

Web survey e strumenti digitali per le community

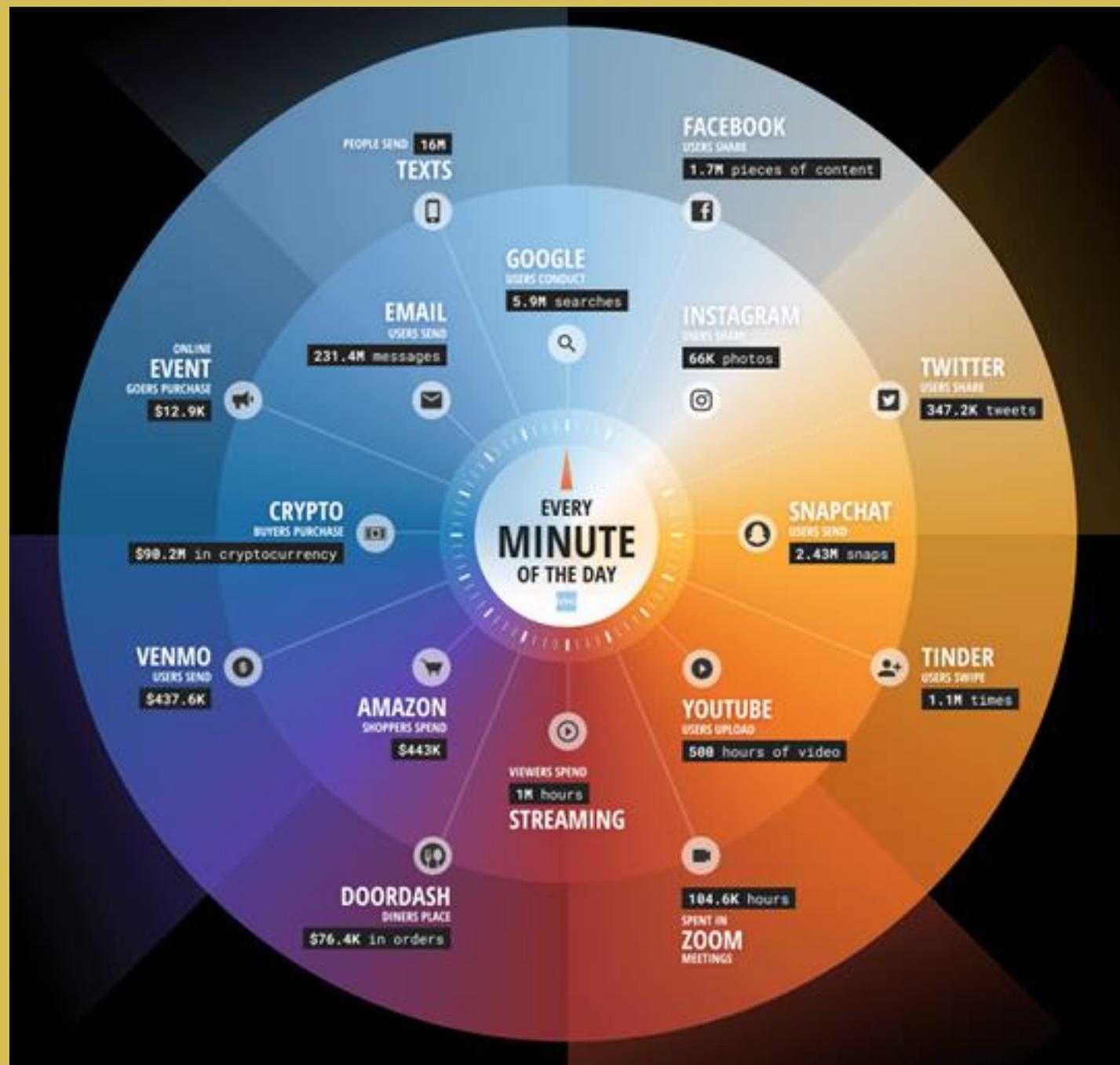


UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

Prof. Biagio Aragona
Università degli studi di Napoli Federico II
mail: aragona@unina.it

I DATI NON DORMONO MAI

Espressioni come “diluvio di dati” (The Economist, 2010), “rivoluzione dei dati” (Kitchin, 2014a), “datificazione” (Espeland & Stevens, 2008) e “società data intensive” (Amaturo & Aragona, 2017) definiscono questa diffusione di dati.



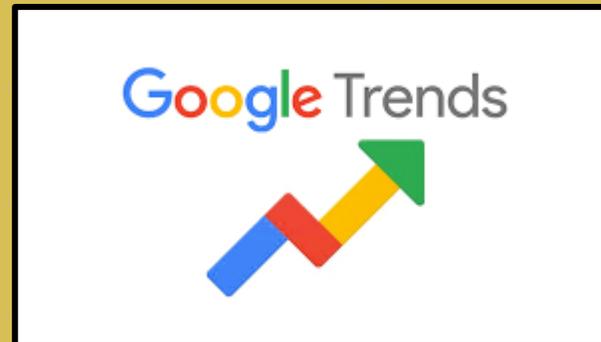
Fonte: Domo, 2023

L'ERA DIGITALE E LA RICERCA SOCIALE

- Nell'era analogica, la maggior parte dei dati utilizzati per la ricerca sociale era creata con l'obiettivo primario di fare ricerca.
- Nell'era digitale invece, una grande quantità di dati, i cosiddetti **big data**, viene creata per fini diversi dalla ricerca: pensiamo ai dati generati dalle imprese con lo scopo di fornire e migliorare i servizi ai clienti, o quelli caricati dagli utenti sui social network o quelli della pubblica amministrazione per l'erogazione e la gestione dei servizi pubblici.
- Questa mole di dati digitali si presta ad un reindirizzamento a fini di ricerca.

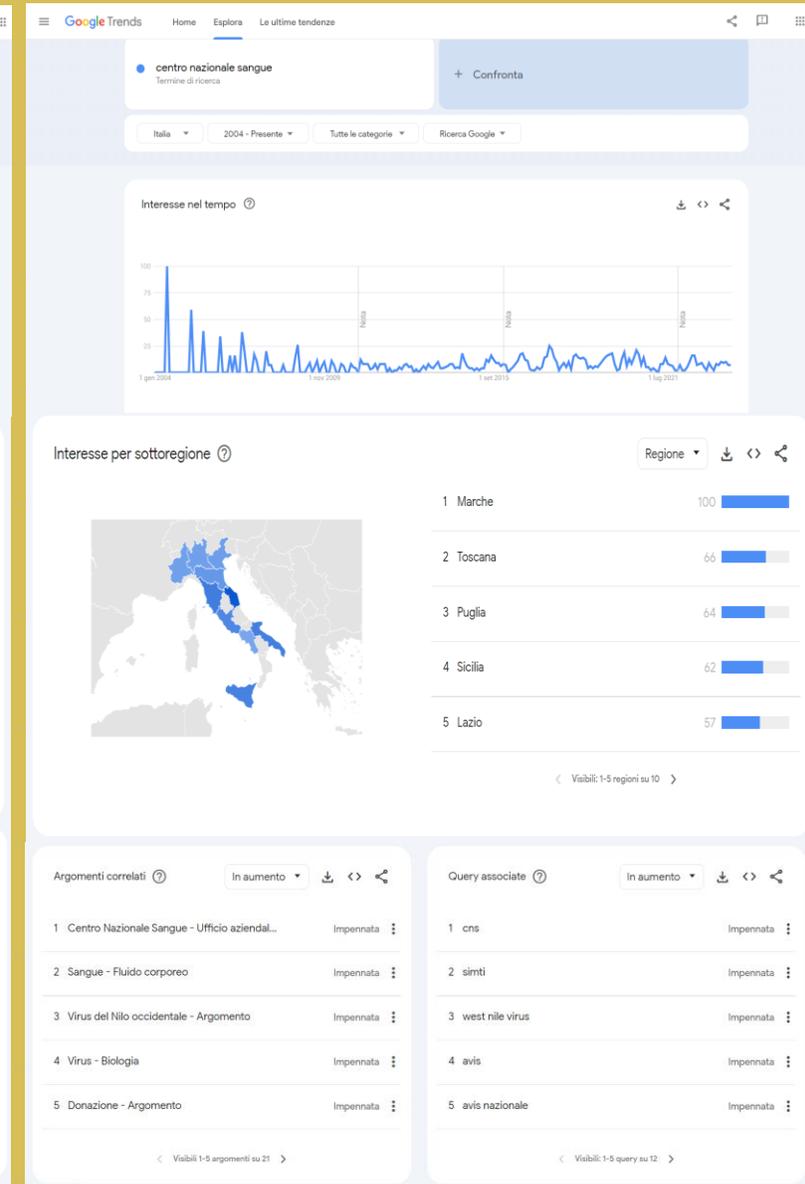
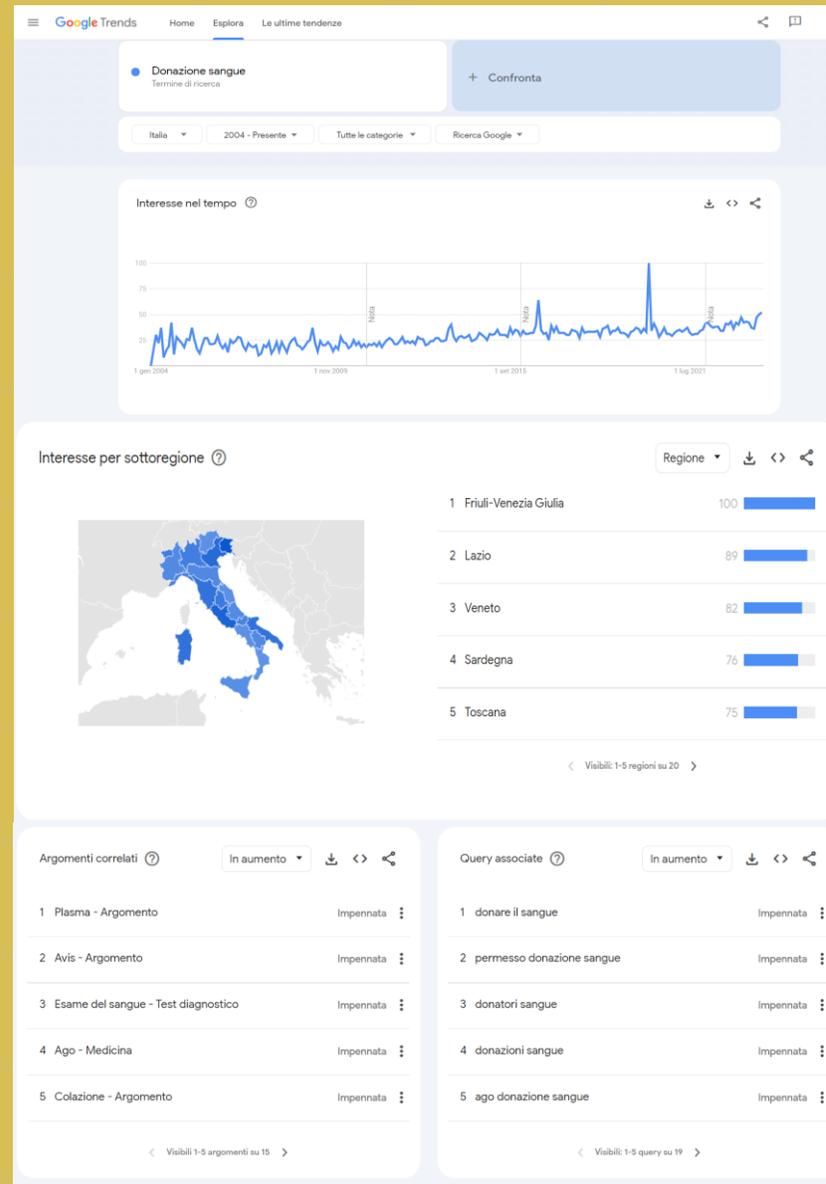
SEARCH AS RESEARCH

- Per Search As Research si intende l'uso dei motori di ricerca per obiettivi di ricerca
- Le ricerche fatte dagli utenti sul motore di ricerca diventano dati che ci possono segnalare l'interesse verso un particolare fenomeno
- Alcuni motori di ricerca hanno al loro interno delle piattaforme per l'analisi delle ricerche degli utenti. Ad esempio,



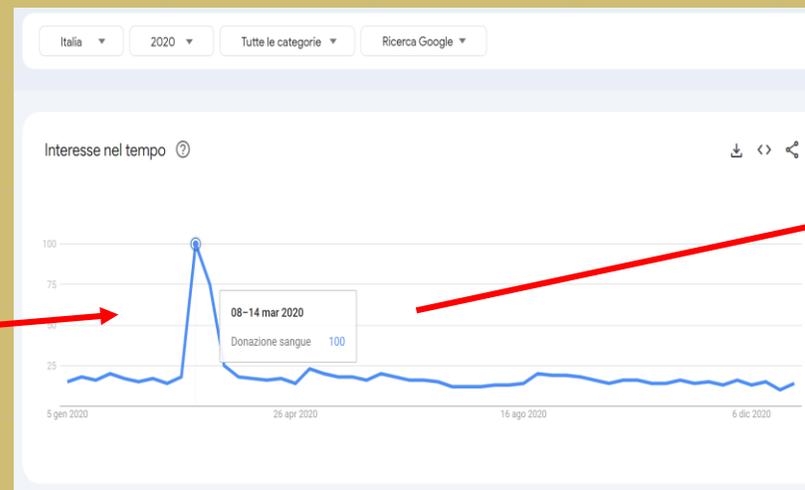
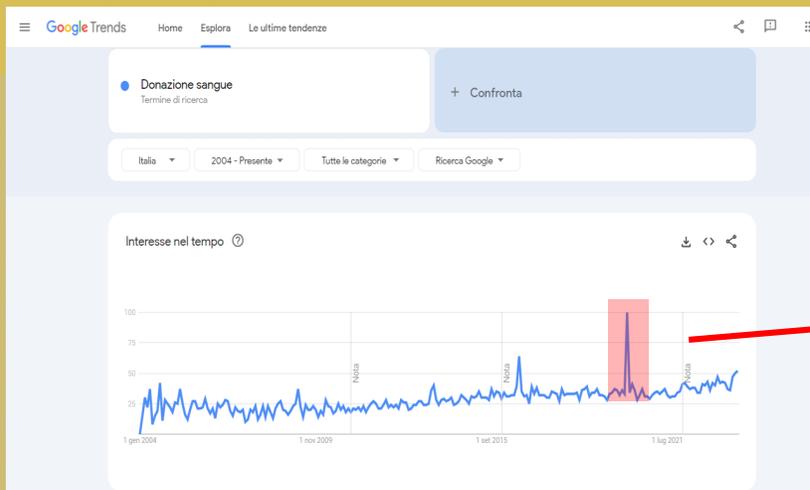
APPLICAZIONE DELLA SEARCH AS RESEARCH

- Risultati di ricerca su Google Trends per “donazione sangue” e “centro nazionale sangue” in Italia dal 2004 ad oggi.
- Analisi longitudinale
- Analisi spaziale



I PICCHI DEL GRAFICO – DONAZIONE SANGUE E CORONA VIRUS

Per comprendere un picco si isola il picco in modo da rinvenire il range temporale in cui si è sviluppato il maggior numero di ricerche sull'intera distribuzione. Si può poi fare ricorso a ulteriori fonti informative per rintracciare quali siano le possibili ragioni del picco. In questo caso ho fatto ricorso all'archivio storico di La Repubblica.



Repubblica.it | 13 marzo 2020 | sez. Medicina e Ricerca

L'allarme Emergenza sangue Scorte quasi finite Mancano i donatori

... relazione logistica con i reparti di cura e il percorso di **donazione** è privo di qualsiasi rischio. Il **sangue** è indispensabile per affrontare le emergenze del pronto soccorso, della chirurgia d'urgenza, delle...

11 marzo 2020 | pag. 07 | sez. CRONACA | CONTENUTO A PAGAMENTO

Coronavirus, calano le donazioni di sangue. L'appello: "Nessun rischio contagio: i centri di raccolta sono sicuri"

...Calano le donazioni di **sangue** in tutta Italia a causa del coronavirus. In questi giorni si stima una riduzione del 10-15% e si teme che le cose vadano a peggiorare. Molti hanno paura di recarsi nelle...

di MICHELE BOCCI

Repubblica.it | 12 marzo 2020 | sez. Cronaca

"Per favore continuate a donare il sangue"

... **donazione di sangue**...

14 marzo 2020 | pag. 02 | sez. CRONACA | CONTENUTO A PAGAMENTO

L'allarme di Avis e Fidas Appello ai donatori il sangue: "Venite, siamo più puliti del supermarket"

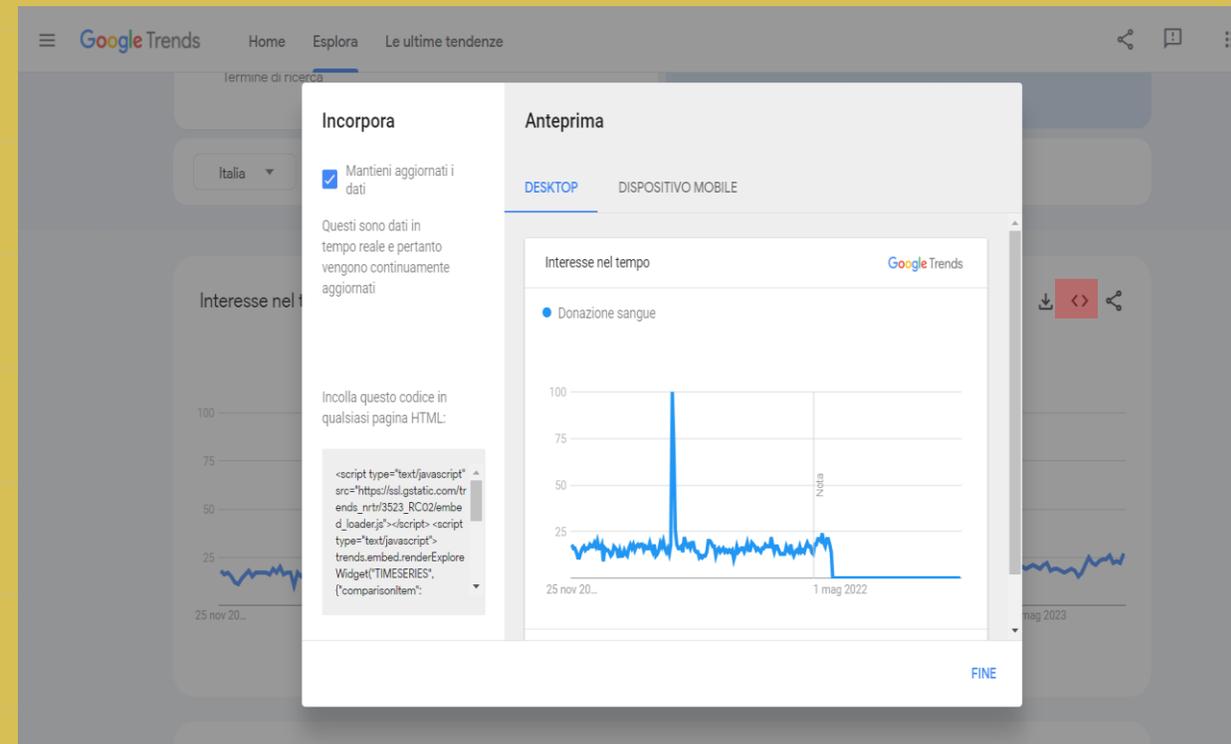
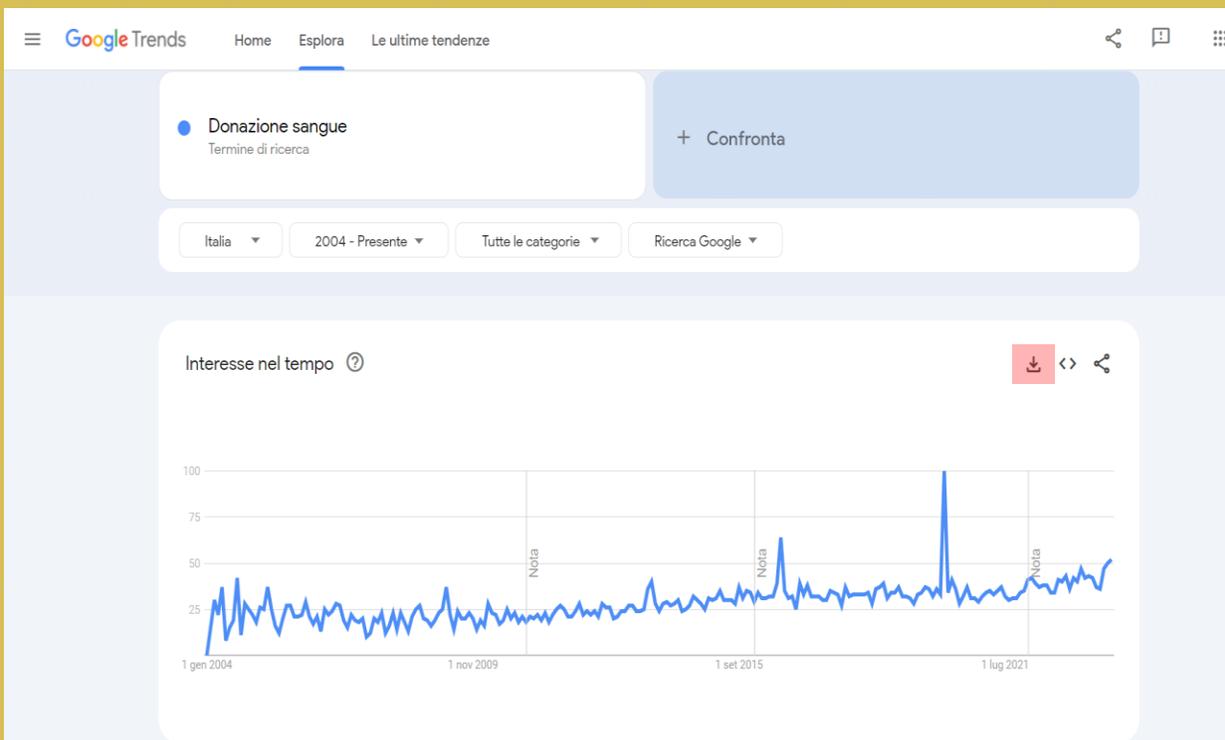
...Calano le donazioni di **sangue**, e le associazioni come Avis e Fidas lanciano l'allarme: "Venite nei nostri centri trasfusionali, sono più puliti del supermercato che frequentate ogni giorno". In piena...

13 marzo 2020 | pag. 07 | sez. CRONACA | CONTENUTO A PAGAMENTO

SEARCH AS RESEARCH – ALCUNE FUNZIONI

Questo strumento ci permette di scaricare i dati sotto forma di matrice dati.

Questo strumento ci permette anche di incorporare all'interno di una pagina web i grafici presentati, sia in forma statica, che in forma dinamica. Basterà cliccare sull'icona "<>" e importare il codice all'interno del codice della propria pagina web.



UN'APPLICAZIONE PARTICOLARE DELLA SEARCH AS RESEARCH - IL NOWCASTING

- Il nowcasting è un termine derivato dall'unione di *now* (ora) e *forecasting* (previsione).
- Il *nowcasting* è la previsione del presente (Choi e Varian, 2012). È un modo di prevedere tempestivo.
- È molto utile per monitorare l'interesse per una particolare campagna di comunicazione o per capire l'interesse per un particolare avvenimento collegato alla donazione.

IL WEB SCRAPING

- La fonte di dati digitali più grande è rappresentata dalle pagine web di internet e dai social network.
- Il web scraping ha l'obiettivo di estrarre dati, elaborare testi, e costruire un *corpus* a partire dalle pagine web di Internet e dai social.
- Il web scraping è una tecnica di raccolta dei dati non intrusiva .
- Un aspetto fondamentale della ricerca con i dati dei social media è la comprensione dei limiti dei dati a cui si accede, perché le piattaforme limitano l'accesso ai propri dati. Rispetto ai social media, bisogna affrontare tre fonti generali di *bias* nella raccolta dei dati:
 1. *Bias* di accesso;
 2. *Bias* nelle caratteristiche della popolazione dei social media;
 3. *Bias* di campionamento e bot.

WEB DATA RESEARCH ASSISTANT

- Plugin compatibile con browser Chrome e Edge, ci permette di ricavare dati da piattaforme digitali. Questo strumento è altamente performante per le piattaforme X (Twitter) e Instagram. Per Facebook invece permette l'estrazione dei testi dei post, i link delle immagini e video contenuti nei post.
- È stato sviluppato dal Web Science Institute dell'University of Southampton



APPLICAZIONE DI WEB SCRAPING

Pagina  del Centro Nazionale Sangue

ID	Impressions	Replies	Retweets	Likes	Author	Retweeted By	In Reply To	Mentions	Hashtags	Embedded Image	Embedded Video	Embedded URL	Embedded Text	Text	URL	Timestamp	Sanitised Text
ID:1724063255256072633	196000	63	415	2600	fedez	centrosangue						https://pbs.twimg.com/media/F-0aYt8W8AAmfGH?format=jpg&name=small	Senza i c	https://t	13/11/2023 13:54	Senza i donatori di sa	
ID:1722582992156074073	151	1	2	6	centrosangue			fedez	#sangue.			https://pbs.twimg.com/media/F-fSvt7W8AAZnVv?format=jpg&name=small	Circa un	https://t	09/11/2023 11:52	Circa un mese fa ha ar	
ID:1721911107034448112	315	1	2	5	centrosangue			luccacandg	panicomicisit av			https://pbs.twimg.com/media/F-V0_XZXEAAbSYx?format=jpg&name=small	Al @Luo	https://t	07/11/2023 15:22	Al ha esordito "Be a h	
ID:1721911111497224439	196	1	2	4	centrosangue				#fumetto,			https://pbs.twimg.com/media/F-V1BIDWUAApDp?format=jpg&name=small	Sono sta	https://t	07/11/2023 15:22	Sono state 30mila le c	
ID:1721911115766984918	211		1	3	centrosangue			marco_rizzo gi	#donazione di			https://pbs.twimg.com/media/F-V007GWsAAbPuP?format=jpg&name=small	Un grazia	https://t	07/11/2023 15:22	Un grazie non solo a F	
ID:1721074872322036198	1100		9	18	crocerossa	centrosangue		centrosangue	panicomicisit			https://pbs.twimg.com/media/F-GP750WYAEeaH6?format=jpg&name=small	"Red Sw	https://t	05/11/2023 08:00	"Red Sword" Ã il nuc	
ID:1719292870044114948	88		1	1	aousenese	centrosangue		centrosangue	#donazione	#Halloween	#CentroEmotrasfusionale			https://t	31/10/2023 09:58	ÃœœUna mostruosaÃœ	
ID:1718934417844928714	1200		7	10	salutelazio	centrosangue		regionelazio	#Halloween			https://pbs.twimg.com/media/F9rhvSCXcAAACrn?format=jpg&name=small	Non ferr	https://t	30/10/2023 10:14	Non fermarti a ãœced	
ID:1714951893993046404	177		2	5	fidasnazionale	centrosangue		ministerosalut	#Sangue	#tras		https://pbs.twimg.com/media/F8y6U5rX0AASQCH?format=jpg&name=small	#Sangue	https://t	19/10/2023 10:29	Ã eÃ Ã sono delle ver	
ID:1714618280873193940	1100		10	16	centrosangue			luccacandg ma	#BeAHero	#dc		https://pbs.twimg.com/media/F8uK_oYXQAA1h_h?format=jpg&name=small	#BeAHei	https://t	18/10/2023 12:23	presenta: Red Sword	
ID:1712741737846714725	3700	1	4	5	edqm_news	centrosangue						https://pbs.twimg.com/ext_tw_video_thumb/17127417118840	A Pan-Ei	https://t	13/10/2023 08:07	A Pan-European Effor	
ID:1712743623664472479	125			2	centrosangue				#comlas	#pati		https://pbs.twimg.com/media/F8TeP8-X0AAoF73?format=jpg&name=small	Anche il	https://t	13/10/2023 08:14	Anche il nostro centro	
ID:1711698086588801110	496	1	8	11	issalute_it	centrosangue			#plasmaferesi			https://pbs.twimg.com/media/F7vFRZCWIAA72ed?format=jpg&name=240x240	Sai che ð	https://t	10/10/2023 11:00	Sai che Ã possibile d	
ID:1711672675372798002	2600		3	8	edqm_news	centrosangue			#EuropeanDonationDay			https://pbs.twimg.com/tweet_video_thumb/F8EVNwYwCAAK	What ex	https://t	10/10/2023 09:19	What exactly is organ	
ID:1711421814633623886	154		2	2	fidasnazionale	centrosangue		ministerosalut	#fidas	#fidatic		https://pbs.twimg.com/card_img/172433337:https://t.co/XCf8Ow0EyP	Musso, F	https://t	09/10/2023 16:42	Musso, Presidente FI	
ID:1710967888964714608	50000	2	171	1600	robertoburioni	centrosangue		fedez				https://pbs.twimg.com/media/F76UDYsXgAA1_8l?format=png&name=small	Sportivi,	https://t	08/10/2023 10:38	Sportivi, attori e cant	
ID:1711298124185030753	63000	87	41	477	corriere	centrosangue						https://pbs.twimg.com/card_img/172419147:https://t.co/kPjCONsyOd	Effetto F	https://t	09/10/2023 08:30	Effetto Fedez, boom	
ID:1709886254433604093	319		6	11	ipopi_info	centrosangue			#IPAW2023	#YouAreMyMedicine		https://pbs.twimg.com/ext_tw_video_thumb/1709884989372	Did you	https://t	05/10/2023 11:00	Did you know? There	
ID:1709110403316187271	111		1	2	pptaeurope	centrosangue						https://pbs.twimg.com/media/F7f6uLiWoAAmo9X?format=jpg&name=small	Some 55	https://t	03/10/2023 07:37	Some 55% of blood is	
ID:1709131144178835948	875		10	20	eubloodalliance	centrosangue						https://pbs.twimg.com/media/F7gNS3YXIAAyC2X?format=jpg&name=small	Every ye	https://t	03/10/2023 08:59	Every year, blood est	
ID:1708814635275616357	336		4	13	ipopi_info	centrosangue			#IPAW2023	#YouAreMyMedicine.		https://pbs.twimg.com/ext_tw_video_thumb/17088129133712	In previc	https://t	02/10/2023 12:02	In previous years, ma	
ID:1708797804703871203	501		6	18	leiresolis	centrosangue		ipopi_info	#YouAreMyMedicine	#immunoglob		https://pbs.twimg.com/ext_tw_video_thumb/17087524997696	#YouAre	https://t	02/10/2023 10:55	- a fantastic way of re	
ID:1708880209146007994	1000	1	5	10	edqm_news	centrosangue						https://pbs.twimg.com/media/F7cpfGmX0ASgdy9?format=jpg&name=small	Today, v	https://t	02/10/2023 16:22	Today, we would like	
ID:1709113565460754882	133		2	6	centrosangue				#Sistra	#trasfu		https://pbs.twimg.com/media/F7f9uxXW0AA3Vtr?format=jpg&name=small	Sta per i	https://t	03/10/2023 07:50	Sta per iniziare la dec	
ID:1708383132037022193	3300	4	19	34	istsupsan	centrosangue			#ISS,			https://pbs.twimg.com/media/F7VITKpXQAAA21e?format=jpg&name=small	Dai dati	https://t	01/10/2023 07:27	Dai dati della sorvegli	
ID:1707315187206795464	216		2	5	centrosangue				#notte deirice			https://pbs.twimg.com/media/F7GVQJ-WkAAAxpnp?format=jpg&name=small	Anche n	https://t	28/09/2023 08:43	Anche noi saremo all	
ID:1707032918906778070	201	1	3	2	edqm_news	centrosangue						https://pbs.twimg.com/media/F7CZS9GXkAAatizf?format=jpg&name=small	Navigati	https://t	27/09/2023 14:02	Navigating the journe	

Impressions: visualizzazioni.
Replies: risposte, commenti.
Mentions: quale account viene menzionato nel tweet.
Sanitised Text: il testo del tweet ripulito.

Fedez: 2600 likes, 196000 visualizzazioni!!!
Burioni 1600 likes, 50000 visualizzazioni

DIGITALE E SURVEY – LA COSTRUZIONE DEI CAMPIONI

- Costruire campioni rappresentativi è sempre più difficile, perché aumentano le percentuali di non risposte.
- Costruire campioni non probabilistici è diventato più veloce, più economico e qualitativamente migliore.
- I contesti digitali offrono l'opportunità di estrarre campioni ampi.

DIGITALE E SURVEY – SOMMINISTRAZIONE E FORMULAZIONE DELLE DOMANDE

- I vantaggi di somministrare i questionari sul web sono tanti:
 - si possono raggiungere individui lontani.
 - si possono prevedere somministrazioni multicanale (social, e-mail, etc.).
 - la costruzione di dati è interamente gestita in modo digitale, riducendo i margini di errore e rendendo subito la matrice disponibile per l'analisi.
 - si riduce la desiderabilità sociale.
- Inoltre, i contesti digitali consentono speciali forme di formulazione delle domande (es. wikisurvey, gamification)

CONCLUSIONI

- Il digitale sta rendendo possibile nuovi modi di configurare l'apparato empirico della ricerca sociologica, con importanti conseguenze sul metodo e sulla metodologia.
- Infrastrutture digitali e tecniche di scraping aprono allo sviluppo di disegni di ricerca in tempo reale e a nuove forme d'uso di dati rilevati attraverso tecniche non intrusive.
- Le tecniche digitali non sostituiranno quelle tradizionali, ma c'è bisogno di un mutuo adattamento e disadattamento tra tecniche di ricerca dentro e fuori i contesti digitali.
- Bisogna prestare attenzione a come configurare le tecniche digitali in modo fruttuoso per la ricerca sociale.
- Per questo va sperimentato l'uso di nuove tecniche.