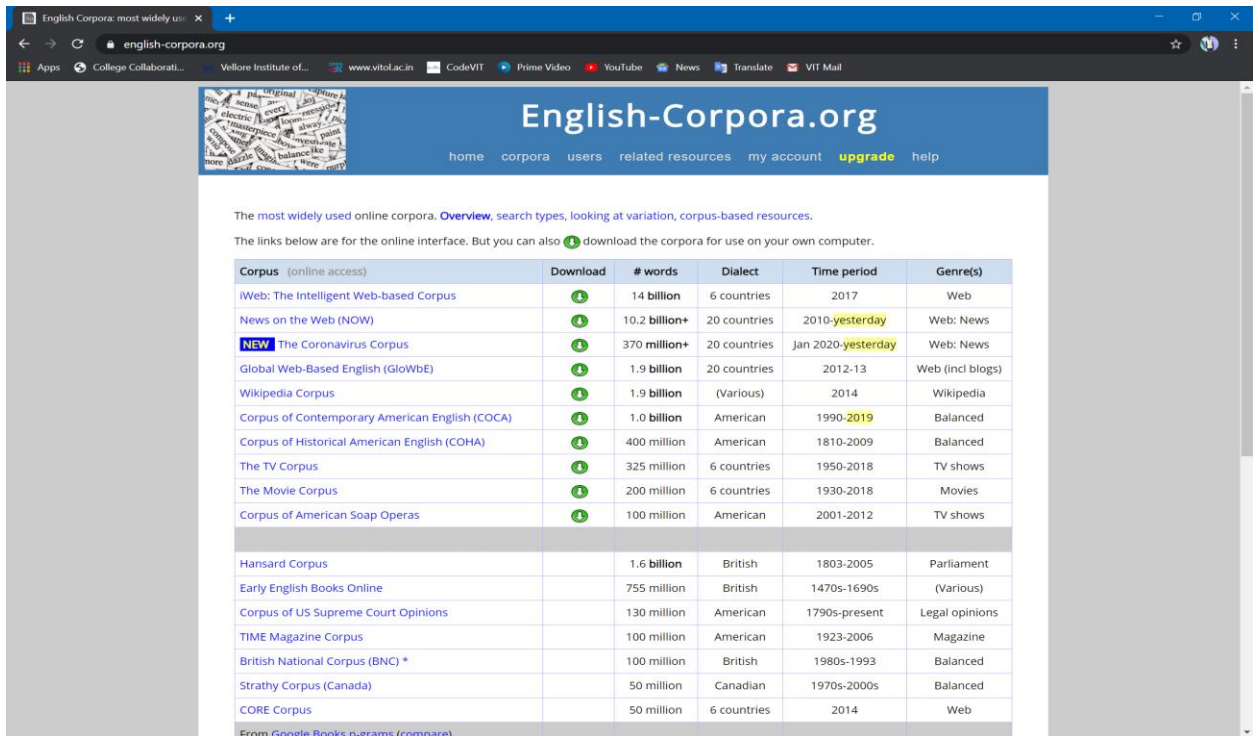


1. Exploring the corpus

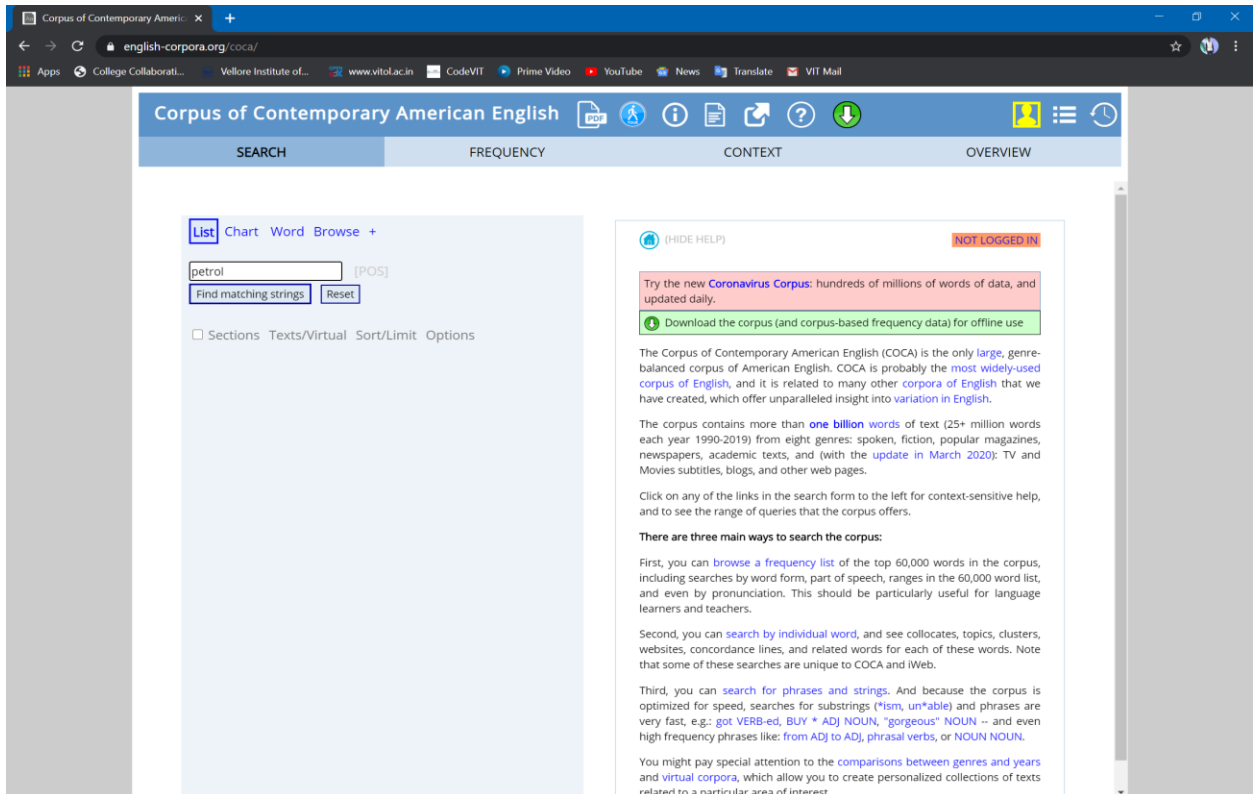


The screenshot shows the English-Corpora.org website. The header includes the site name and navigation links: home, corpora, users, related resources, my account, upgrade, and help. Below the header, a text block states: "The most widely used online corpora. Overview, search types, looking at variation, corpus-based resources. The links below are for the online interface. But you can also download the corpora for use on your own computer." Below this is a table listing various corpora.

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		10.2 billion+	20 countries	2010- yesterday	Web: News
NEW The Coronavirus Corpus		370 million+	20 countries	Jan 2020- yesterday	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		1.0 billion	American	1990- 2019	Balanced
Corpus of Historical American English (COHA)		400 million	American	1810-2009	Balanced
The TV Corpus		325 million	6 countries	1950-2018	TV shows
The Movie Corpus		200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas		100 million	American	2001-2012	TV shows
Hansard Corpus		1.6 billion	British	1803-2005	Parliament
Early English Books Online		755 million	British	1470s-1690s	(Various)
Corpus of US Supreme Court Opinions		130 million	American	1790s-present	Legal opinions
TIME Magazine Corpus		100 million	American	1923-2006	Magazine
British National Corpus (BNC) *		100 million	British	1980s-1993	Balanced
Strathy Corpus (Canada)		50 million	Canadian	1970s-2000s	Balanced
CORE Corpus		50 million	6 countries	2014	Web

From: Google Books n-grams (compare)

2. Exploring the coca corpus



The screenshot shows the Corpus of Contemporary American English (COCA) interface. The header includes the site name and navigation links: SEARCH, FREQUENCY, CONTEXT, and OVERVIEW. Below the header, there is a search form with a text input field containing "petrol" and a "[POS]" dropdown menu. Below the search form, there are checkboxes for "Sections", "Texts/Virtual", "Sort/Limit", and "Options". To the right of the search form, there is a sidebar with a "NOT LOGGED IN" message and a "Download the corpus (and corpus-based frequency data) for offline use" button. Below this, there is a text block describing the COCA corpus.

Try the new **Coronavirus Corpus**: hundreds of millions of words of data, and updated daily.

Download the corpus (and corpus-based frequency data) for offline use

The Corpus of Contemporary American English (COCA) is the only large, genre-balanced corpus of American English. COCA is probably the **most widely-used corpus of English**, and it is related to many other **corpora of English** that we have created, which offer unparalleled insight into **variation in English**.

The corpus contains more than **one billion words** of text (25+ million words each year 1990-2019) from eight genres: spoken, fiction, popular magazines, newspapers, academic texts, and (with the **update in March 2020**): TV and Movies subtitles, blogs, and other web pages.

Click on any of the links in the search form to the left for context-sensitive help, and to see the range of queries the corpus offers.

There are three main ways to search the corpus:

First, you can **browse a frequency list** of the top 60,000 words in the corpus, including searches by word form, part of speech, ranges in the 60,000 word list, and even by pronunciation. This should be particularly useful for language learners and teachers.

Second, you can **search by individual word**, and see collocates, topics, clusters, websites, concordance lines, and related words for each of these words. Note that some of these searches are unique to COCA and iWeb.

Third, you can **search for phrases and strings**. And because the corpus is optimized for speed, searches for substrings (*ism, un*able) and phrases are very fast, e.g.: **got VERB-ed**, **BUY * ADJ NOUN**, "gorgeous" **NOUN --** and even high frequency phrases like: **from ADJ to ADJ**, phrasal verbs, or **NOUN NOUN**.

You might pay special attention to the **comparisons between genres and years** and **virtual corpora**, which allow you to create personalized collections of texts related to a particular area of interest.

3. Searching for a word in coca corpus.

Here the word we had searched is petrol

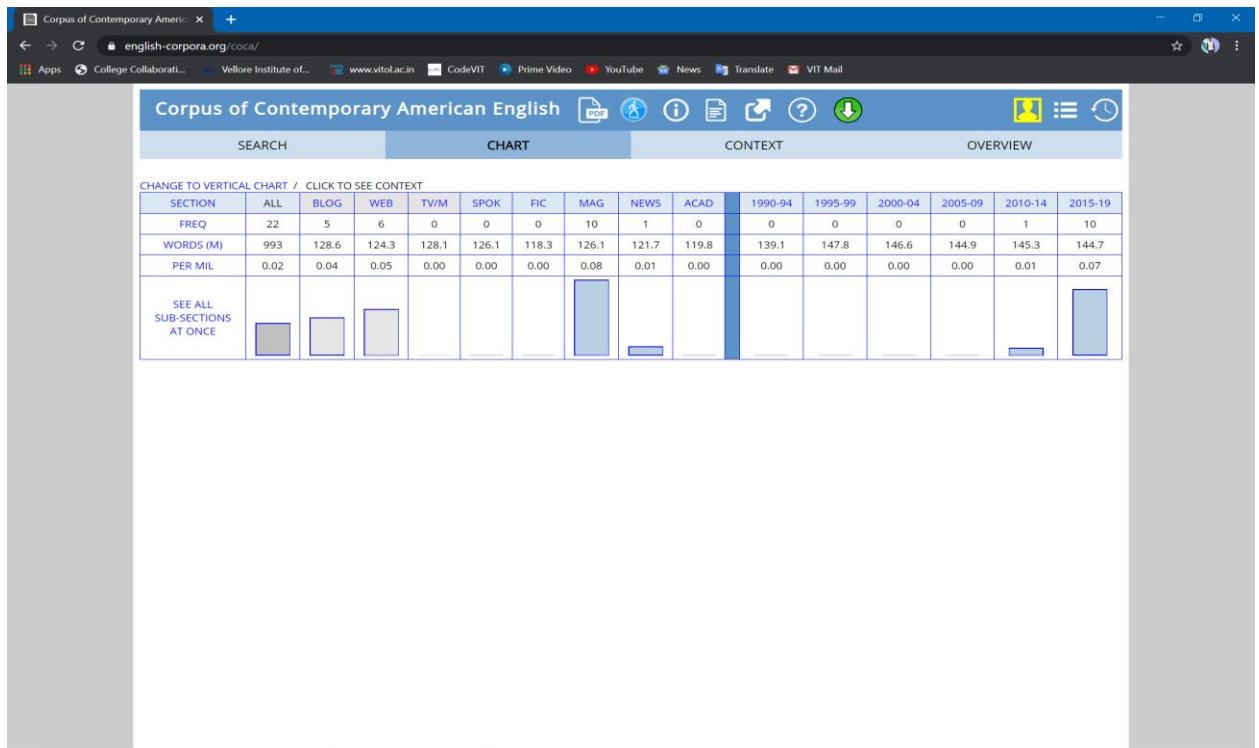
The screenshot shows the COCA website interface. The search bar at the top contains the word 'petrol'. Below the search bar, the 'CONTEXT' tab is selected, showing a frequency of 934. The 'OVERVIEW' tab is also visible. The search results are displayed in a table with columns for '1' and '2'. The word 'petrol' is highlighted in the first column.

4. Now we see the context that gives information about the word from where the source is and what is the context is

The screenshot shows the COCA website interface with the 'CONTEXT' tab selected. The search results are displayed in a table with columns for '1' and '2'. The word 'petrol' is highlighted in the first column. The context is provided in the second column, showing various sentences where the word 'petrol' is used. The table includes a 'SHOW DUPLICATES' button and a 'CLICK FOR MORE CONTEXT' button.

1	2
1	any work in the field. I am currently working as a cashier in a petrol station in the United Kingdom but use my spare time to study truth and run
2	deny him access to his children if there are any.... or just pour petrol over him and torch him after splattering his balls with a hammer and hacking his
3	without modification, requires no additives, and can use the existing distribution system of petrol stations. # The claims sound even bolder when you lea
4	- # Eager to get out of its way, I turn right for the petrol station and am irritated when it does likewise, driving close enough to pluck the
5	enough to pluck the GB sticker from my bumper. # I turn into the petrol station, hoping to lose the screaming blue vehicle. But no, it follows
6	in a while I would bump into him in Budgens supermarket, or at the petrol station, or at Glyndebourne. And as luck would have it I had the
7	turn from there. # My guess is the real motivation for electric over a petrol engine is packaging that helped it break the record. Electric motors are extrao
8	lot of people in the UK cycle or walk on a regular basis, and petrol runs the equivalent of \$8 to \$9 per gallon (which still doesn't seem
9	taken the long view. A generator is nice, until you run out of petrol. I'm fortunate enough to live in a very, very secluded area on
10	each coach had to stop a few yards down the road to fill up with petrol, prompting the logical question from one of our number: "Why couldn't
11	: " Why couldn't they have filled up before setting off, when the petrol station was only a few yards from the center? " As with so many
12	that there's the substantial cost of insurance, as well as topping up the petrol. Moreover, living in Manchester means a high risk of the car either being
13	toasts of our blood with the same politicians who wash away our blood stains with petrol and dollars. All theocratic poles in Yemen are growing in the o
14	in use to reprocess the actual Used Four tires directly into things like energy resource petrol for certain vehicles, herbal oils, char (also referred to as gocc
15	any work in the field. I am currently working as a cashier in a petrol station in the United Kingdom but use my spare time to study truth and run
16	motorcycles shuffling in and out of the fast food chain premise. There was a petrol station beside the restaurant, also having the sight of vehicles shufflin
17	3215602 # Full-size electric cars cost significantly more than their petrol and diesel equivalents. Attempts to bring the costs down fall into two camps:
18	More promising is the work at Middle Tennessee State University, which has converted a petrol car into a hybrid by adding a DC electric motor to each re
19	a longer period, making the financial commitment more akin to buying fuel for a petrol car: pay-as-you-go, not pay-everything-at-once. # Renault is the or
20	# Biostatistician and R user Matt Cooper noticed recently that the price he pays for petrol (gasoline) at the pump in Perth, Australia was about the same a
21	. Nonetheless, inflation has marched on over the years, so does that mean petrol is effectively cheaper now than it used to be? And how does the price
22	(in constant 2001 dollars), you see a much shallower trend. And petrol prices four years ago (just before the correction) were actually more expensive tha
23	drive, up and down steep hills. It chews up quite a bit of petrol BUT the last part of the drive is very pretty. # Once we got
24	in short term gains. On Wall Street, the pharmaceutical industry is third after petrol and banking, and each year it increases by 20%. It's more profitable
25	new Mercedes A-Class can choose between 6 engines, spread over three diesel and three petrol: the first 109, 136 and 170 hp of power, and the gasolinee
26	east in the past? do you know that libya is the 3rd country in petrol in Africa.... do you see how poor the people are????
27	.The government subsidizes the price of cars enabling everyone to own a vehicle. Petrol and bread cost a penny while no taxes are imposed on those wh
28	The fuel shortage has worsened, as evidenced by the queue of cars at petrol stations. # Petrol in the middle east is not as expensive as it once was. No need

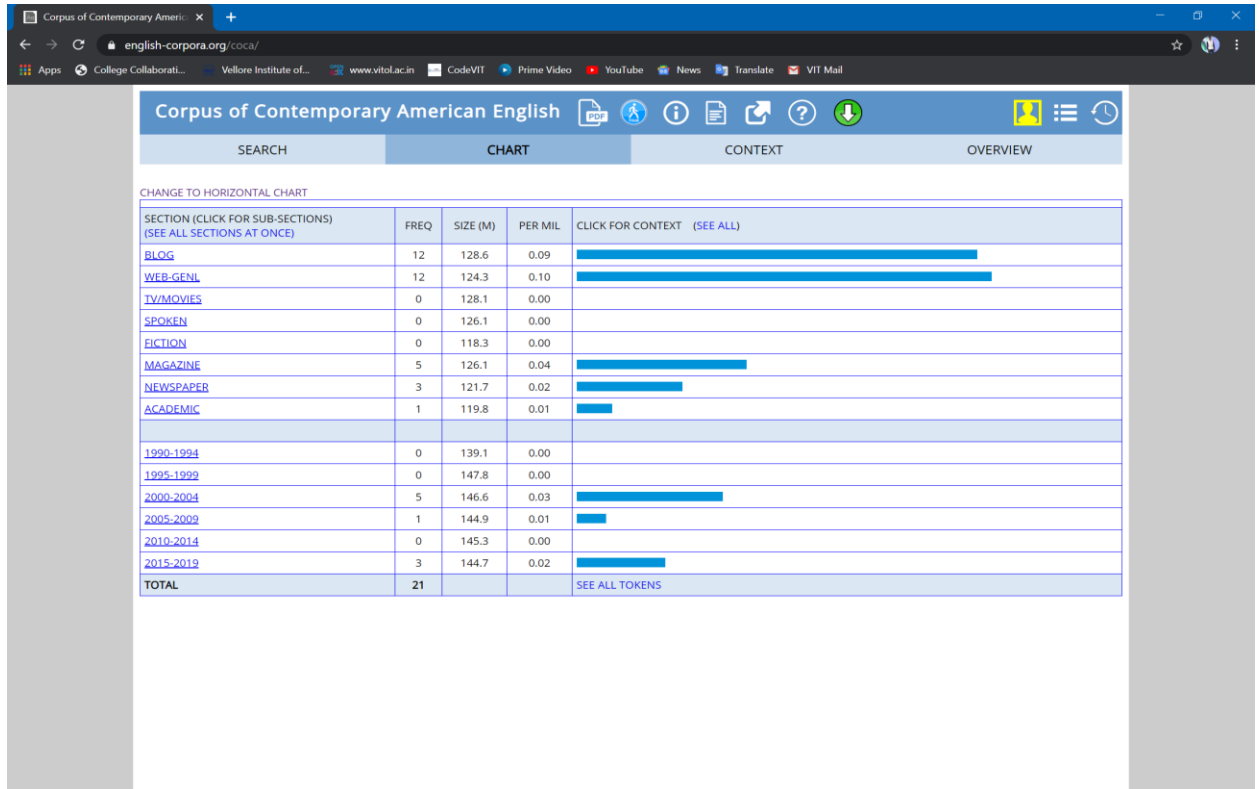
5. Here is the frequency distribution of the word that has occurred when, where and its frequency occurrence in the specific years.



6. Here we are searching for the word Niranjana

The screenshot shows the COCA interface with the 'SEARCH' tab selected. The search results for the word 'Niranjana' are displayed. The interface includes a search bar with the word 'Niranjana' entered, a 'See frequency by section' button, and a 'Reset' button. The results are shown in a table with columns for 'List', 'Chart', and 'Word Browse'. The 'List' column shows the word 'Niranjana' and its frequency across different sections. The 'Chart' column shows a bar chart of the frequency distribution. The 'Word Browse' column shows the word 'Niranjana' and its frequency across different sections. The interface also includes a 'NOT LOGGED IN' status and a 'CHART display' section with instructions on how to use the chart.

7. Here we can see the frequency distribution of the above word



8. The context and the source that the work occurred

SECTION: NO LIMITS (33) (SHUFFLE)

CLICK FOR MORE CONTEXT				SHOW DUPLICATES
1	2019	MAG	Medical Xpress	A B C and issues of data consent # by Jessica Bassano, The Lead # Associate Professor Niranjan Bidargaddi. Credit: Flinders University # The MINDick app was deve
2	2019	MAG	Medical Xpress	A B C pattern can trigger professional mental health intervention when necessary. # MINDick co-creator Associate Professor Niranjan Bidargaddi said the ability to
3	2019	NEWS	USA TODAY	A B C chariots and pulled to an open square after which rituals and festivities are performed. Niranjan Shrestha, AP # Indian election officials and paramilitary soldie
4	2012	WEB	yogajournal.com	A B C feel amazing stillness and can still feel the vibrations around and inside you. # Niranjan Prabhu # Karim: Its Om Tat Sat which mean Om That is the Truth
5	2012	WEB	aidnycforum.org	A B C brick kiln owner in J & K; employed to keep her bonded labourers from leaving. Niranjan Sahu, a survivor, tells his story # Ajachi Chakrabarti New Delhi # Photo
6	2012	WEB	aidnycforum.org	A B C tells his story # Ajachi Chakrabarti New Delhi # Photo: Ankit Agrawal # When Niranjan Sahu, 34, a labourer at a brick kiln in Nawanshahr in Jammu and
7	2012	WEB	aidnycforum.org	A B C years working at different kilns in the state. Instead of being granted permission, Niranjan was told that he owed Singh Rs 3 lakh and was beaten up. Scared
8	2012	WEB	aidnycforum.org	A B C a loan he had never taken, and of being held a prisoner indefinitely. Niranjan ran away, leaving behind his four-month pregnant wife and 18-month-old son at
9	2012	WEB	aidnycforum.org	A B C at the kiln, in order to get help to rescue his family. # Niranjan made his way to another kiln in Budgam, near Srinagar, where his father
10	2012	WEB	aidnycforum.org	A B C Budgam labourers and their families -- 51 people in total -- added their names to Niranjan 's petition and pooled in money to pay for his ticket to Raipur. Niran
11	2012	WEB	aidnycforum.org	A B C Niranjan's petition and pooled in money to pay for his ticket to Raipur. Niranjan , who left Budgam under the pretext of selling some land to pay off his
12	2012	WEB	aidnycforum.org	A B C debt, appealed to Chhattisgarh-based social activist Pyare Lal Diwakar for help. Diwakar took Niranjan to Delhi to Bandhua Mukti Morcha (BMM), an organisat
13	2012	WEB	aidnycforum.org	A B C them at least before they killed us all." # Meanwhile in Nawanshahr, Niranjan 's wife was facing Nayak's wrath. Days after Niranjan's escape, Nayak
14	2012	WEB	aidnycforum.org	A B C Meanwhile in Nawanshahr, Niranjan's wife was facing Nayak's wrath. Days after Niranjan 's escape, Nayak told her that her husband had left her and that she
15	2012	WEB	aidnycforum.org	A B C marry Manoj Kumar Chauhan, one of the kiln's enforcers who had beaten up Niranjan . When she refused, she was sent to Chauhan's hut where he raped
16	2012	BLOG	lostremote.com	A B C mobile gaming, immersive worlds, real-life brands, and social networking. Keynote speaker Niranjan Nagar, Head of Games & Apps at Facebook, will show you
21	2012	BLOG	...change.wordpress.com	A B C . His uncle, Rajni Ajmera, has been the president of MCHI. Niranjan Hiranandani, too, has held the post of president of MCHI in the past
22	2012	BLOG	...change.wordpress.com	A B C thus the homebuyer automatically becomes aware of the carpet area, " he said. Niranjan Hiranandani of the Hiranandani Group said: " All our deals on paper
23	2012	BLOG	mediabistro.com	A B C mobile gaming, immersive worlds, real-life brands, and social networking. Keynote speaker Niranjan Nagar, Head of Games & Apps at Facebook, will show you
26	2012	BLOG	fora.tv	A B C Minister of Rural Development for India, joins Dr. Abhijit Banerjee of MIT and Niranjan Rajadhyaksha of Mint. * Rs. 250 est. Bio Abhijit Vinayak Banerjee
27	2012	BLOG	fora.tv	A B C documentary film, " The Name of the Disease, " in 2006. Niranjan Rajadhyaksha Jairam Ramesh # Jairam Ramesh is an Indian economist and politician belongi
28	2008	ACAD	EarNoseThroat	A B C . Laryngoscope 2000; 110(11):1850-6. # -- (n7.) Lunsford LD, Niranjan A, Flickinger JC, et al. Radiosurgery of vestibular schwannomas: Summary of
29	2001	MAG	People	A B C , 51, and his two siblings, Princess Shruti, 24, and Prince Niranjan , 21. He also fatally wounded an uncle and killed two aunts and a
30	2001	MAG	People	A B C with their three children (from left) Princess Shruti, Prince Dipendra and Prince Niranjan . // The dispute apparently came to a head shortly before the massacr
31	2001	MAG	People	A B C catch was that, if he did so, they would anoint his younger brother Niranjan as crown prince and begin grooming him to take over as king. A royal
32	2000	NEWS	NYTimes	A B C 5084 # " IVE always had this house in my head, " said Niranjan Rao, an Indian-born surgeon. With its three-story turret of rustic brown stone and
33	2000	NEWS	NYTimes	A B C " # Farther out still, in a more rural landscape, live Meena and Niranjan Rao, who met when they were students at college in India. A native