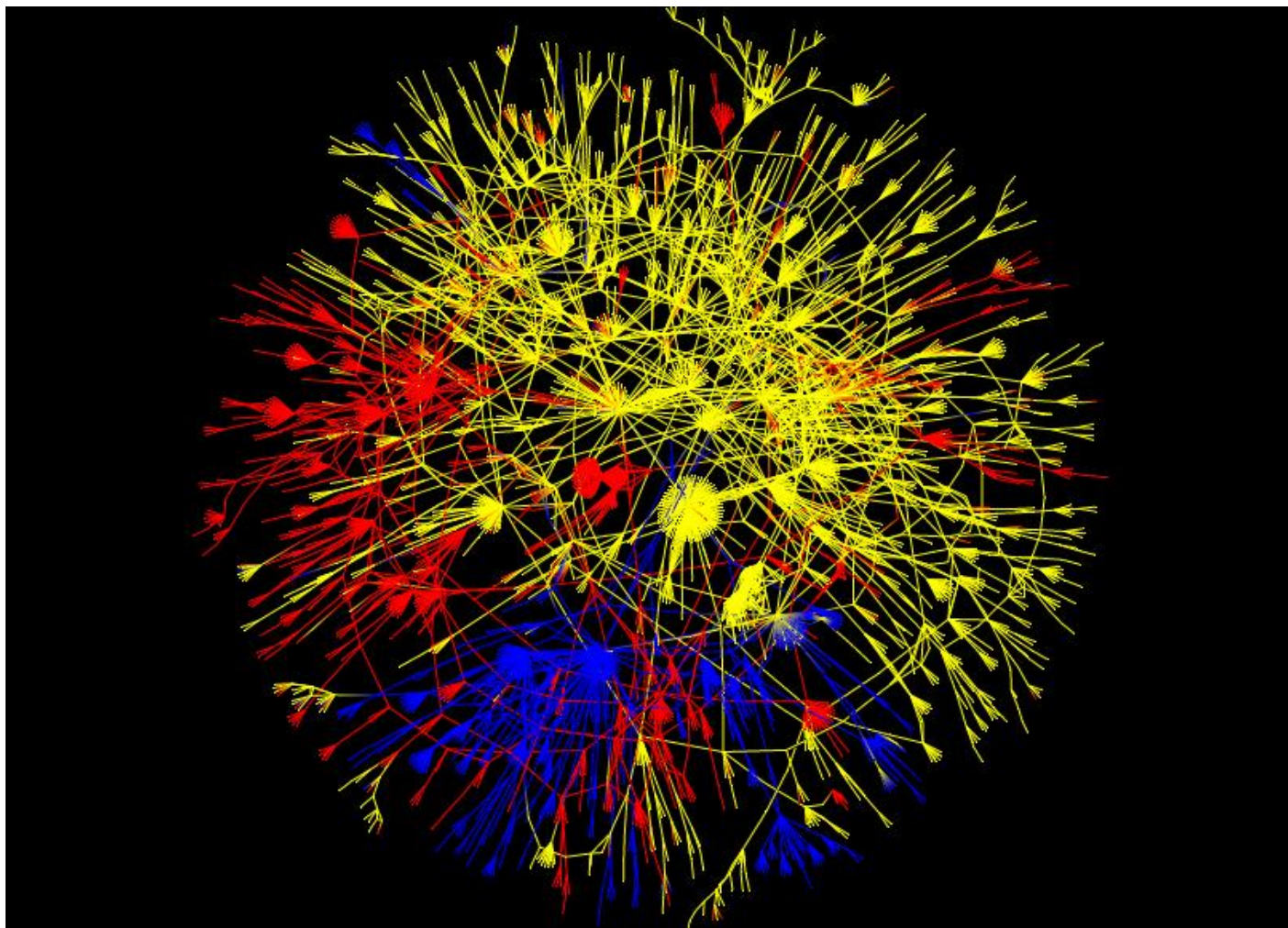


آیا حباب‌های زبان آنلاین، شکل دیگر مرزها در اینترنت هستند؟

اسکات هال به تأثیر زبان در جست‌وجو و به اشتراک‌گذاری اطلاعات در فضای گسترده اینترنت می‌پردازد.



اولین اصل «بحث آزادی بیان» به این حق می‌پردازد که "... بتوانیم فارغ از مرزها، اطلاعات و نظرات گوناگون را دریافت کرده و به اشتراک بگذاریم". یکی از واضح‌ترین مرزهای آنلاین که کمتر مورد مطالعه قرار گرفته، زبان است. پروژه «بحث آزادی بیان» بر این امر واقف بوده و در همین راستا برنامه مؤثری جهت **ترجمه محتویاتش به 13 زبان دیگر** در دستور کار داشته است.

با توجه به این امر، تأثیر زبان در جست‌وجو و هم‌خوان کردن اطلاعات در فضای گسترده اینترنت چیست؟ تحقیقات موجود اساساً به این سؤال نپرداخته‌اند و در این نوشته هم نمی‌توان به طور کامل آن را بررسی کرد. امروزه موتورهای جست‌وجوگر اینترنت، صفحاتی با محتویات گوناگون و در زبان‌های مختلف فراهم می‌کنند. وقتی برای عبارتی، به جست‌وجوی عکس می‌پردازید، این موتورها سعی می‌کنند تا کلمه مورد جست‌وجو را با متنی که در کنار عکس‌های موجود در اینترنت آمده، یا نام فایل خود عکس‌ها یا کلمات لینک شده به آن عکس‌ها انطباق دهند. از یک سو، ما انتظار عکس‌هایی مشابه در زبان‌های گوناگون داریم؛ چراکه عکس‌ها غالباً به طور مستقل و بدون توضیح و حاشیه‌نگاری تشخیص داده می‌شوند. اما هنوز عکس‌ها در تنظیمات فرهنگی-زبانی جداگانه‌ای آپلود و بارگذاری می‌شوند. هرچند که گوگل موتور جست‌وجوی غالب در همه بازارها نیست (یاهوی ژاپن، یاندکس و بایدو به ترتیب سهم بیشتری در بازارهای ژاپن، روسیه و چین دارند) اما هنوز پیش‌تاز جست‌وجوی جهانی به شمار می‌رود، این موتور جست‌وجوگر مطالب و محتویات

گسترده‌ای را فهرست می‌کند و احتمالاً الگوریتم‌های مشابهی، اگر نگوئیم که همان الگوریتم را برای جست‌وجوی مطالب در زبان‌های گوناگون به کار می‌بندد. این امر باعث می‌شود که نتایج حاصل از جست‌وجو در مکان‌های گوناگون؛ بسیار جالب از آب در آید.

تصاویر 1 و 2 نتایج حاصل از جست‌وجوی عکس گوگل برای میدان تیانانمن به زبان انگلیسی و چینی را نشان می‌دهد. تمامی عملیات جست‌وجو با همان عبارت و همان کامپیوتر کمی بعدتر در `com.google` بریتانیا انجام گرفت؛ اما نتایج گاهی بدون قصد خاصی و گاهی بنا بر برخی ملاحظات به طور عجیبی متفاوت بوده است. به عنوان مثال، نتایج جست‌وجوی عبارت «میدان تیانانمن» ناهمخوانی مشهودی در عکس‌های فهرست شده در زبان‌های انگلیسی و چینی پیرامون تظاهرات 1989 را نشان می‌دهد.

«تظاهرات 1989 در میدان تیانانمن»

2022:20222022 20222022 2022 2022 2022 20222022 «202220222022 20222022» 2022 2022

[مطالعات سیستماتیک بیشتر در خصوص تفاوت‌های زبان‌های گوناگون دانش‌نامه ویکی‌پدیا](#) نشان می‌دهد که “به‌طور شگفت‌انگیزی محتویات اندکی در زبان‌های گوناگون ویکی‌پدیا با هم منطبق هستند.” (گرگل و هجت، 2010) جالب‌تر اینکه حتی در نسخه انگلیسی آن، که با فاصله زیادی بزرگ‌ترین نسخه این دانش‌نامه محسوب می‌شود، به‌طور میانگین تنها شامل 60% از مفاهیمی است که در دیگر زبان‌های ویکی‌پدیا آمده. (بالاترین میزان میان زبان انگلیسی و عبری است با 75%). به علاوه، ویکی‌پدیای انگلیسی فقط شامل نیمی از مسایل کار شده در زبان آلمانی، که از لحاظ گستردگی در رتبه قرار دارد، می‌باشد و این در حالی است که زبان آلمانی تنها حاوی 16% از مقاله‌های انگلیسی است. البته، حتی جایی که “دو زبان گوناگون به مفهوم واحدی می‌پردازند (به‌طور کاملاً مشخصی) هر دو به‌طور متفاوتی آن را توصیف می‌کنند.” موضوعی که پیش‌تر [گرگل و هجت \(2010\) در مقاله‌ی خود بررسی می‌کنند و ابزارهای جدیدی چون امنیپدیا \(Omnipedia\)](#) که به کاربران این اجازه را می‌دهد تا تفاوت و ویرایش‌های گوناگون زبانی را کشف کنند، به آن می‌پردازند.

هر چند اوضاع آن قدرها هم بد نیست، تعدادی از پایگاه‌های اینترنتی حقیقتاً به جایگاه‌های جهانی رسیده‌اند (به طور مثال فیسبوک، تویتر و یوتیوب)؛ این پایگاه‌ها که در درجه‌ی اول ارتباط در آنها مبتنی بر زبان است، ظرفیتهایی برای گسترش اطلاعات ورای مرزها و با سرعتی بی‌سابقه فراهم کرده‌اند. [ت تحقیقات من بر روی بلاگ‌هایی که به زلزله 2010 هایتی پرداخته بودند](#) (شکل شماره 3) و [اشتراک‌گذاری لینک‌ها در ویکی پدیا و تویتر پس از سونامی سال 2011 ژاپن](#) و همچنین مطالعه [آیرن التاپیرامون تویتر](#) نشان می‌دهد که چگونه در برخی موارد اطلاعات در زبان‌های گوناگون پخش می‌شود و کاربران چندزبانه چون «پل‌های گره‌ای» میان گروه‌های زبانی عمل می‌کنند و موجب جریان اطلاعات میان این زبان‌ها می‌گردند.

برقراری و توسعه جریان اطلاعات میان زبان‌های گوناگون، هر دو بعد تکنیکی و اجتماعی را در بر دارد. [ماشین‌ها و نرم‌افزارهای ترجمه بدون خطا نیستند](#)، اما به هر حال تا حد زیادی ما را در جریان مطالب دیگر زبان‌ها می‌گذارند، و تحقیقات بیشتر و منابع نوین داده‌های آموزشی می‌توانند به بهبود مداوم سطح کار این نرم‌افزارهای ترجمه بیانجامند. علاوه بر این، مطالعه پیرامون تدابیری که بتواند در کمک کردن به کاربران جهت کشف و دسترسی به اطلاعات در دیگر زبان‌ها نقش داشته باشد ([مانند تحقیق خود من](#)) مورد نیاز است و شرکت‌های رسانه‌ای اجتماعی نیز باید از نتایج حاصل از این مطالعات و دیگر حوزه‌های علوم اجتماعی در تأسیس پایگاه‌های جدید اینترنتی خود بهره گیرند. و در نهایت، ابزارهای جذاب و جدیدی می‌توانند مهارت و توانایی کامپیوترها و کاربران در گذشتن از این مرزهای زبانی را به کار گیرند و تقویت کنند. [دولینگو](#) و [مونوترانس 2](#)، مثال‌هایی از این دست هستند که کاربران یک‌زبانه را قادر می‌سازند که مطالب را ترجمه کنند، در مورد [دولینگو](#) حتی کاربر می‌تواند در همان حین، زبان جدیدی هم بیاموزد. گذشته از این، همیشه جایگاهی برای [ترجمه انسانی](#) وجود دارد و [سازمان‌های فعال در امر رسانه](#) هم می‌توانند در شناسایی و تأیید اطلاعات مربوط به رویدادهای مهم به زبان‌های دیگر نقش ایفا کنند.

[اسکات هال دستیار پژوهش و دانشجو دکتری](#)

منتشر شده در جولای 25، 2012