



فرکانس استفاده از يك کلمه

به تازگی آزمایشگاه گوگل با همکاری 2 جامعه‌شناس، محصول جالبی بیرون داده است. گوگل 5.2 میلیون کتاب را دیجیتالی کرده است. این تعداد کتاب 500 میلیارد کلمه دارند. کتاب‌های دیجیتالی شده بین سال‌های 1500 تا 2008 نوشته شده‌اند.

جام جم آنلاین: به تازگی آزمایشگاه گوگل با همکاری 2 جامعه‌شناس، محصول جالبی بیرون داده است. گوگل 5.2 میلیون کتاب را دیجیتالی کرده است. این تعداد کتاب 500 میلیارد کلمه دارند. کتاب‌های دیجیتالی شده بین سال‌های 1500 تا 2008 نوشته شده‌اند.

محصول جدید گوگل مشخص می‌کند که در هر سال، چه میزان از يك کلمه مشخص استفاده شده است! به این ترتیب معلوم می‌شود که در يك بازه زمانی چندقرنی فرکانس استفاده از يك کلمه چقدر فرق کرده است و از این طریق، روزهایی برای بررسی فرهنگ و جامعه باز می‌شود.

من در نخستین آزمایش، کلمه war را در دو قرن اخیر، جستجو کردم:

مشخص است که در سال‌های دور و بر جنگ‌های جهانی اول و دوم، فرکانس استفاده از این کلمه بیشتر شده است.

حالا تست دیگری می‌کنیم و مردان و زنان را در کتاب‌ها جستجو می‌کنیم:

این جستجو هم جالب است و نشان می‌دهد که قبل از دهه 70 میلادی، میزان استفاده از کلمه ««زنانه» به شکل معناداری کمتر از ««مردانه» بود.

حالا بیابید، ««ایران» را جستجو کنیم:

تفسیر این نمودار، بسیار دشوارتر است. پیداست که فرکانس استفاده از واژه ایران بعد از انقلاب و جنگ ایران و عراق و در سال‌های اخیر به دلایل سیاسی باید بیشتر شده باشد.

ولی در این لحظه نمی‌توانم ابراز نظر کنم که چرا در سال‌های دور و بر 1800 یا 1750، در نمودار قله‌هایی ظاهر شده است.

دست آخر ««کامپیوتر» را جستجو می‌کنیم:

معلوم می‌شود که از 1945 به بعد، سر و کله این کلمه پیدا شده است، از سال 1977، شتاب استفاده از آن تند شده است، از سال 2000 به بعد هم، احتمالاً به خاطر رواج کلمات هم‌ارزی مثل لپ‌تاپ، میزان استفاده از آن کمتر شده است.

وبلاگ يك پزشك