

In and out: Work flows between library data and linked-data at the National Library of Spain

Semantic Web in Libraries
November, 25-27th, 2019, Hamburg

Ricardo Santos
National Library of Spain



@3186Ricardo
#datosbne

WHAT IS DATOS.BNE?

A fully functioning linked-data based gateway into the library resources.

- Library resources (as any other catalogue)
- Data about authors
- Data about vocabularies
- Data about items
- Connected to usual library services (digital library items, loan services, document reproduction services)

MAIN FEATURES

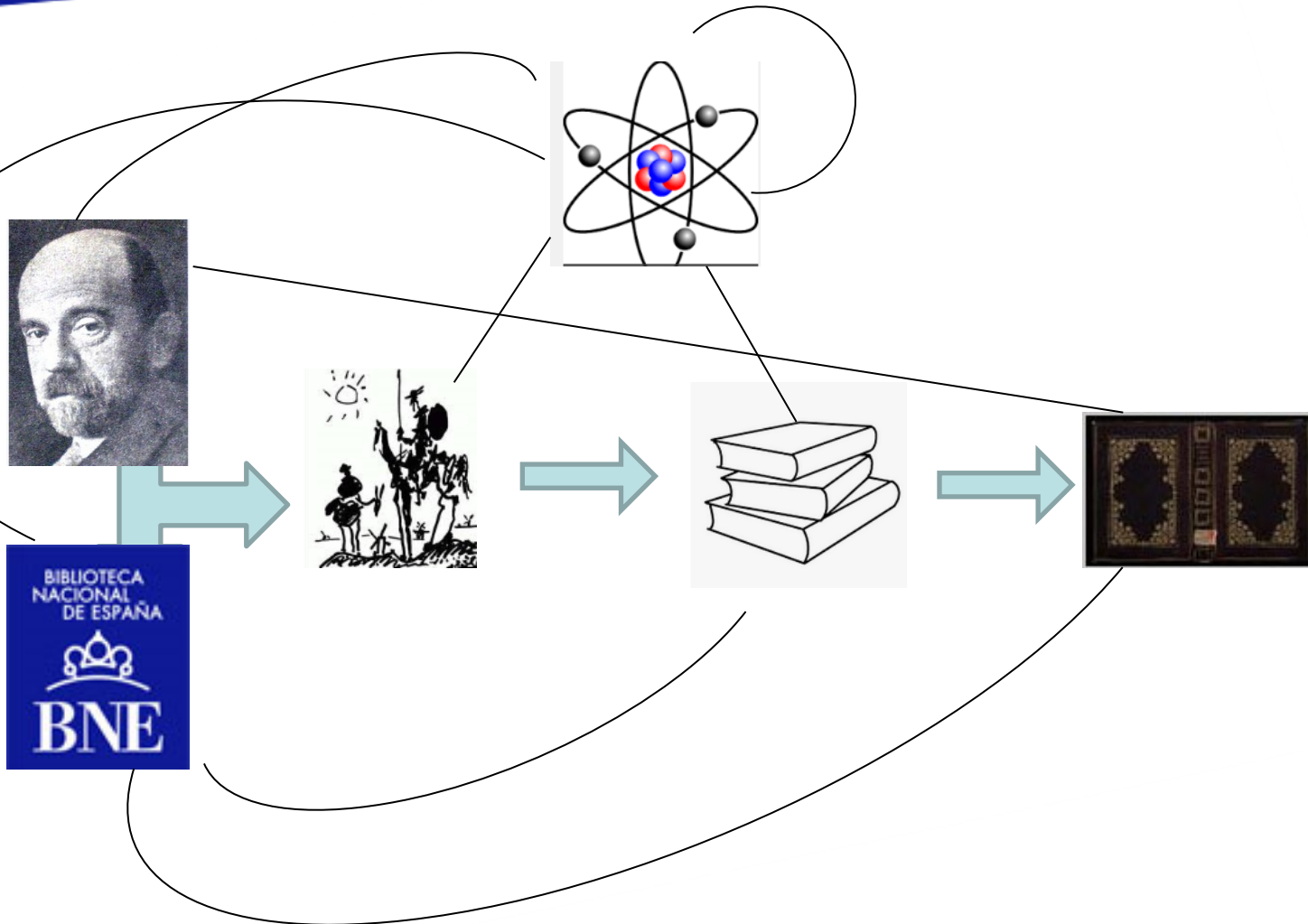
FRBR-Based

- Entity based
- Entity driven search
- Entity-based search results ranking

Linked and enriched

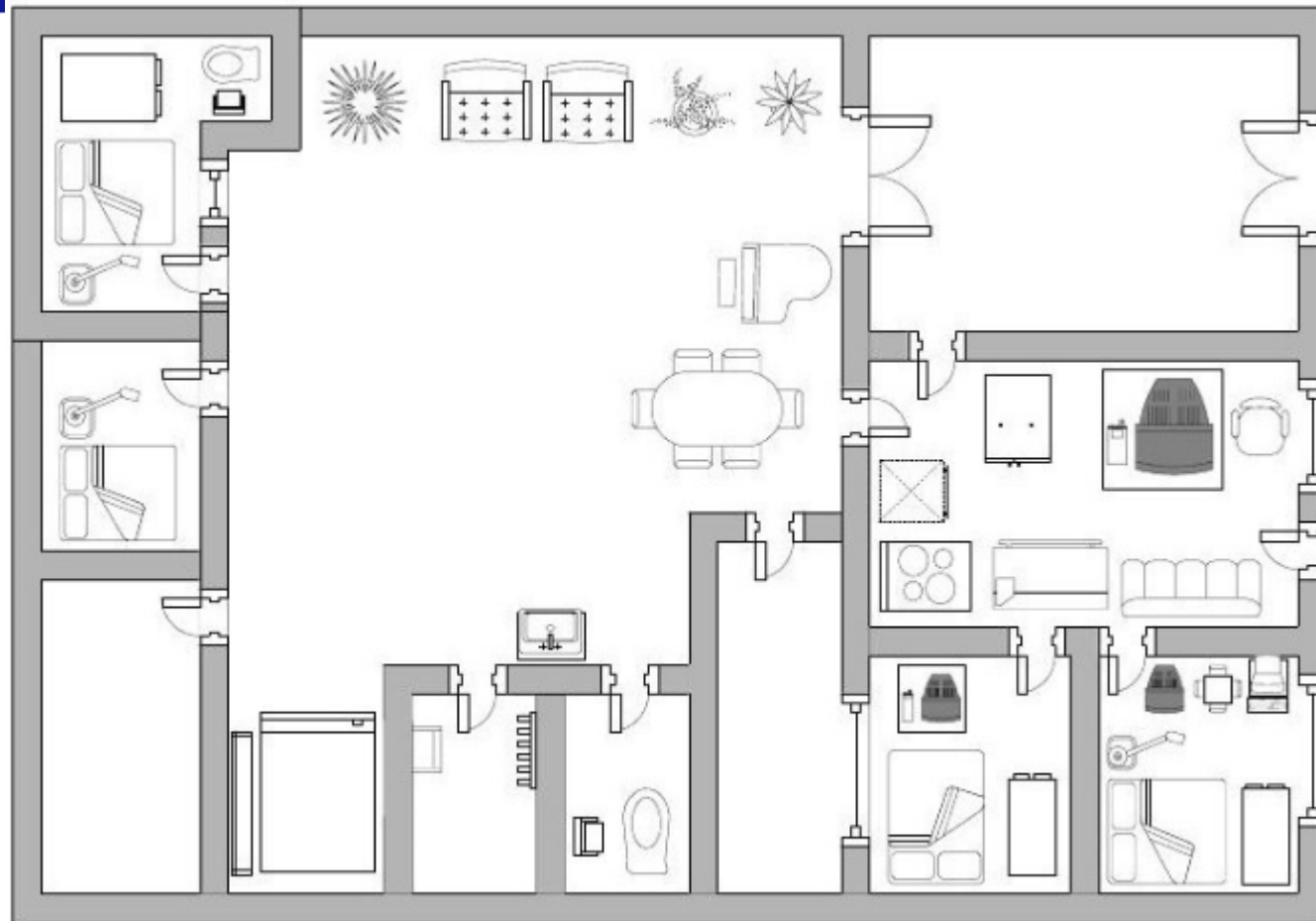
- Interlinks between FRBR entities
- Linked to prominent equivalent concepts
- Enriched with outside data

BASIC DATA MODEL



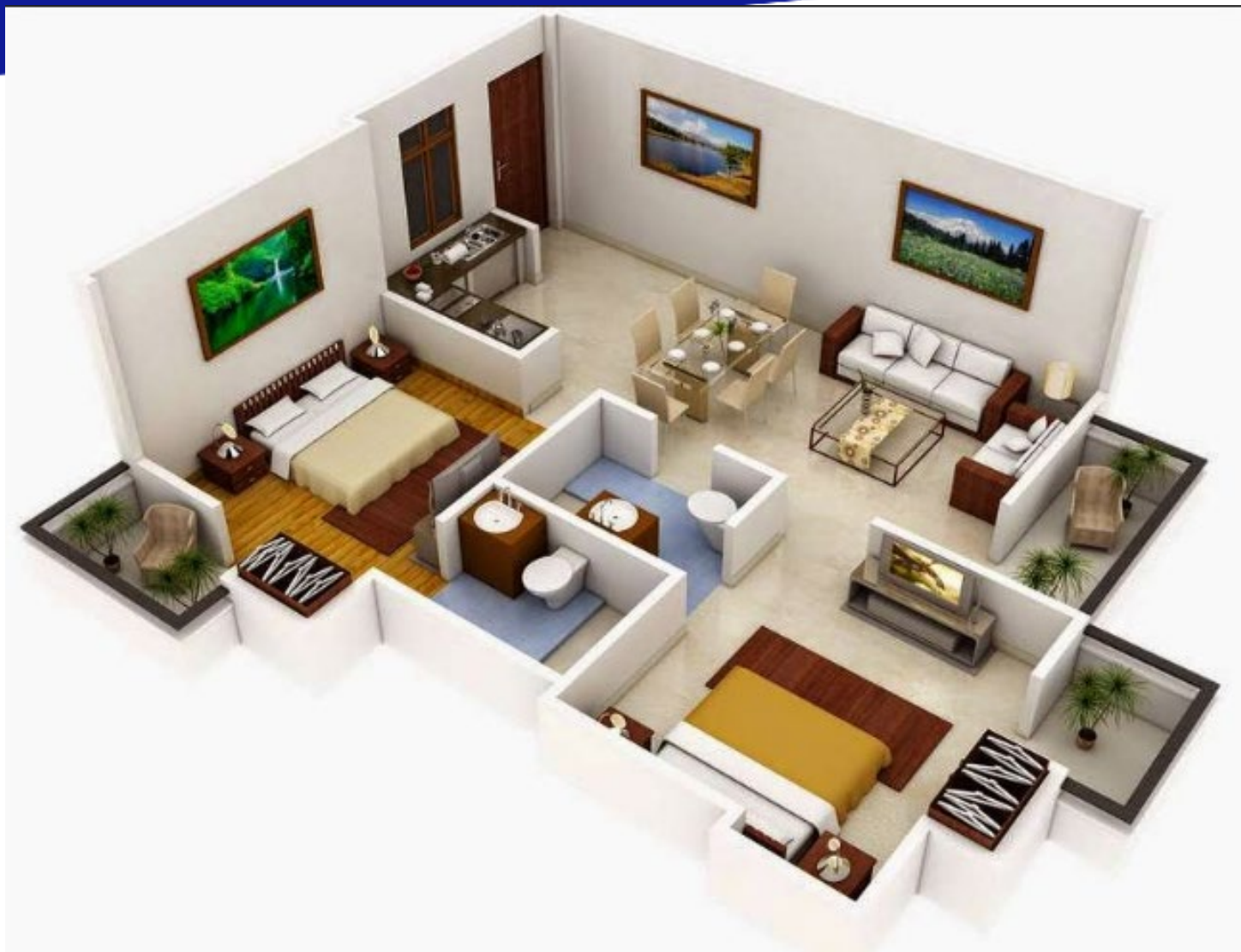
From MARC to FRBR entities

5



From a 2-D structure to....

From MARC to FRBR entities



Behind the scenes

Data
generation

Data (Marc)

MARiMbA

Data analysis
Mapping & sorting
Linked Data generation
Linking and enrichment

Publication
pipeline

Indexing (1)

Ranking based in FRBR
relationships

Double storing

Front-end (2)
Triple store (3)

- (1) Elasticsearch
- (2) MongoDB
- (3) Virtuoso

Peeking inside Marimba

Ingestion and analysis of
MARC21 records

Mapping 1: sorting

Mapping 2: model entity linking

Mapping 3: data properties

RDF generation

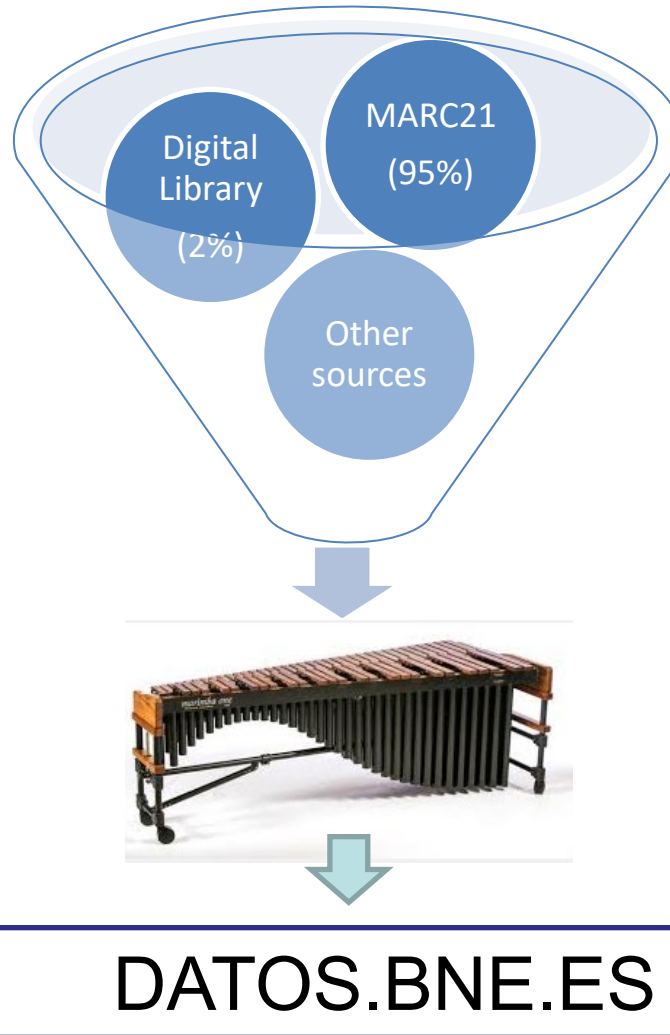


Data properties are based on

subfields:

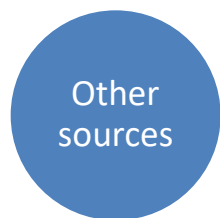
260 \$a → P3003 (place of
publication)

Dataflows



Dataflows

1
0



Identifiers from VIAF

Links: sameAs

Enlaces relacionados	http://es.wikipedia.org/wiki/Johann_Wolfgang_von_Goethe
Biblioteca Nacional Francesa	http://data.bnf.fr/11905269
Biblioteca Nacional Alemana	http://d-nb.info/gnd/118540238
VIAF	http://viaf.org/viaf/24602065
Biblioteca del Congreso de los EEUU	http://id.loc.gov/authorities/names/n79003362
Libris (Suecia)	http://libris.kb.se/resource/auth/56318
Sudoc	http://www.idref.fr/026895528
DBpedia	http://dbpedia.org/resource/Johann_Wolfgang_von_Goethe

Dataflows

Goethe, Johann Wolfgang von (1749-1832)

Johann Wolfgang von Goethe (AFI: ['jo:han 'vɔlfɡaŋ fɔn 'ɡø:tə]; Fráncfort del Meno, 28 de agosto de 1749-Weimar, 22 de marzo de 1832) fue un poeta, novelista, dramaturgo y científico germánico, contribuyente fundamental del Romanticismo, movimiento al que influenció profundamente. En palabras de George Eliot (1819-1880) fue «el más grande hombre de letras alemán... y el último verdadero hombre universal que caminó sobre la tierra». Su obra, que abarca géneros como la novela, la poesía lírica, el drama e incluso controvertidos tratados científicos, dejó una profunda huella en importantes escritores, compositores, pensadores y artistas posteriores, siendo incalculable en la filosofía alemana posterior y constante fuente de inspiración para todo tipo de obras. Su novela Wilhelm Meister fue citada por Arthur Schopenhauer como una de las cuatro mejores novelas jamás escritas, junto con Tristram Shandy, La Nouvelle Héloïse y Don Quijote. Su apellido da nombre al Goethe-Institut, organismo encargado de difundir la cultura alemana en todo el mundo.

Información extraída de Wikipedia



Other
sources

→ Identifiers from VIAF

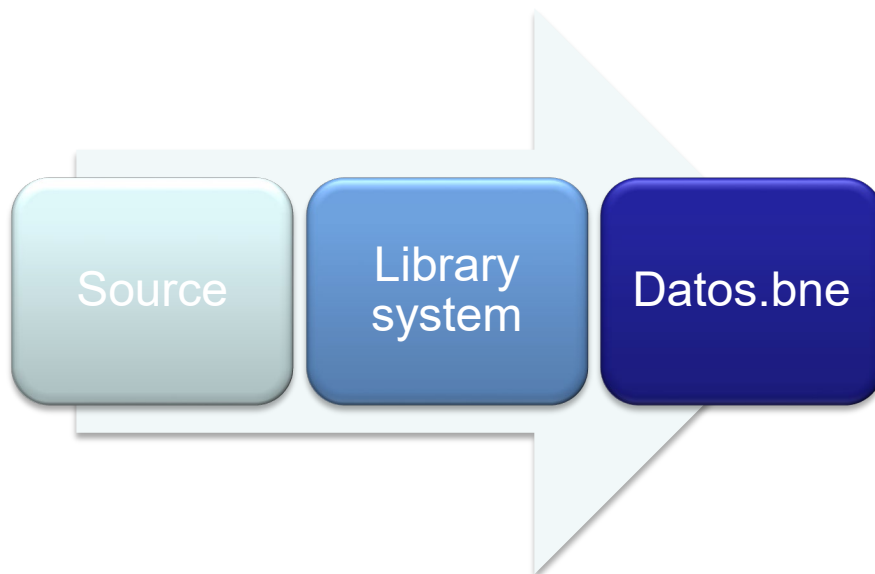
→ Dbpedia

THEY REMAIN ONLY IN DATOS.BNE

Enrichments and data imports

Main enrichment: EVERYDAY JOB

Other enrichments and imports:



Sources: Libraries and beyond

Data imported: identifiers, equivalencies and attributes

2019: Wikidata

Achievement: more than 80.000 persons authority records enriched with data from Wikidata

Mapping BNE Ids with Wikidata IDs, using VIAF identifier datadump

`void:dataDump`

<http://viaf.org/viaf/data/viaf-20191104-links.txt.gz>

File Details

`rdfs:comment`

This is a text file showing the correspondence between source IDs in clusters,

Extracting selected properties (in Spanish where available), using the Wikidata API:

- Retrieved selected properties from every ID in JSON
- Sorting and grouping
- Convert to MARC21 fields according to mapping

Performed by datosbne technological partner, [Memorandum Multimedia](#)

2019: Wikidata

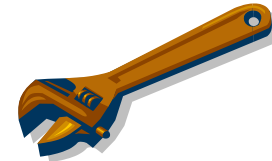
1
4

Quality check-ups:

General overview → *disregard property P737 (influenced by)*

Work on property *Occupation (P106)* :

- Sorting, grouping and arranging for revision by vocabulary experts.
- Decision on every value:
 - Keep it
 - Align it with BNE Subject Vocabulary
 - Disregard it.



2019: Wikidata

1
5

Occupation quality check-up:

0	Anticuario	Anticuarios
1	Antropología social	BORRAR
2	Antropólogo	Antropólogos
3	Antólogo	BORRAR
4	Apicultor	
5	Apneísta	
6	Apologeta	BORRAR
7	Apóstol	BORRAR
8	Arabista	Arabistas
9	Aracnólogo	Entomólogos
00	Archivero	Archiveros
01	Archivista paleógrafo	Archiveros
02	Aristócrata	BORRAR
03	Armada de los Estados Unidos	BORRAR

From Wikidata to datosbne

A funny thing happened on my way to datos.bne



WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

Tools

- What links here
- Related changes
- Upload file
- Special pages
- Permanent link
- Page information
- Wikidata item**

Article [Talk](#)

[Read](#) [Edit source](#) [View history](#) [☆](#)



This November is Wikipedia Asian month.
Join the contest and win a postcard from Asia.
[\[Help with translations!\]](#)

Ice-T

From Wikipedia, the free encyclopedia

This article is about the American rapper. For other uses, see [IceT](#).

Tracy Lauren Marrow (born February 16, 1958),^[1] better known by his stage name **Ice-T**, is an American musician, rapper, songwriter, actor, record producer, and author. He began his career as an underground rapper in the 1980s and was signed to [Sire Records](#) in 1987, when he released his debut album *Rhyme Pays*; the second hip-hop album to carry an [explicit content sticker](#) after [Slick Rick's](#) *La Di Da Di*. The following year, he founded the record label Rhyme \$yndicate Records (named after his collective of fellow hip-hop artists called the "Rhyme \$yndicate") and released another album, *Power*, which went on to go Platinum. He also released several other albums that went Gold.

He co-founded the [heavy metal](#) band [Body Count](#), which he introduced on his 1991 rap album *O.G.: Original Gangster*, on the track titled "Body Count". The band released their [self-titled debut album](#) in 1992. Ice-T encountered controversy over his track "[Cop Killer](#)", the lyrics of which discussed killing police officers. Ice-T asked to be released from his

Ice T



From Wikidata to datosbne

occupation

actor

▶ 1 reference

edit

rapper

▼ 0 references

edit

+ add reference

musician

▶ 1 reference

edit

television actor

▼ 0 references

edit

Related changes

Upload file

Special pages

Permanent link

Page information

Wikidata item

He co-founded the [heavy metal](#) band [Body Count](#), which he introduced on his 1991 rap album [O.G.: Original Gangster](#), on the track titled "Body Count". The band released their [self-titled debut album](#) in 1992. Ice-T encountered controversy over his track "[Cop Killer](#)", the lyrics of which discussed killing police officers. Ice-T asked to be released from his



Ice-T in *Marked Men* as the cop *Sam S. Cooper*

From Wikidata to datosbne

1
8

occupation

rapper (Q2252262)

artist who speaks rhymes in a rhythm over music

▼ In more languages

Configure

Language	Label	Description
English	rapper	artist who speaks rhymes in a rhythm over music
Spanish	rapero	artista que habla en rimas sobre un fondo musical
Catalan	raper	artista que parla amb rimes sobre un fons musical

▼ 0 references

Related changes

Upload file

Special pages

Permanent link

Page information

Wikidata item

He co-founded the [heavy metal](#) band [Body Count](#), which he introduced on his 1991 rap album [O.G.: Original Gangster](#), on the track titled "Body Count". The band released their [self-titled debut album](#) in 1992. Ice-T encountered controversy over his track "[Cop Killer](#)", the lyrics of which discussed killing police officers. Ice-T asked to be released from his

Add reference



Ice-T in [Mark Wahlberg](#) as the cop [Fleming "Son of Sam"](#)

From Wikidata to datosbne

1
9

A funny thing happened on my way to datos.bne

1399	Publiciste	Publicistas
1400	Publicitario	Publicistas
1401	Párroco	Sacerdotes
1402	Quiropráctico	
1403	Qutb	BORRAR
1404	Químico	Químicos
1405	Rabino	Rabinos
1406	Radiotelegrafista	
1407	Radiólogo	Radiólogos
1408	Rapero	Raperos
1409	Rapsoda	Poetas
1410	Rapsodia	BORRAR
1411	Reade	BORRAR
1412	Realizador	
1413	Rebe	BORRAR
1414	Recolector de música popular	Folkloristas
1415	Rector	Rectores universitarios
1416	Redactor de discursos	
1417	Redactor en jefe	Periodistas
1418	Reformador protestante	Protestantes

From Wikidata to datosbne

A funny thing happened on my way to datos.bne

1399	Publiciste	Publicistas
1400	Publicitaria	Publicistas
024	7	http://viaf.org/viaf/93602067 2viaf
024	7	https://www.wikidata.org/wiki/Q309941 2wikidata
024	7	http://www.isni.org/isni/0000000115782584 2isni
040		SpMaBN bspa cSpMaBN erdc fembne
046		f1958
100	1	Ice-T d1958-
370		Newark vWikidata
372		Gangsta rap vWikidata
373		Body Count vWikidata
374		Actores aPodcaster aActores aCantantes aMúsicos aRaperos aActores vWikidata
375		Masculino vWikidata
377		eng Ingles
400	1	Marrow, Tracy d1958-
1410	Redactor de discursos	
1417	Redactor en jefe	Periodistas
1418	Reformador protestante	Protestantes

From Wikidata to datosbne

A funny thing happened on my way to datos.bne

Nombre				Ordenar
Año de nacimiento (desde - hasta)				Año de fallecimiento (desde - hasta)
Lugar nacimiento		Lugar fallecimiento		
Títulos y otras características		Profesión/Ocupación		
		raperos		
Área de actividad		Institución asociada		

datos.bne Persons' advanced search

From Wikidata to datosbne

A funny thing happened on my way to datos.bne

Ramone, Dee Dee (1952-2002)

Punk,Hip hop
Bajistas,Músicos,Compositore
Fort Lee

Persona



Falco (1957-1998)

Pop,Rock,Hip hop
Músicos,Compositores,Raperc
Viena

Persona



Wonder Mike (1957-)

Hip hop
Raperos
Nueva York

Persona

Big Bank Hank (1958-)

Hip hop
Raperos
The Bronx

Persona

Ice-T (1958-)

Gangsta rap
Cantantes,Músicos,Raperos,Podcaster,Actores
Newark

Persona

Flavor Flav (1959-)

Hip hop
Raperos,Músico
Long Island

Persona

Chuck D (1960-)

Hip hop



Marl, Marley (1962-)

Hip hop



Shock-G (1963-)

Hip hop

Wikidata : concerns

PROVENANCE

Provenance is explicit in the *library side* of the data

024	7	http://viaf.org/viaf/93602067 2viaf
024	7	https://www.wikidata.org/wiki/Q309941 2wikidata
024	7	http://www.isni.org/isni/00000000115782584 2isni
040		SpMaBN bspa cSpMaBN erdc fembne
046		f1958
100	1	Ice-T d1958-
370		Newark vWikidata
372		Gangsta rap vWikidata
373		Body Count vWikidata
374		Actores aPodcaster aActores aCantantes aMúsicos aRaperos aActores vWikidata
375		Masculino vWikidata
377		eng Inglés
400	1	Marrow, Tracy d1958-

No mapping or device has been implement in the RDF side

Wikidata : concerns

2
4

QUALITY

100	1	Jovellanos, Gaspar Melchor de d1744-1811
372		Economía vWikidata
373		Real Academia de la Historia aReal Academia Española vWikidata
374		Actores aBarrenderos aYoutuber aVideojugador vWikidata
375		Masculino
377		lEspañol
370		Gijón, Asturias, España v(Wikidata) 0XX451245
400	1	Jovellanos d1744-1811
400	1	Jovellanos, Gaspar de d1744-1811
400	1	Jovellanos, Gaspar Melchor María d1744-1811
400	1	Jovellanos y Ramírez, Gaspar Melchor de d1744-1811
400	1	Suárez de Langreo, Toribio d1744-1811
670		Poesía, teatro, prosa literaria, 1993 bport. (Gaspar Melchor de Jovellanos)

Keep calm and Don't be afraid of errors

Thanks!!

Ricardo Santos
Head of Technical Services Department

ricardo.santos@bne.es



[@3186ricardo](https://twitter.com/3186ricardo)

Pº de Recoletos 20-22
28071 Madrid
España
T +34 915 807 800

www.bne.es

BIBLIOTECA NACIONAL DE ESPAÑA
NATIONAL LIBRARY OF SPAIN