

## ماده تاریک ویکی پدیا

### مقاله‌های یتیم ویکی پدیا قابل بازیابی هستند!

باعث مشکلات زیادی می‌شود؛ چون ممکن است ویراستار، کارشناس موضوع‌ها و مقالات دیگر نباشد.

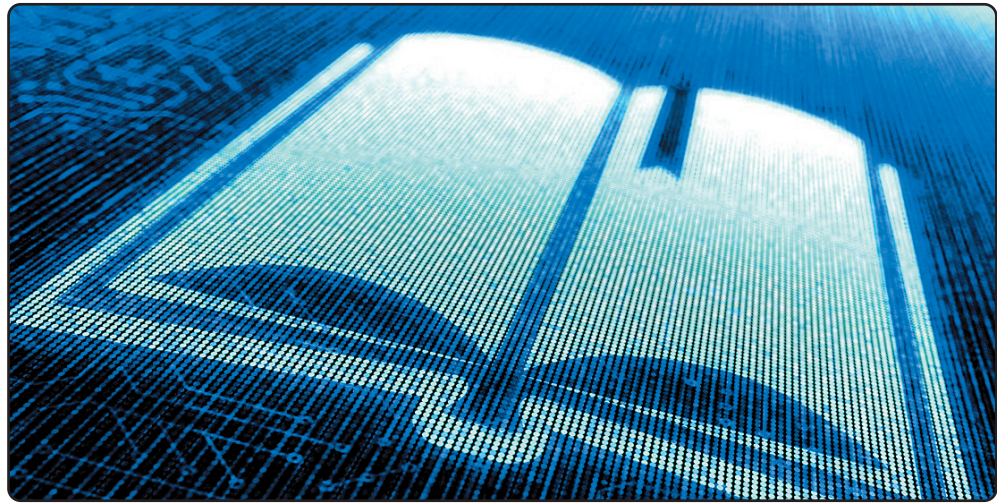
مطالعه سیستماتیک روی این مقالات همچنین نشان داده است که اختلاف زیادی بین نسخه‌های ویکی پدیا که به زبان‌های مختلف هستند وجود دارد. در ویکی پدیاهایی که به بیش از ۱۰۰ زبان مختلف در دسترس هستند، تعداد مقالات یتیم بیش از ۳۰ درصد است که ویکی پدیا زبان عربی مصری درصد بسیار بالایی از این مقالات را دارد (۷۸ درصد) و ویکی پدیای زبان ویتنامی دارای ۵۰ درصد مقالات یتیم است. هر دوی آن‌ها در زمره ۲۰ نسخه زبانی ویکی پدیا که بیشترین حجم را دارند هستند. این امر از عدم توانایی ویراستارها در بعضی از زبان‌ها حکایت دارد و نشان دهنده نیاز به ارتقاء ابزارهای فعلی از قبیل «فاین‌لینک» است که در این کار به ویراستارها کمک می‌کنند.

مطلب جالب توجه این است که یک مقاله یتیم در یک زبان همیشه در زبان‌های دیگر به شکل مقاله یتیم نیست. بنابراین برای شناسایی مقاله‌هایی که از آن‌ها می‌توان به وسیله ترجمه به مقاله‌های یتیم لینک زد رویکرد دیگری لازم است.

اگر یک مقاله در زبان دیگری یک مقاله یتیم نیست، معنایش این است که ویراستارهای زبان دیگر توانسته‌اند مقاله‌های دیگری را پیدا کنند که به مقاله یتیم لینک داده می‌شوند. این مشکل با انتقال لینک از زبان‌های دیگر به زبانی که مقاله مورد نظر در آن یک مقاله یتیم است حل شده است. با این راهکار لینک‌هایی برای بیش از ۶۳ درصد از مقاله‌های یتیم ارائه شده‌اند.

این رویکرد می‌تواند در نقش یک ابزار، کیفیت تجربه خوانندگانی که از ویکی پدیا کسب اطلاعات می‌کنند را بالا ببرد. هوش مصنوعی از دو ناحیه می‌تواند به این امر کمک کند. نخستین آن‌ها این است که پژوهشگران روی شبکه‌های عصبی گراف کار می‌کنند تا پیشنهاداتی برای لینک‌ها سازماندهی کنند. دوم این که ابزار دیگری را طراحی می‌کنند که می‌تواند ویراستارها را راهنمایی کند در کجای یک صفحه متنی باید مفاهیم جدید را بیافزایند. در مرحله بعد از طریق این مفاهیم جدید و با استفاده از هوش مصنوعی زبانی متن‌های دیگری به خوانندگان پیشنهاد می‌شوند.

مهم‌تر این که همواره ویراستارهای داوطلبی هستند که کاری را که هوش مصنوعی انجام داده را با بازبینی و ویرایش بهتر می‌کنند. ویراستارها به مخاطبین ویکی پدیا در سراسر دنیا خدمت‌رسانی می‌کنند اما تعداد آن‌ها در زبان‌هایی که گستره کوچک‌تری دارند کافی نیست. از این رو، حمایت از ویراستارها اهمیت زیادی دارد چون نگارش و اصلاح مقالات کار سختی است.



پنهان نیز قابل مشاهده شوند و با پیوند خوردن با مقالات دیگر در دسترس مخاطبین قرار گیرند. برای این منظور، نخستین مطالعه سیستماتیک روی مقالات یتیم در تمام ۳۱۹ نسخه زبانی ویکی پدیا انجام شده است. به‌طور کلی، تعداد Pageview (تعداد مشاهدات صفحه) های مقالات غیر یتیم (قابل لینک شدن) دو برابر تعداد پیچ و یوهای مقالات یتیم (غیر قابل لینک شدن) است. بنابراین یک رابطه علت و معلولی بین افزودن n-link به مقالات یتیم و افزایش تعداد پیچ و یوها تعریف شده است. مشاهده نشدن مقالات یتیم به شیوه‌های جستجو و مشاهده صفحات ویکی پدیا بستگی دارد. شیوه نخست از طریق موتور جستجو است که در آن یک صفحه خاص از ویکی پدیا به خواننده نمایش داده می‌شود. شیوه دوم زمانی است که خواننده از ویکی پدیا مانند یک دایر فال‌معرف استفاده می‌کند و از یک مقاله به مقاله دیگر کلیک می‌کند. شیوه سوم ترکیبی از دو شیوه نام برده است.

در هر سه مورد، ویراستار نه تنها باید لینک‌هایی به خارج از مقاله‌ای که ویرایش می‌کند اضافه کند بلکه همچنین لازم است از همه مقالات مرتبط که احتمال لینک خوردن به آن دارند اطلاع داشته باشد که این کار آسانی نیست. وقتی ویراستاری در حال ویرایش مقاله‌ای است که معلومات زیادی درباره‌اش دارد، می‌تواند لینک‌هایی به خارج از آن مقاله که آن را به مقالات دیگر پیوند می‌دهند اضافه کند. برعکس کردن جهت ارجاع لینک‌ها از لینک در حال ویرایش شدن به بیرون از آن

ویکی پدیا نیز مانند جاده‌ها، اینترنت، ترکیبات شیمیایی، ژن‌ها و هر شبکه دیگری قابلیت جهت‌دهی و ناوبری دارد. در نتیجه، می‌توان در این شبکه از نقطه‌ای به نقطه دیگر رفت. شبکه‌های اطلاعات در سلسله مراتب خاصی سازماندهی می‌شوند اما شاید تعجب کنیم چرا مقالاتی در ویکی پدیا هستند که دست هیچ خواننده‌ای به آن‌ها نمی‌رسد! محتوای کنونی و قابل دسترس ویکی پدیا که مشتمل بر ۶۰ میلیون مقاله در بیش از ۳۰۰ زبان مختلف است همچنان روبه افزایش است و هر ماه حدود ۲۰۰ هزار مقاله جدید به آن اضافه می‌شوند. خوانندگان مطالب ویکی پدیا در بیشتر مواقع اطلاعات تازه‌ای از آن به دست می‌آورند و با کلیک روی آبر پیوندها (هایپر لینکس) که یک مقاله را به مقاله‌ای دیگر مرتبط می‌کنند از موضوع مورد نظر خود مطالب بیشتری یاد می‌گیرند. شاید بسیاری از ما ندانیم که اگر چه ویکی پدیا بزرگ‌ترین پلتفرم برای دسترسی رایگان به اطلاعات آنلاین است اما حدود ۹ میلیون مقاله یا به عبارتی ۱۵ درصد از محتوای آن برای خوانندگانی که صفحات ویکی پدیا را در همه موضوع‌ها مرور می‌کنند نامرئی و غیر قابل رؤیت است. به‌طور معمول و در ترجمه تحت‌اللفظی به آن‌ها «مقالات یتیم» (orphan articles) می‌گویند؛ چرا که هیچ مقاله‌ای به آن‌ها لینک نمی‌شود، انگار که خانواده‌ای ندارند و تنها مانده‌اند. شاید بتوانیم به آن‌ها مقالات بی‌پیوند یا غیر قابل لینک شدن بگوییم. از این رو، پژوهشگران سوئیس ابزار ساختند تا این مقالات

## موتور جستجوی گوگل به هوش مصنوعی مجهز می‌شود!

شرکت گوگل در حال تکمیل یک ویژگی جدید به نام A.I. Overviews برای موتور جستجوی خود است. این ابزار توصیف مختصری از موضوعی که کاربر جستجو می‌کند به وجود می‌آورد و لینک‌هایی را برای مطالعه بیشتر درباره آن موضوع به کاربر ارائه می‌دهد. این ویژگی بخش‌هایی از اطلاعات را به شکل تجزیه شده نمایش می‌دهد تا کاربر سریع‌تر به پاسخ برسد. در حال حاضر، A.I. Overviews تمامی موضوعاتی که جستجو می‌شود نمایان نمی‌شود اما کاربرهایی که در ایالات متحده هستند از هفته سوم ماه می ۲۰۲۴ توانستند بالا آمدن این ویژگی را هنگام جستجوهای خود مشاهده کنند. شرکت گوگل پیش‌بینی می‌کند که این ویژگی تا پایان سال جاری میلادی در دسترس بیش از یک میلیارد کاربر در سراسر جهان قرار خواهد گرفت. تجهیز موتور جستجو به هوش مصنوعی اقدام هیجان‌انگیزی از جانب گوگل است اما از آنجا که هنوز به‌طور کامل آماده نشده است، عده‌ای با آن مخالف هستند و معتقدند که می‌تواند پیامدهایی را به دنبال داشته باشد. شرکت گوگل در عمل تمامی دنیا را به تست‌های بتا برای محصول خود تبدیل می‌کند. با توجه به این که جستجو عملی است که کاربرها روزانه از طریق آن از اینترنت استفاده می‌کنند، ارائه محصولی که هنوز آماده نیست می‌تواند به نوعی تحمیل کردن یا وادار کردن افراد از استفاده از آن باشد.

به علاوه، فناوری افزوده شده به موتور جستجو هوش مصنوعی از نوع مدل‌های زبانی است اما هنوز به‌طور قطع اطمینانی وجود ندارد که مدل‌های زبانی دروغ نگویند. این یکی از مواردی است که هنوز در ویژگی جدید گوگل آزمایش نشده است یا دست‌کم به‌طور کامل مورد بازبینی قرار نگرفته است. انتقاد دیگری که به این اقدام گوگل شده این است که موتور جستجوی مجهز به هوش مصنوعی اطلاعاتی از وبسایت‌های دیگر فراهم می‌آورد اما کاربرها را به آن سایت‌ها هدایت نمی‌کند. بنابراین کاربرها نمی‌توانند از سایت‌های منبع بازدید کنند.

