.org/resource/Blue_Jasmine	100,0%
mic:Film/1143	http://dbpedia.org/resource/Blowup	100,0%
mic:Film/1142	http://dbpedia.org/resource/Blown_Away_(1994_film)	100,0%
mic:Film/1141	http://dbpedia.org/resource/Blow_(film)	100,0%
mic:Film/851	http://dbpedia.org/resource/Blowing_Wild	100,0%

Comparison: jaroWinkler (jaroWinkler1) 100,0%

Transform: lowerCase (lowerCase3) ballata selvaggia

Input: rdfs:label (sourcePath2) Ballata Selvaggia

Transform: lowerCase (lowerCase4) wilde glut le souffle sauvage ballata selvaggia blowing wild blowing wild

Input: rdfs:label (targetPath2) Wilde Glut Le Souffle sauvage Ballata selvaggia Blowing Wild Blowing Wild

Prev 1 2 3 Next

Linking Statistics

Number of source entities: 9734
Number of target entities: 15703
Number of links: 286

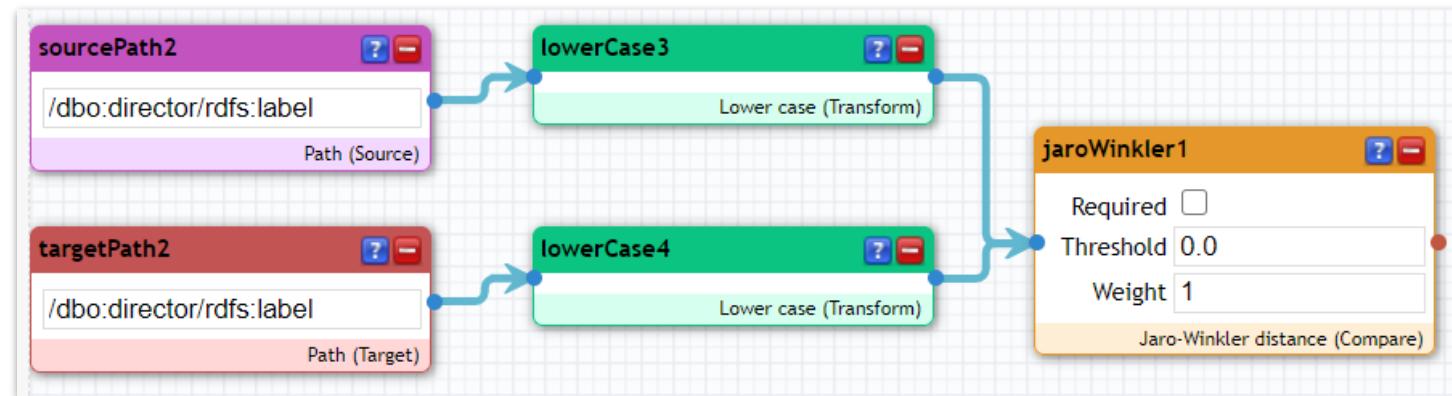
Matching: 1372 tasks 203 links. (74,5%)

Start

Stop

Config

4- SILK: linking pipeline 2



- Compariamo solo le label dei registi.

Source Entity	Target Entity	Score	Actions
mic:Film/8816	http://dbpedia.org/resource/10_30_PM_Summer	100,0%	<input checked="" type="checkbox"/> ? X
mic:Film/6674	http://dbpedia.org/resource/12_(2007_film)	100,0%	<input checked="" type="checkbox"/> ? X
mic:Film/3320	http://dbpedia.org/resource/008_Operation_Exterminate	100,0%	<input checked="" type="checkbox"/> ? X
mic:Film/6833	http://dbpedia.org/resource/101_Dalmatians_(1996_film)	100,0%	<input checked="" type="checkbox"/> ? X
mic:Film/723	http://dbpedia.org/resource/12 Angry_Men_(1957_film)	100,0%	<input checked="" type="checkbox"/> ? X
mic:Film/1169	http://dbpedia.org/resource/13_Assassins	100,0%	<input checked="" type="checkbox"/> ? X

Comparison: jaroWinkler (jaroWinkler1) 100,0%
Transform: lowerCase (lowerCase3) (takashi miike)
Input: dbo:director/rdfs:label (sourcePath2) Takashi Miike
Transform: lowerCase (lowerCase4) (takashi miike) takashi miike takashi miike takashi miike takashi miike 三池崇史 takashi miike takashi miike takashi miike takashi miike ミニク, タカシ 三池崇史
Input: dbo:director/rdfs:label (targetPath2) Takashi Miike Takashi Miike Takashi Miike Takashi Miike Takashi Miike 三池崇史 Takashi Miike Takashi Miike Takashi Miike ミニク, タカシ 三池崇史

Prev 1 2 3 4 5 6 7 8 9 10 11 Next

Linking Statistics
Number of source entities: 9734
Number of target entities: 11471
Number of links: 7197

Matching: 1232 tasks 7028 links. (90,9%)

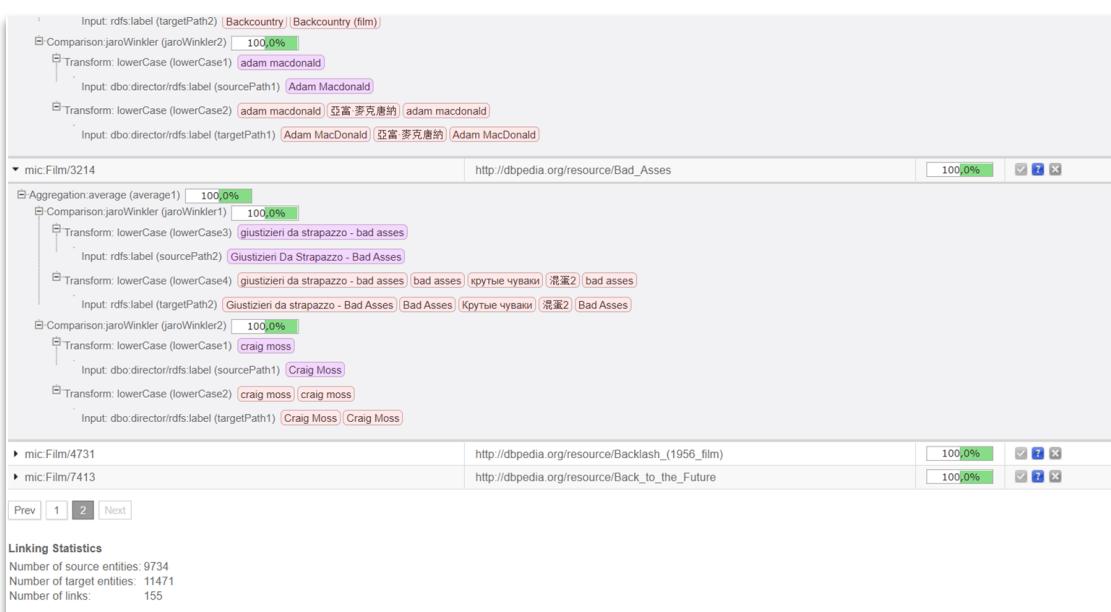
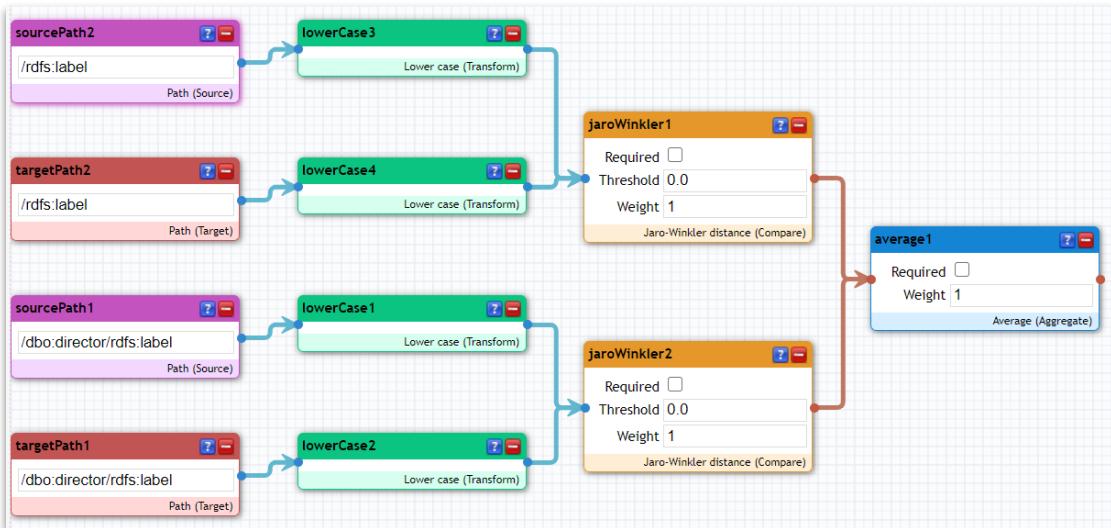
Start

Stop

Config

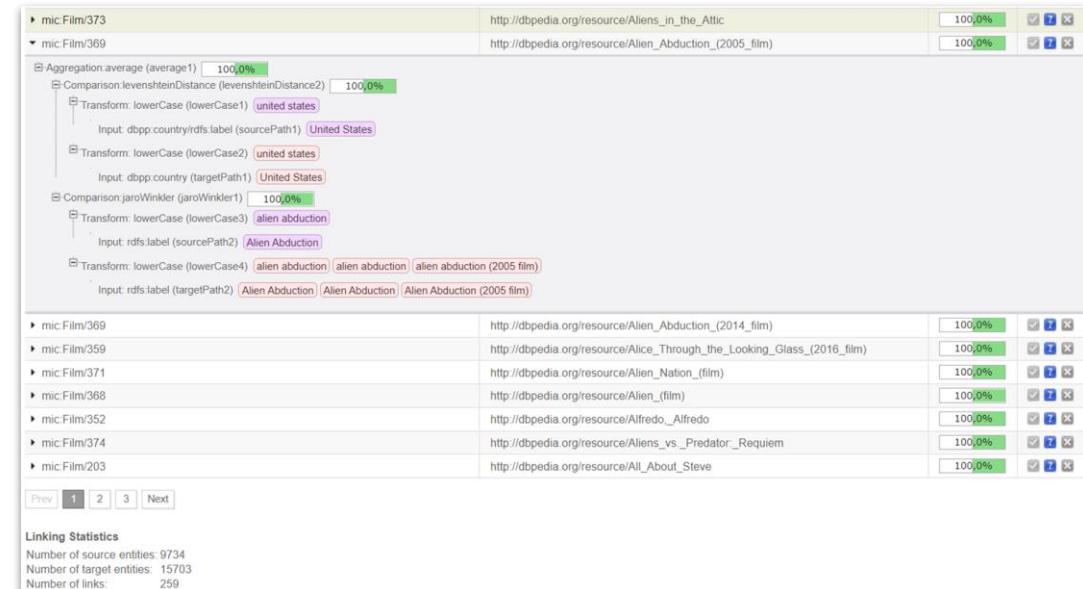
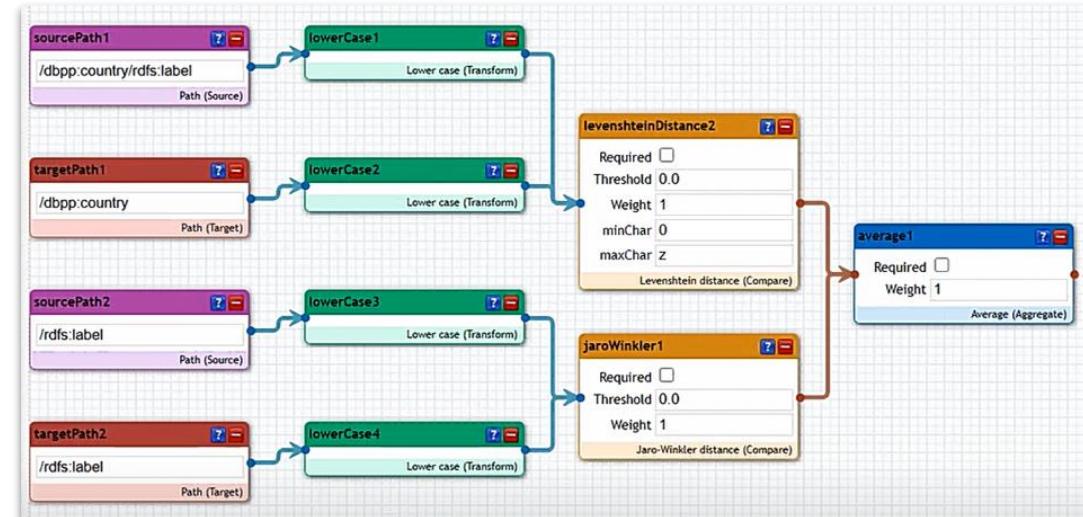
4- SILK: linking pipeline 3

- Compariamo la label dei film e quella dei registi.



4- SILK: linking pipeline 4

- Compariamo la label dei film e i paesi di produzione.



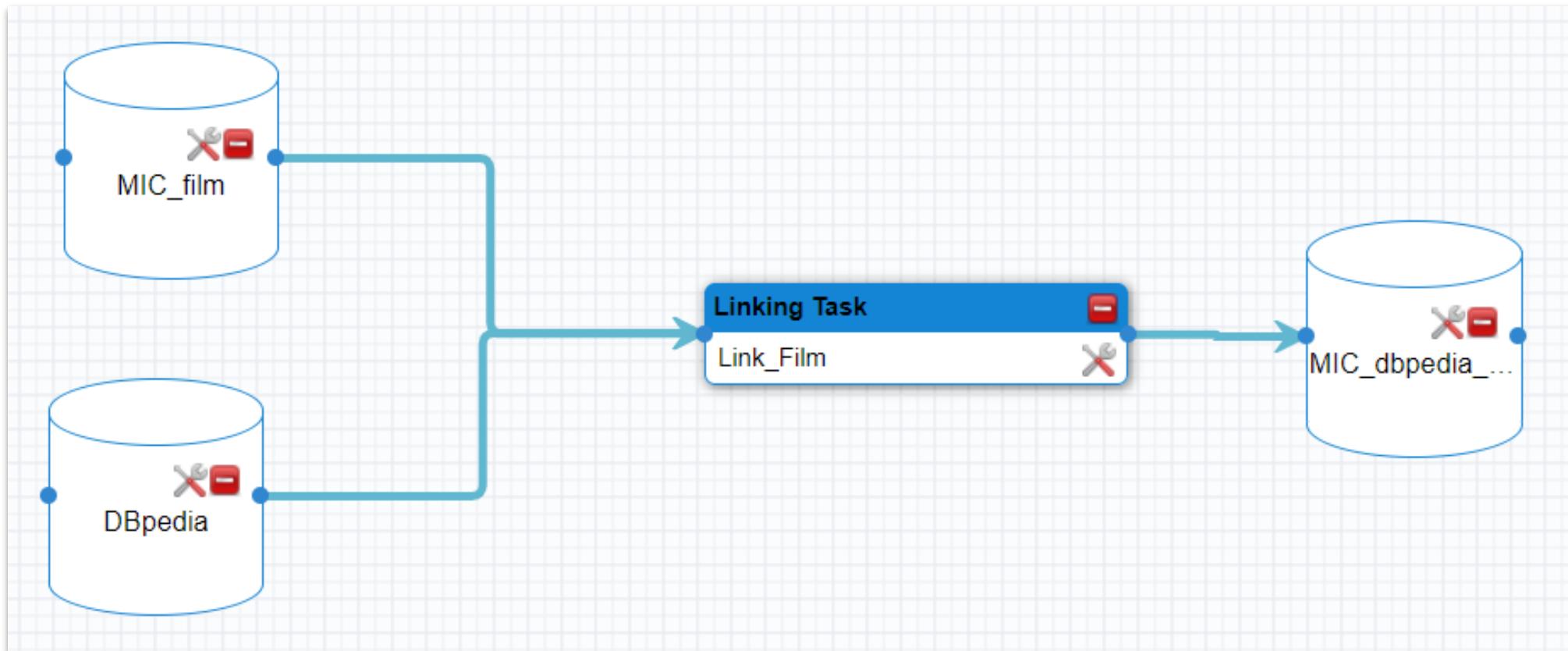
4- SILK: workflow

- Errore di salvataggio del Workflow.

localhost:9000 dice

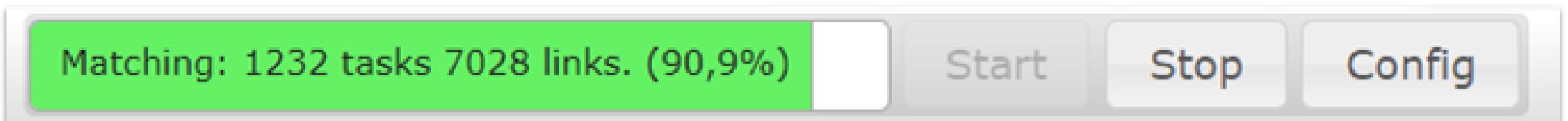
Error committing workflow to backend: For input string:
"372.00000762939453"

OK



4- SILK: conclusioni

- Vengono trovate tutte le 10 mila entità (circa) del file MIC_File_Output.ttl;
- Vengono trovate le entità in DBpedia;
- Il numero di match (`dbo:film/rdfs:label owl:sameAs dbo:film/rdfs:label`) è molto inferiore rispetto al numero totale di entità (generando circa 290 link su 10 mila entità);
- Facendo una media anche con le Nazioni (`dbpp:Country owl:sameAs dbo:Country`) vengono creati 259 link;
- Il risultato non cambia con il variare dei comparatori (quelli testati: Levenshtein, Jaccard, Jaro Distance, Jaro Winkler) o della threshold;
- I link restituiti si fermano alla lettera b (circa 900 righe) → problema dell'endpoint di DBpedia.

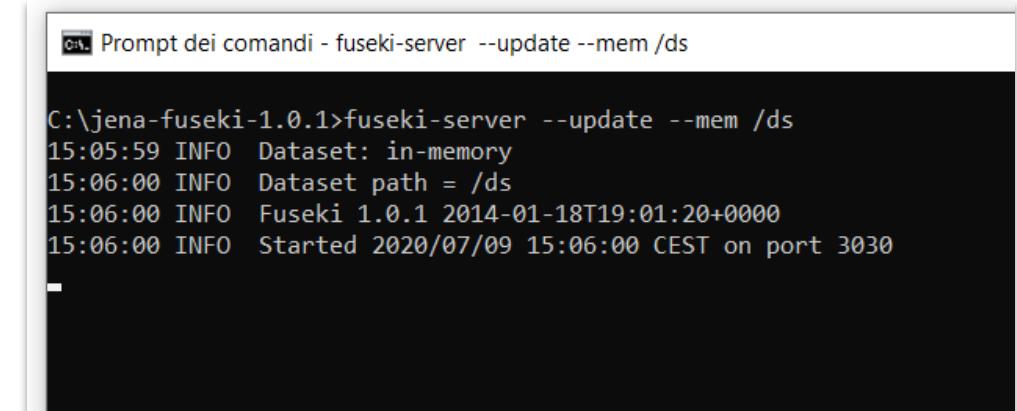


5- SPARQL QUERY (Jena Fuseki)



5- SPARQL QUERY: installazione Jena Fuseki

- Download Jena Fuseki dal sito di [Apache Jena](#);
- Dopo aver riscontrato dei problemi nell'installazione dell'ultima versione si è optato per una vecchia versione (la 1.0.1), trovata sempre sul sito di [Apache Jena](#);
- Dopo aver estratto la cartella zippata si avvia il tool tramite il comando “fuseki-server --update --mem /ds”.



```
Prompt dei comandi - fuseki-server --update --mem /ds
C:\jena-fuseki-1.0.1>fuseki-server --update --mem /ds
15:05:59 INFO Dataset: in-memory
15:06:00 INFO Dataset path = /ds
15:06:00 INFO Fuseki 1.0.1 2014-01-18T19:01:20+0000
15:06:00 INFO Started 2020/07/09 15:06:00 CEST on port 3030
```

5- SPARQL QUERY: caricamento dataset

- Una volta avviato, col browser ci si reca a <http://localhost:3030/>;
- Nel control Panel si seleziona il file locale che vogliamo interrogare;
- Si vede il numero di triple contenute nel file.

The screenshot shows the Fuseki Control Panel interface. At the top, there's a navigation bar with links for Server Management, Documentation, and Validators. Below that is a section for the General SPARQL Service with links to SPARQL query form and standards like SPARQL 1.1 Query, Update, and Protocol. A large "Success" message is prominently displayed in the center. To the right, a file browser window is open, showing a list of files with details like name, date, type, and size. One file, "MIC_Film_Output.nt", is highlighted. In the foreground, there are two main panels: "SPARQL Query" and "SPARQL Update". The "SPARQL Query" panel has a text input field for the query and a "Get Results" button. The "SPARQL Update" panel has a "Perform update" button. At the bottom, there's a "File upload" section with a "Scegli file" button and an "Upload" button. A "Back to Fuseki" button is also visible.

Tipo	Dimensione
Cartella di file	59.512 KB
Executable Jar File	10.424 KB
File N3	17.319 KB
File NT	10.424 KB
File TTL	616 KB
Foglio di lavoro di...	508 KB
File con valori sep...	775 KB
Foglio di lavoro di...	749 KB
File con valori sep...	1.075 KB
Executable Jar File	59.512 KB

5- SPARQL QUERY: query

- Nel control panel si trova la sezione per eseguire le SPARQL Query;
- Importante è l'attribuzione dei prefissi non ancora esistenti, quelli della sorgente dati del MIC.

SPARQL Query

```
PREFIX micReg: <https://mic-ontology.herokuapp.com/MovieDirector/>
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?titolo ?film
WHERE {
    ?film dbo:director micReg:Lana_Wachowski, micReg:Lilly_Wachowski .
    ?film rdfs:label ?titolo .
}
```

Output:

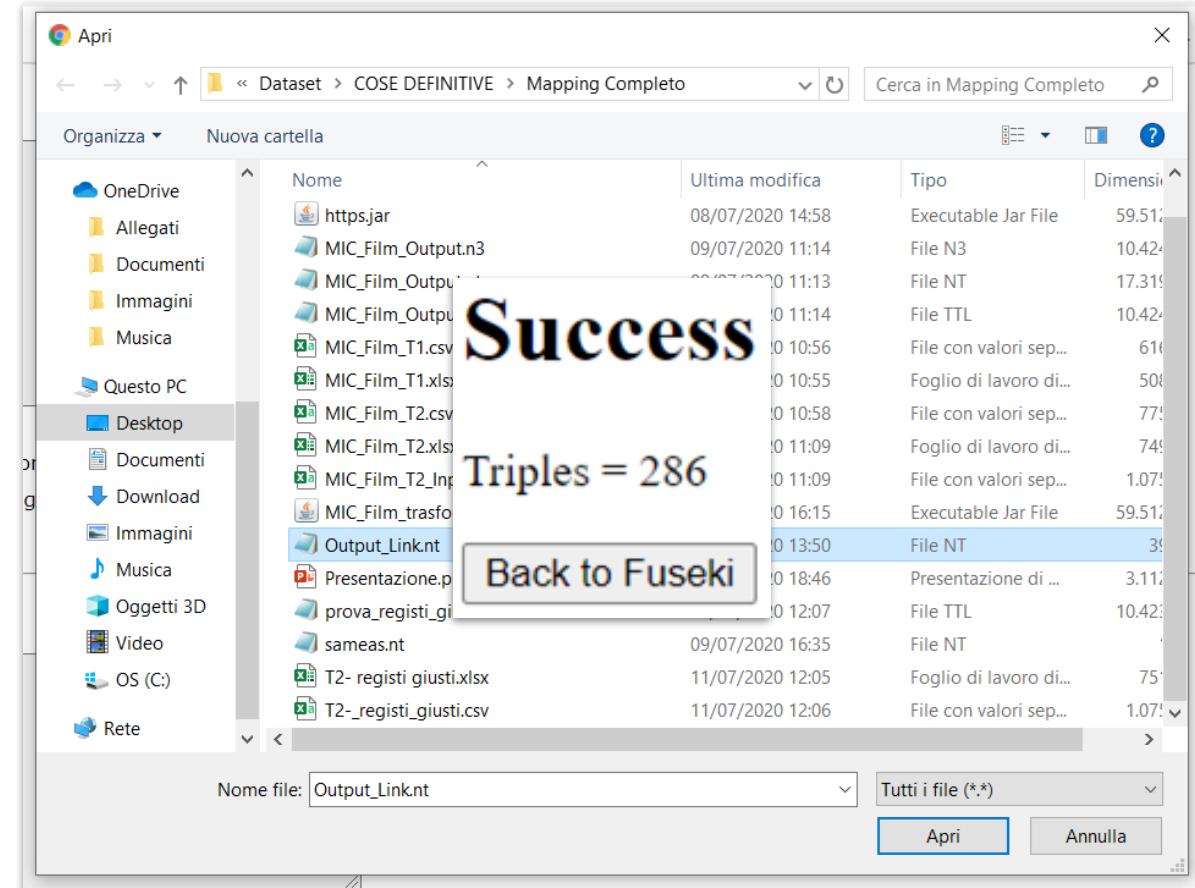
If XML output, add XSLT style sheet (blank for none):

Force the accept header to text/plain regardless.

titolo	film
"Cloud Atlas"@it	< https://mic-ontology.herokuapp.com/Film/1702 >
"Matrix Revolutions"@it	< https://mic-ontology.herokuapp.com/Film/5769 >
"Matrix Reloaded"@it	< https://mic-ontology.herokuapp.com/Film/5768 >
"Speed Racer"@it	< https://mic-ontology.herokuapp.com/Film/8109 >
"Matrix"@it	< https://mic-ontology.herokuapp.com/Film/5767 >
"Jupiter - Il Destino Dell'Universo"@it	< https://mic-ontology.herokuapp.com/Film/4499 >
"Jupiter - Il Destino"@it	< https://mic-ontology.herokuapp.com/Film/4498 >

5- SPARQL QUERY: caricamento documento link

- Caricamento del secondo documento (Output_Link.nt), contenente i link sameAs ottenuti con Silk;
- Jena Fuseki segnala 286 nuove triple caricate.



5- SPARQL QUERY: query con owl:sameAs

- A questo punto si possono effettuare delle query sfruttando i link sameAs e andare a recuperare dei dati non presenti nel dataset iniziale.

SPARQL Query

```
PREFIX micNat: <https://mic-ontology.herokuapp.com/Country/>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX dbp: <http://dbpedia.org/property/>
PREFIX dbr: <http://dbpedia.org/resource/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?mic_Id ?title ?c ?actors
WHERE {
    ?mic_Id dbo:country micNat:United_States ;
        rdfs:label ?title ;
        owl:sameAs ?c .
SERVICE <http://dbpedia.org/sparql>
{
    ?c dbo:starring ?attori .
    ?attori rdfs:label ?actors .
    FILTER (lang(?actors) = 'en')
}
ORDER BY ?title
LIMIT 20
```

Output:

If XML output, add XSLT style sheet (blank for none):

Force the accept header to text/plain regardless.

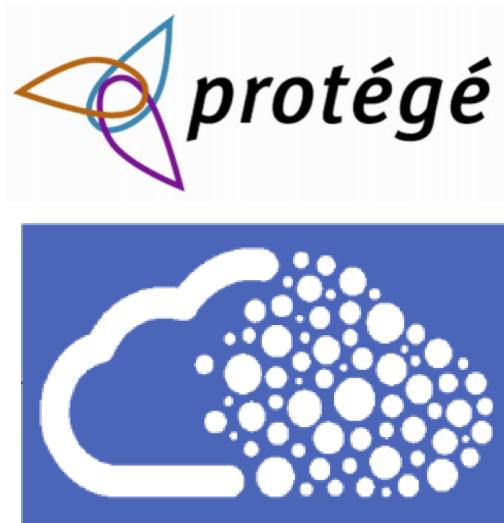
mic_Id	title	c	actors
micFilm:24	"10 Cloverfield Lane"@it	dbr:10_Cloverfield_Lane	"John Gallagher, Jr."@en
micFilm:24	"10 Cloverfield Lane"@it	dbr:10_Cloverfield_Lane	"John Goodman"@en
micFilm:24	"10 Cloverfield Lane"@it	dbr:10_Cloverfield_Lane	"Mary Elizabeth Winstead"@en
micFilm:27	"10.000 Ac"@it	<http://dbpedia.org/resource/10,000_BC_(film)>	"Affif Ben Badra"@en
micFilm:27	"10.000 Ac"@it	<http://dbpedia.org/resource/10,000_BC_(film)>	"Camilla Belle"@en
micFilm:27	"10.000 Ac"@it	<http://dbpedia.org/resource/10,000_BC_(film)>	"Cliff Curtis"@en
micFilm:27	"10.000 Ac"@it	<http://dbpedia.org/resource/10,000_BC_(film)>	"Marco Khan"@en
micFilm:27	"10.000 Ac"@it	<http://dbpedia.org/resource/10,000_BC_(film)>	"Steven Strait"@en
micFilm:27	"10.000 Ac"@it	<http://dbpedia.org/resource/10,000_BC_(film)>	"Tim Barlow"@en
micFilm:34	"12 Round"@it	<http://dbpedia.org/resource/12_Rounds_(film)>	"Aidan Gillen"@en
micFilm:34	"12 Round"@it	<http://dbpedia.org/resource/12_Rounds_(film)>	"Ashley Scott"@en
micFilm:34	"12 Round"@it	<http://dbpedia.org/resource/12_Rounds_(film)>	"Billy Slaughter"@en
micFilm:34	"12 Round"@it	<http://dbpedia.org/resource/12_Rounds_(film)>	"Brian J. White"@en
micFilm:34	"12 Round"@it	<http://dbpedia.org/resource/12_Rounds_(film)>	"John Cena"@en
micFilm:34	"12 Round"@it	<http://dbpedia.org/resource/12_Rounds_(film)>	"Steve Harris (actor)"@en
micFilm:34	"12 Round"@it	<http://dbpedia.org/resource/12_Rounds_(film)>	"Taylor Cole"@en
micFilm:35	"127 Ore"@it	dbr:127_Hours	"James Franco"@en
micFilm:40	"17 Again"@it	<http://dbpedia.org/resource/17_Again_(film)>	"Leslie Mann"@en
micFilm:40	"17 Again"@it	<http://dbpedia.org/resource/17_Again_(film)>	"Matthew Perry"@en
micFilm:40	"17 Again"@it	<http://dbpedia.org/resource/17_Again_(film)>	"Michelle Trachtenberg"@en

6- COMMENTI

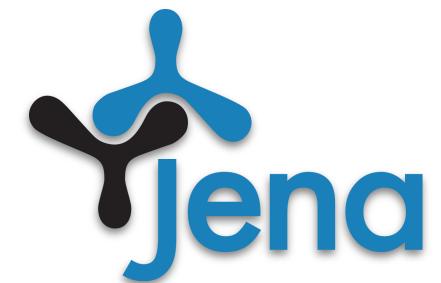


6- Commenti 1

Sono stati riscontrati dei problemi con tutti i tool:



Silk Workbench



6- Commenti 2



- Grande affidabilità;
- Facilità di utilizzo;
- Ampia documentazione online;
- Rapidità del programma anche con tanti dati;
- Tante funzioni per automatizzare la pulizia dei documenti.

- Problemi con l'esportazione in csv dei documenti (non si può selezionare il carattere separatore);
- Problemi con importazione dei csv da Windows.

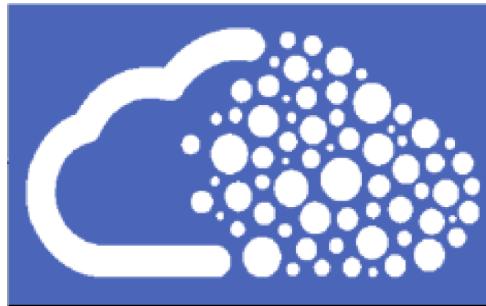
6- Commenti 3



- Interfaccia intuitiva;
- Plug-in mapper già incluso;
- Possibilità di costruire un'ontologia e poi importare i dati;
- Possibilità di separare trasformazioni da dataset (riutilizzo trasformazioni).

- Il mapper ha problemi con grandi quantità di dati;
- Il mapper ha problemi con la ripetizione di entità;
- Il mapper non permette di fare operazioni di pulizia sui dati.

6- Commenti 4



- Possibilità di lavorare su grandi quantità di dati (scaricando il Jar);
- Possibilità di effettuare operazioni di pulizia dei dati;
- Possibilità di separare trasformazioni da dataset (riutilizzo trasformazioni).
- Interfaccia spesso ambigua;
- Mancanza di feedback;
- Lentezza;
- Vari bug;
- Problemi di gestione dei datatype (possibile problema di comunicazione col backend).

6- Commenti 5

Silk Workbench



- Interfaccia abbastanza intuitiva.
- Lentezza anche in computer con grandi risorse computazionali;
- Limitazioni stringenti su dimensioni dei file da caricare;
- Svariati errori dati dal tool durante l'utilizzo;
- Errore di salvataggio del workflow;
- Limite di circa 900 righe;

6- Commenti 6



- Interfaccia intuitiva;
- Facilità di utilizzo;
- Possibilità di fare query sul dataset da locale;
- Possibilità di fare query su più dataset da locale.

- Problemi nell'installazione dell'ultima versione (risolti con l'installazione della versione 1.0.1).

GRAZIE PER L'ATTENZIONE!