

STORIE CHE NASCONO DAI DATI

COME CAMBIA IL GIORNALISMO NELL'ETÀ DELLA RETE

Alessio Cimorelli

dataninja.it | [@jenkin27](https://twitter.com/jenkin27)

<http://dtnj.it/nexa2014>

Torino, 10 settembre 2014



Nexa Center for Internet & Society
Politecnico di Torino

DALLA FISICA AL GIORNALISMO



IL NETWORK DATANINJA

PERSONE E COMPETENZE



Un incontro



Un interesse



L'idea



Un nome e un logo



Un server



Un blog



Un dominio



IL NETWORK DATANINJA

TECNOLOGIA E OPENNESS



Il premio



L'inchiesta



Il grant europeo



THE MIGRANTS' FILES

LE 23 MILA VITTIME DELLA FORTEZZA EUROPA

<http://themigrantsfiles.com>

<http://stories.dataninja.it/themigrantsfiles/>



#MIGRANTSFILES

<http://dtnj.it/tmf>

Un'inchiesta europea sui migranti morti dal 2000 nel tentativo di raggiungere la Fortezza Europa

CC-BY 3.0 Italia

dataninja.it

DATANINJA.IT / CHI SIAMO

- Dataninja è un network di data journalism fondato nel 2012
- Il nostro obiettivo è raccontare storie partendo dall'analisi dei dati, il **data-driven story-telling**
- Ci occupiamo di giornalismo e formazione con i principali gruppi editoriali italiani e con le pubbliche amministrazioni più avanzate in tema di Open Data.



IL CONTESTO: LE INFORMAZIONI SUL WEB

From the dawn of civilization until 2003, humankind generated five exabytes of data. Now we produce five exabytes every two days... and the pace is accelerating!



Eric Schmidt, Google's executive chairman | Fonte: <http://goo.gl/pgle3G>

PERCHÉ VISUALIZZARE I DATI?

Come non era mai successo nelle nostre vite, oggi abbiamo accesso a un'infinità di informazioni libere e gratuite. Con gli strumenti giusti possiamo cominciare a dare un senso a questi dati per vedere schemi e trend che altrimenti per noi sarebbero invisibili. Trasformando i numeri in forma grafica, permetteremo ai lettori di conoscere le storie che quei numeri nascondono.

Alberto Cairo, "The functional Art" | Fonte: <http://goo.gl/KHm3hT>

L'INCHIESTA SU L'ESPRESSO



The screenshot shows the top portion of a web article. At the top right, there is a 'NETWORK' dropdown menu. The L'Espresso logo is on the left. Social media icons for Facebook, Twitter, and Pinterest are on the right with the text 'SEGUICI SU'. Below the logo, a red sub-header reads 'IL PIÙ AMPIO MONITORAGGIO EUROPEO REALIZZATO DA UN TEAM INTERNAZIONALE DI 10 GIORNALISTI'. The main title 'Migranti, la guerra del Mediterraneo' is in large white font. The text below describes a report on migrant deaths in the Mediterranean, mentioning 23,000 deaths and a 50% increase over official estimates. It lists authors and a photo credit. A red link for an interactive map is provided. At the bottom, a white bar contains social sharing buttons for 'Consiglia', '2mila' (Facebook), 'Tweet' (Twitter), '299' (Twitter), 'Pinterest', '2' (Google+), '15' (Google+), 'Email', and 'Stampa'.

Publication day: in contemporanea su 9 giornali
in sei Paesi europei il 31 marzo 2014

LA NOTIZIA



**DAL 2000 AL 2013 PIÙ DI 23 MILA MIGRANTI MORTI
NEL TENTATIVO DI RAGGIUNGERE L'EUROPA**

LE ROTTE

entered in possession of valid travel documents and a visa whose validity period they have since overstayed.

Main migratory routes into the EU / land & sea

- ▶ Western Mediterranean route
- ▶ Central Mediterranean Route
- ▶ Eastern Mediterranean Route
- ▶ Eastern Land Borders Route
- ▶ Western Balkans Route
- ▶ Western African Route
- ▶ Circular Route from Greece and Albania
- ▶ Apulia and Calabria
- Schengen area
- Schengen associate countries

Canary Islands



OTTO ROTTE DI INGRESSO

LA MORTALITÀ DELLE ROTTE

È stato possibile stimare un tasso di mortalità delle singole rotte, calcolo che evidenzia che la criticità maggiore si verifica nel transito di migranti tra il Magreb e la Sicilia

1	Route	Type	2008	2009	2010	2011	2012
24	Apulia & Calabria	mortality			0.22%	1.60%	0.17%
25	Central Mediterranean route	mortality	2,43%	12.21%	1.69%	4.82%	4.09%
26	Circular route from Greece and Albania	mortality	0,00%	0.00%	0.00%	0.00%	0.00%
27	Eastern Land borders Route	mortality		1.20%	0.29%	0.00%	0.00%
28	Eastern Mediterranean route Land	mortality		0.01%	0.11%	0.03%	0.08%
29	Eastern Mediterranean route Sea	mortality		0.25%	0.49%	3.00%	3.50%
30	Western African route	mortality	2,37%	1.96%	0.00%	2.94%	1.15%
31	Western Balkan route	mortality	0,04%	0.03%	0.00%	0.00%	0.00%
32	Western Mediterranean route	mortality	3,88%	3.55%	1.86%	2.71%	3.13%

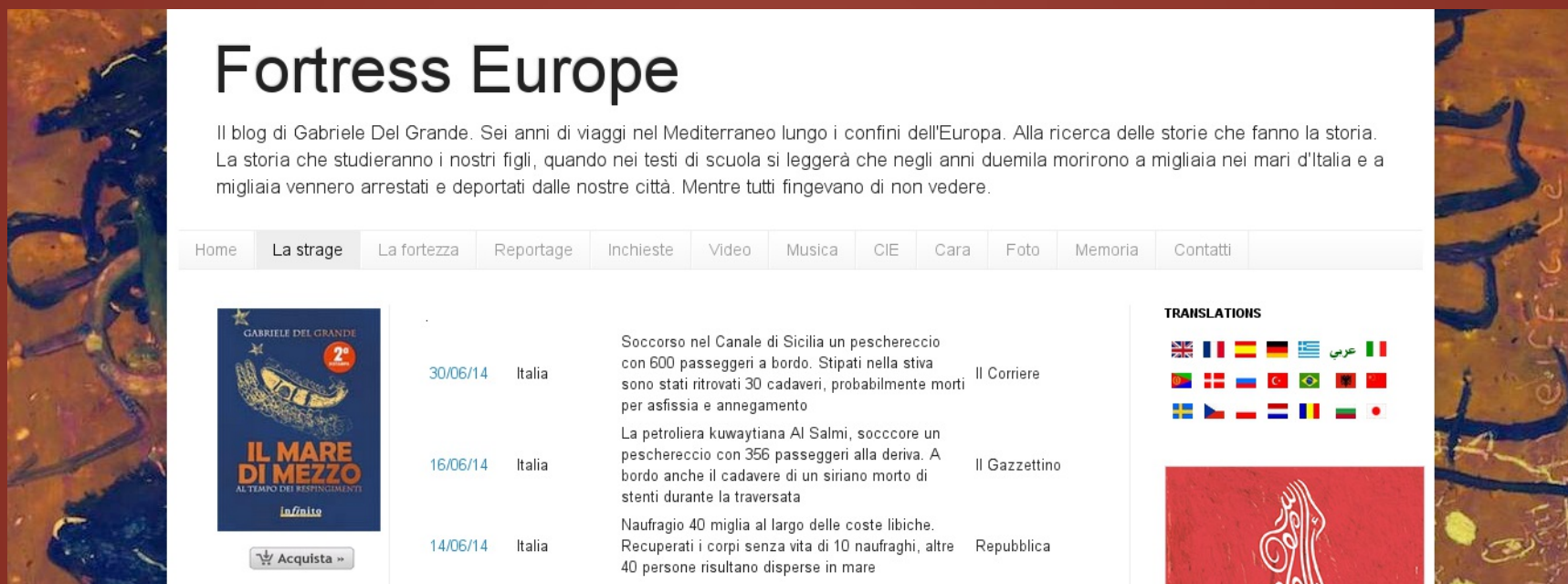
1. I DATI

LE BASI DELL'INCHIESTA

- **Non esistono** monitoraggi ufficiali delle vittime (né dello Stato italiano, né dell'Unione Europea)
- Due agenzie UE **monitorano** le frontiere (*Frontex, Eurosur*)
- **Esistono** raccolte dati basate su articoli e post pubblicati da mezzi di informazione e blog di NGO (*Fortress Europe, United for Intercultural Action, PULS dell'università di Helsinki*)
- **Esistono** studi su aree territoriali (*Sicilia, Canarie*) o su precise sezioni temporali (*un singolo anno, una stagione*)

FORTRESS EUROPE


Dataninja ha estratto i dati dal sito *Fortress Europe*.



Fortress Europe


Il blog di Gabriele Del Grande. Sei anni di viaggi nel Mediterraneo lungo i confini dell'Europa. Alla ricerca delle storie che fanno la storia. La storia che studieranno i nostri figli, quando nei testi di scuola si leggerà che negli anni duemila morirono a migliaia nei mari d'Italia e a migliaia vennero arrestati e deportati dalle nostre città. Mentre tutti fingevano di non vedere.

Home La strage La fortezza Reportage Inchieste Video Musica CIE Cara Foto Memoria Contatti

Home	La strage	La fortezza	Reportage	Inchieste	Video	Musica	CIE	Cara	Foto	Memoria	Contatti
											
30/06/14	Italia	Soccorso nel Canale di Sicilia un peschereccio con 600 passeggeri a bordo. Stipati nella stiva sono stati ritrovati 30 cadaveri, probabilmente morti per asfissia e annegamento		Il Corriere							
16/06/14	Italia	La petroliera kuwaytiana Al Salmi, soccorre un peschereccio con 356 passeggeri alla deriva. A bordo anche il cadavere di un siriano morto di stenti durante la traversata		Il Gazzettino							
14/06/14	Italia	Naufragio 40 miglia al largo delle coste libiche. Recuperati i corpi senza vita di 10 naufraghi, altre 40 persone risultano disperse in mare		Repubblica							


TRANSLATIONS

UK FR DE GR ES عربي IT
PT DK RU TR BR IN JP
FI PL CH NL BG RO



UNITED FOR INTERCULTURAL ACTION / PULS

Journalism++ ha estratto i dati di United for Intercultural Action e PULS

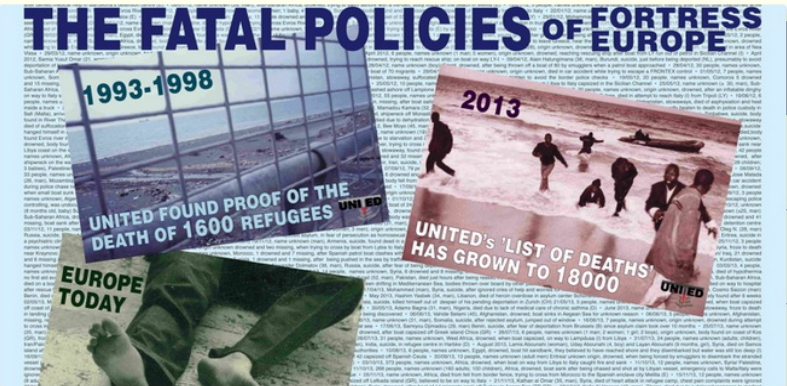


UNITED for Intercultural Action
European Network against nationalism, racism, fascism and in support of migrants and refugees

[Donate](#)
VISA MASTERCARD

[News](#) [About UNITED](#) [Campaigns](#) [Conferences](#) [Projects](#) [Support Us](#) [Calendar](#) [Publications](#) [Addressbook](#)

THE FATAL POLICIES OF EUROPE



1993-1998

UNITED FOUND PROOF OF THE DEATH OF 1600 REFUGEES

EUROPE TODAY

2013

UNITED'S 'LIST OF DEATHS' HAS GROWN TO 1800

Order Free Posters for the International Refugee Day Campaign 2014




Victory for some, defeat for more...



Vote and you can make a difference! – Caravanning in Hungary



'Your Vote Can Unite, You Choose' – Press Release, 21-05-2014

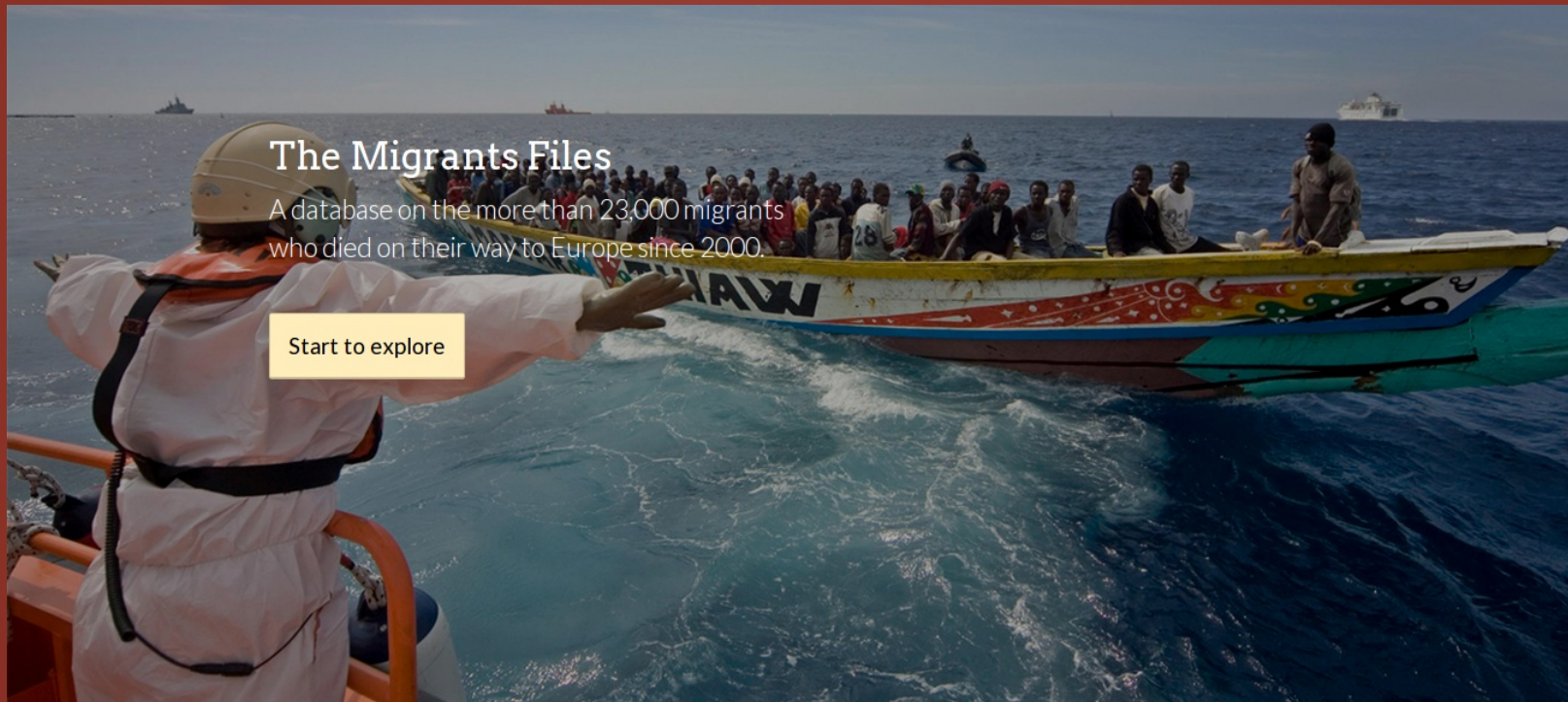


SCRAPING / MERGING / DEDUPLICATION

- È l'attività necessaria per estrarre i dati dai loro *contenitorie* strutturarli in tabella per renderli *machine readable*
- I dati sono stati estratti da: pagine html e pdf
- Sono stati riorganizzati replicando la struttura originale delle colonne (data / luogo / descrizione / fonte / ...)
- I dati sono stati aggregati in un'unica tabella e ripuliti attraverso un processo di deduplicazione per individuare ed eliminare i doppi.

IL NUOVO DATABASE

Contiene 2700+ eventi (nel corso dei quali sono morti 23000+ migranti dal 2000 al 2013) ed è *esplorabile*



<http://themigrantsfiles.com>

ENTITY EXTRACTION

I singoli eventi (es.: i naufragi) vengono messi in relazione per l'estrazione di entità (es.: tipi di eventi, località, tipi di vittime)

| Link: <http://goo.gl/xy0h0Q>

Detective.io Search for events, territories, detention centers and migrants [Login](#)

Boat capsizes off Tripoli on May 4, 2014
Event

Migrants shot by Libyan border guards in June 2013

Boat capsizes off Tripoli on May 4, 2014

Mal

Location	Off Libya's coast
Cause of death	Drowned
Latitude	33.89
Source	Guardian
Migration route	Central Mediterranean Route
Dead	36
Source URL	http://www.theguardian.com/world/2014/may/11/migrant-boat-sinks-coast-libya

[JSON](#) [XML](#)

Jump to...

- [Happened in](#)
- [Crossing from](#)
- [Crossing to](#)

2. IL TEAM

GIOCO DI SQUADRA

L'inchiesta è stata sviluppata da un team di oltre 10 giornalisti, analisti, sviluppatori sparsi per tutta Europa, tramite una piattaforma di lavoro condivisa.

The screenshot shows a Trello board interface for a project named "Fortress Europe". The board is organized into four main columns: Literature, Data Cleaning, Articles, and Organization. Each column contains several cards representing different tasks or documents. The "Data Cleaning" column features a card with a table of data. The "Articles" column includes a card with a map. The "Organization" column has a card with a bubble chart. A right-hand sidebar shows the board's menu, including members and activity.

Boards **Trello**

Fortress Europe ★

Literature

- The Charter of Lampedusa
- Borders, Mobility and Technologies of Control
- Xénophobie business : A quoi servent les contrôles migratoires
- Bilal: Als Illegaler auf dem Weg nach Europa
- Del Grande's 2007 report
- International Centre on Migration Policy Development

Data Cleaning

- Error estimate for the deduplication process
- 2013 detections Data
- New Italian Observatory on Migrants Deaths

Articles

- Follow up
- Spanish Article
- Publication Plan Jan 21st
- Maps
- Documented migrants death 50% higher than previously believed

Organization

- Possible external contributions
- Re-using our data
- Links of/about TMF
- Official Hashtag

Menu

Members

Activity

Michail Batikas on Official Hashtag

thank you @nelsonmau ! I have to add that geocoding and language are not exactly the same, especially because of the expats (like myself :) I have them all (#MigrantsFiles) since Sunday afternoon, when some of us did some kind of promo.

yesterday at 15:06

Andrea Nelson Mauro

RUOLI E METODOLOGIA

- Un **project manager** si è occupato di coordinare il processo e garantire lo stato di avanzamento
- I **giornalisti** hanno cercato dati ed effettuato verifiche a campione sui singoli eventi (*fact-checking*)
- **Dataninja** ha effettuato verifiche per stimare la possibilità di errore all'interno dei dati (stimato in massimo il **10%**)

METODOLOGIA SU L'ESPRESSO

Nell'inchiesta pubblicata su L'Espresso è stata pubblicata anche la metodologia di lavoro | Link: <http://goo.gl/9gpvIA>

Come nasce l'inchiesta

L'inchiesta de l'Espresso è parte dell'inchiesta europea "The Migrants Files". Il lavoro investigativo è stato condotto da 10 giornalisti, tra i quali i network di datajournalism [Dataninja](#) e [Journalism++](#). L'inchiesta e il database sono co-pubblicati da vari media europei, tra cui El Confidencial, Neue Zürcher Zeitung, Sydsvenskan e Le Monde Diplomatique. Per l'Italia il lavoro è stato curato dai giornalisti Andrea Nelson Mauro, Alessio Cimarelli e Jacopo Ottaviani e il sito dell'inchiesta è raggiungibile all'indirizzo <http://stories.dataninja.it/themigrantsfiles>. Al progetto hanno contribuito anche gli studenti del Laboratorio di Data Journalism dell'Università di Bologna sotto la supervisione di Carlo Gubitosa (giornalista e docente del laboratorio), verificando un campione degli eventi censiti nel database. La realizzazione del progetto è stata possibile grazie al supporto dello [European Cross-border Grant Programme](#) di [JournalismFund.eu](#).

Nota metodologica

I tre database originali (Fortress Europe, United e Puls) sono stati ripuliti e uniformati prima di essere uniti in quello finale. Analizzando date, numero di morti e dispersi, luoghi menzionati e descrizione degli eventi sono stati identificati gli eventi duplicati (presenti in più di un database) ed eliminati. Vari controlli incrociati a campione hanno permesso di integrare dati mancanti, aggiungere dettagli degli eventi, recuperare fonti non più esistenti o raggiungibili. Dai nomi dei luoghi si sono ricavate le coordinate geografiche con cui si è costruita la mappa interattiva. Laddove le coordinate precise dei naufragi non erano disponibili, gli eventi sono stati mappati e raggruppati sulle località le più vicine possibile (per esempio, in un ipotetico naufragio avvenuto nel Mar Libico, le vittime sono state geolocalizzate sul baricentro della Libia, dove erano avvenute anche altre tragedie).

FACT-CHECKING ALL'UNIVERSITÀ

Gli studenti di data journalism dell'Università di Bologna hanno contribuito a verificare i singoli eventi (grazie a Carlo Gubitosa)



3. I RISULTATI

31 MARZO 2014: IL PUBLICATION DAY

L'inchiesta è stata pubblicata in contemporanea su 9 giornali di 6 diversi paesi europei il 31 marzo 2014, generando una diffusione capillare in ogni Paese

About the project

The Migrants' Files project was launched in August 2013 by a group of European journalists who joined forces to accurately calculate and report the deaths of emigrants seeking refuge in Europe. This pan-European consortium of journalists is partially funded by the European non-profit organization Journalismfund.eu.

They know their lives are at risk, yet each year thousands of people from Africa, the Middle East and beyond – war refugees, asylum seekers and economic migrants – leave their homes and try to reach the promised land of Europe. On the third of October 2013, more than 360 would-be emigrants drowned off the coast of the Italian island of Lampedusa. A catastrophe of this dimension grabbed the media's attention for a while and won the sympathy of the general public.

In response, later that month, the European Council decided to implement measures aimed at preventing a repeat of such a tragedy at the European Union's borders. The Council called for strengthening the EU's border security co-ordination system, Frontex, more formally known as the European Agency for the Management of Operational Cooperation at the External Borders of the Member States of the European Union. And the Europe-wide surveillance system Eurosur began operations on December 2, 2013. Thus, once again, a large tragedy spurred real, if belated, action.

Well-intended though they doubtless were, these measures only address the tip of the migration iceberg. Little is known about how many men, women and children actually have lost their lives on their journey to Europe. Believing that policy unsupported by facts cannot be optimal, a consortium of European journalists committed themselves to systematically assembling and analyzing the data on the deaths of Europe's would-be migrants. The Migrants' Files project is partially funded by the European non-profit organization Journalismfund.eu.

Neue Zürcher Zeitung

El Confidencial

radiobubble.gr
intergalactic radio

SYDSVENSKAN

L'Espresso

HBL

Le Monde
diplomatique

euobserver.com

MO
* BONDIAL NEWS

Neue Zürcher Zeitung

[Die Toten vor Europas Türen](#)

Neue Zürcher Zeitung

[Daten einer Tragödie](#)

Radio Bubble

[Τα θανατηφόρα τείχη της Ευρώπης](#)

El Confidencial

[Trece años de inmigración: más de 23.000 muertos por intentar alcanzar Europa](#)

El Confidencial

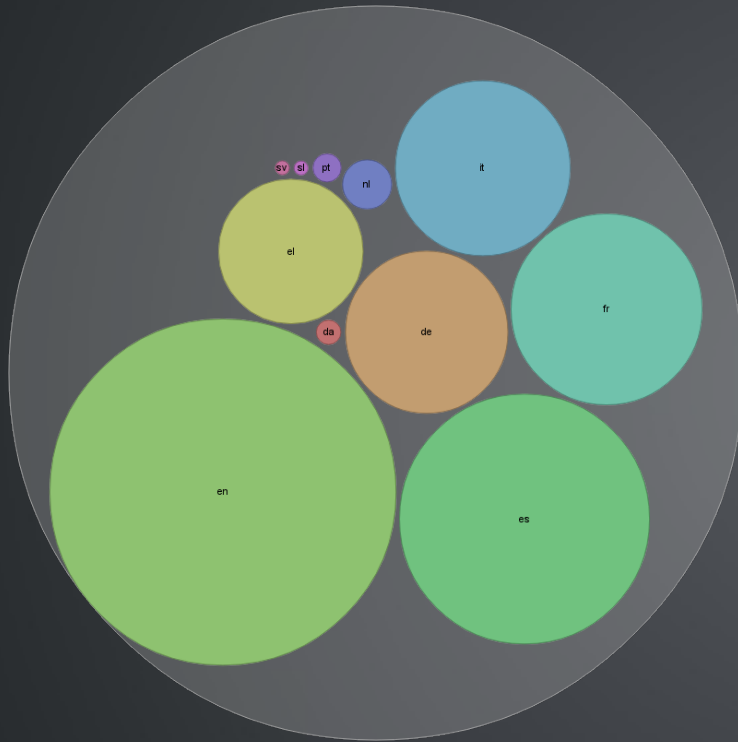
[Ante el cierre de fronteras, los inmigrantes se vuelcan hacia rutas más peligrosas](#)

Hufvudstadsbladet

[Fästning Europa kräver tusentals liv \(€\)](#)

L'IMPATTO

In 24 ore sono stati prodotti 2000+ tweet in 11 lingue diverse



Prevalentemente: inglese, francese, italiano, tedesco, spagnolo e greco | <http://goo.gl/cTZyqq>

I PREMI: BRUSSELS E BARCELONA

L'inchiesta è stata presentata durante la DataHarvest+ Conference a Brussels e ha vinto i Data Journalism Awards, assegnati a Barcelona (giugno 2014)

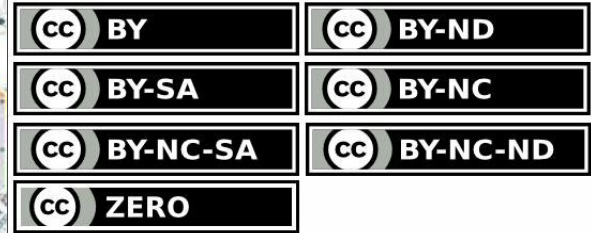


IL VIDEO

#migrantsfiles: <http://vimeo.com/97210043>

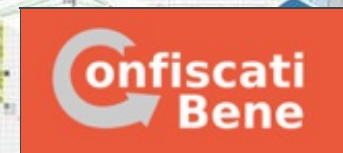
OPEN DATA & OPEN SOURCE

TRASPARENZA E CONDIVISIONE



Gli open data, istituzionali e non, sono preziosi per il data journalism, che a sua volta non può disattendere i principi di trasparenza e falsificabilità.

I giornalisti come produttori di open data, oltre che utilizzatori, in contatto con diverse comunità.



CAWABONGA!

Alessio Cimorelli / jenkins

jenkin@dataninja.it

@jenkin27

Dataninja

www.dataninja.it

<http://school.dataninja.it>

info@dataninja.it

Newsletter

<http://dataninja.it/newsletter>

